Relevé numérique d'art pariétal : définition d'une approche innovante combinant propriétés géométriques, visuelles et sémantiques au sein d'un environnement de réalité mixte

Violette Abergel

▶ **To cite this version:**

Violette Abergel. Relevé numérique d'art pariétal : définition d'une approche innovante combinant propriétés géométriques, visuelles et sémantiques au sein d'un environnement de réalité mixte. Synthèse d'image et réalité virtuelle [cs.GR]. HESAM Université, 2020. Français. NNT : 2020HESAE021 . tel-02950598

HAL Id: tel-02950598
https://theses.hal.science/tel-02950598

Submitted on 28 Sep 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

HESAM UNIVERSITÉ

Arts et Métiers Sciences et Technologies

ÉCOLE DOCTORALE SCIENCES DES MÉTIERS DE L'INGÉNIEUR
[UMR MAP CNRS/MC – Campus d'Aix-en-Provence]

THÈSE

présentée par : **Violette ABERGEL**

soutenue le : **01 juillet 2020**

pour obtenir le grade de : **Docteur d'HESAM Université**

préparée à : **École Nationale Supérieure d'Arts et Métiers**

Spécialité : **Conception**

Relevé numérique d'art pariétal : définition d'une approche innovante combinant propriétés géométriques, visuelles et sémantiques au sein d'un environnement de réalité mixte

THÈSE dirigée par :
[M. DE LUCA Livio]
[M. VERON Philippe]

et co-encadrée par :
[M. JACQUOT Kévin]

Jury

M. Simon RICHIR, Professeur, LAMPA, Arts et Métiers	Président
Mme Nadine COUTURE, Professeur, UMR 5800 LaBRI, ESTIA	Rapportrice
M. Laurent LESCOP, Professeur, UMR 1563 AAU, ENSA Nantes	Rapporteur
M. Oscar FUENTES, Ingénieur de recherche, Centre National de Préhistoire	Examinateur
M. El Mustapha MOUADDIB, Professeur, EA 4290 MIS, Université de Picardie	Examinateur
M. Livio DE LUCA, Directeur de recherche au CNRS, UMR 3495 MAP	Examinateur
M. Philippe VERON, Professeur, EA 7515 LISPEN, Arts et Métiers	Examinateur
M. Kévin JACQUOT, Maître de Conférence, UMR 3495 MAP, ENSA Lyon	Examinateur
Mme Geneviève PINÇON, Directice du CNP, Ministère de la Culture	Invitée

T H È S E

Remerciements

Je tiens tout d'abord à remercier Nadine COUTURE et Laurent LESCOP pour avoir accepté la charge de rapporteur et pour leurs remarques constructives, ainsi que les autres membres du jury Oscar FUENTES, El Mustapha MOUADDIB, Simon RICHIR, pour le temps qu'ils ont bien voulu consacrer à mon travail.

Je souhaite remercier chaleureusement les membres de mon équipe encadrante pour leur disponibilité et leur bienveillance. Je souhaite en particulier exprimer toute ma gratitude envers mon directeur de thèse Livio DE LUCA, pour la confiance qu'il a bien voulu m'accorder dès mon entrée au laboratoire MAP, pour ses nombreux conseils, sa disponibilité, mais aussi pour avoir toujours su trouver les mots pour me pousser à aller toujours plus loin. Je suis très fière d'avoir été ta doctorante, et de continuer à travailler à tes côtés. Je remercie Philippe VERON pour ses précieux conseils méthodologiques et pratiques, qui m'ont permis de démarrer sur de bons rails lors de ma première année, et de l'achever dans de bonnes conditions malgré le contexte difficile du confinement. Enfin, je ne saurais trop remercier Kévin JACQUOT pour ses conseils bien sûr, mais surtout pour sa présence amicale, et pour m'avoir supportée (dans tous les sens du terme !) au quotidien durant ces trois années de thèse.

Ce travail n'aurait pas été possible sans le soutien du Ministère de la Culture et de l'ENSAL, qui m'ont permis, grâce à une bourse doctorale, de me consacrer pleinement et sereinement au développement de cette thèse. Je remercie Hervé LEQUAY pour m'avoir initiée à la recherche durant mes études d'architecture, puis accueillie au MAP-ARIA. Ce travail n'aurait pas été ce qu'il est sans le soutien des membres du Centre National de Préhistoire, en particulier Geneviève PINÇON et Stéphane KONIK, qui m'ont ouverte au domaine de l'art pariétal. Je garderai d'excellents souvenirs de mes passages à Périgueux, en espérant qu'ils se reproduisent. Je tiens à adresser également mes remerciements à Frédéric MERIENNE pour sa participation enrichissante à mes comités de suivi et pour m'avoir aimablement ouvert les portes du master MTI3D.

Je tiens bien entendu à remercier tous mes collèges du MAP pour tous les bons moments passés ensemble : missions, séminaires, mais aussi sorties ski, randonnées, pique-niques, apéritifs, restaurants… La bonne ambiance permanente au laboratoire a largement contribué au plaisir pris à mener ce travail à bien, merci pour cela. La liste est longue, désordonnée et non-exhaustive mais essayons tout de même :

- Du côté de Lyon, merci à Renato (Nato), Kévin (je suis sympa, je tairais le surnom affectueux dont Renato t'a affublé), Alexandra (Tatiana), Dimitri (Mimi), Xavier (Smoumi), Églantine, Hervé, mais aussi les anciens : Sébastien (Gneeeb), Inès (Topo), Quentin (Piou-piou) et les autres ;

- Du côté de Marseille : Merci à Livio, Menehould (la maman du GAMSAU quoi que tu en dises !), Pascal (ou plutôt « Pascaaaaal ?? »), Anas (il faut qu'on retourne à SF !), Anthony (Antho-la-scoumoune, j'attends ton recueil en 5 tomes), Laurent (Lolo), Ariane, Noémie, Jean-Yves, Ivona, Philippe, Antoine, Jean-Marc, Odile, et en particulier Adeline et Tommy, mes prédécesseurs, pour leurs conseils avisés et leur expérience en matière de démarches administratives (outch !). Je souhaite également une bonne continuation à Valeria et probablement Roxane, qui prennent la relève !

Je remercie sincèrement mes amis pour leur présence réconfortante, Chloé mon amie de toujours, (tu es la prochaine !), Clémence (ma blonde !), Delphine (Fifine), Alexandra, Nolwenn, Bertrand et toute la troupe des Couzonnais, avec une mention spéciale pour les Barranco pour leur club de soutien et leur dynamisme extraordinaire, les Tata Zéro/Pasta Zéro (ma Gigi !) pour les bons moments des vendredis soir, et tous ceux qui me sont chers et que je n'aurais malencontreusement pas cité.

Merci du fond du cœur aux membres de ma famille pour leur amour inconditionnel, en particulier mes parents, Fabien, Jérôme, Julie (Van Crrr), Matthieu (Spouty), Rémy (Mimi), Yui, mes adorables neveux, ma belle-mère Dorina, mais aussi mes oncles et tantes, que je ne vois pas assez souvent mais qui sont toujours à mes côtés.

Enfin, les mots me manquent pour remercier Renato autant qu'il le mériterait et exprimer combien sa présence m'est précieuse. Merci, merci, merci.

V.

Sommaire

Introduction **1**

Contexte général 1

Cas d'étude 2

Objectifs 3

Portées, limites, apports principaux 4

Structure du document 5

Note concernant l'objet d'expérimentation 6

Partie I.

Chapitre 1. Le relevé **8**

1.1. Le relevé du bâti 8

1.1.1. Historique : de la représentation d'architecture au relevé du bâti 8

1.1.2. Les moments du relevé : Décrire, Figurer, Représenter 14

1.2. Le relevé d'art pariétal 15

1.2.1. Historique : vers une recherche d'exactitude 16

1.2.2. La complexité du relevé d'art pariétal 22

1.3. Bilan : Les exigences des méthodes de relevé 31

Chapitre 2. Le relevé numérique **34**

2.1. Vers un changement de paradigme 34

2.2. Vers la numérisation massive de l'existant 36

2.2.1. Outils et méthodes d'acquisition 2D 36

2.2.2. Outils et méthodes d'acquisition 3D 41

2.2.3. Bilan des méthodes d'acquisition de données numériques 46

2.3. Extraction d'informations géométriques et visuelles 46

2.3.1. Traitement des données 2D 47

2.3.2. Traitement des données 3D 50

2.3.3. Bilan des méthodes d'extraction d'informations 59

2.4. Enrichissement sémantique de données numériques 60

2.4.1. Approches 2D 61

2.4.2. Approches 3D 64

2.4.3. Approches hybrides 2D/3D 66

2.4.4. Bilan des méthodes d'enrichissement sémantique 70

2.5. Bilan du relevé numérique : dispersion des données et rupture réel/virtuel 71

Chapitre 3. Les systèmes d'information et l'hybridation réel/virtuel dans le domaine du patrimoine 74

3.1. Systèmes d'information patrimoniaux ... 74
3.1.1. ArTaPOC ... 74
3.1.2. Approches basées sur l'utilisation de SIG ... 75
3.1.3. CHER-Ob ... 77
3.1.4. Agata ... 78
3.1.5. Netptune IS ... 79
3.1.6. NUBES ... 80
3.1.7. Aïoli ... 81
3.2. Bilan sur les systèmes d'informations patrimoniaux ... 83
3.3. Hybridation réel / virtuel ... 85
3.3.1. L'expérience de Brunelleschi ... 85
3.3.2. Le continuum Réel –Virtuel ... 86
3.3.3. Définitions ... 86
3.3.4. Technologies ... 88
3.3.5. Applications au patrimoine ... 92
3.3.6. Le suivi et le recalage ... 95
3.3.7. Bilan des méthodes de suivi ... 104
3.3.8. Les composantes applicatives d'un système de réalité augmentée ... 105
3.4. Bilan des méthodes d'hybridation ... 107

Chapitre 4. Discussion et approche proposée pour la définition d'un environnement numérique multimodal dédié au relevé d'art pariétal ... 109

4.1. Un moment privilégié de connaissance ... 109
4.2. Une modification profonde du processus de relevé : le rôle de l'observateur ... 110
4.3. Vers une approche intégrée ... 112
4.4. Replacer l'objet réel au cœur du processus d'interprétation ... 113

Partie II.

Chapitre 5. Spatialisation et recalage des données numériques de relevé ... 118

5.1. Point de départ de notre travail : état actuel de la plateforme Aïoli ... 118
5.1.1. Architecture applicative ... 119
5.1.2. Données produites par la plateforme ... 119
5.2. Méthode proposée pour la mise en relation de données spatialisées : un *viewer* hybride 2D/3D 122
5.2.1. Structure proposée pour spatialiser les données dans une même scène ... 122
5.2.2. Spatialisation des données 2D autour des nuages de points de référence ... 124

5.2.3. Bilan ... 130
5.3. Méthode proposée pour le recalage en réalité augmentée des ressources d'un projet ... 131
5.3.1. Formalisation du problème ... 131
5.3.2. Approche semi-automatique proposée pour le recalage d'un projet ... 133
5.3.3. Spatialisation d'une *frame* par photogrammétrie ... 134
5.3.4. Persistance AR des projets ... 137
5.4. Bilan ... 139
Chapitre 6. Modalités d'augmentation géométrico-visuelle et d'annotation ... 141
6.1. Augmentation géométrico-visuelle des ressources du projet ... 141
6.1.1. Stratégies d'augmentation basées sur les données 2D ... 141
6.1.2. Stratégies d'augmentation basées sur les données 3D ... 146
6.1.3. Traitement temps-réel du flux vidéo ... 150
6.2. Modalités d'annotation ... 152
6.2.1. Annotation de données 2D ... 152
6.2.2. Annotation de données 3D ... 155
6.2.3. Perspectives : vers de nouvelles modalités d'annotation ... 156
6.3. Codification graphique des annotations ... 164
6.3.1. Paramétrage de matériaux d'annotation ... 164
6.3.2. Génération d'une légende graphique ... 165
6.4. Bilan ... 166
Chapitre 7. Implémentation informatique de la preuve de concept ... 168
7.1. De la réalité augmentée sur le web avec l'API WebXR ... 168
7.2. Implémentation du *viewer* hybride ... 169
7.2.1. Interface ... 169
7.2.2. Structure du *viewer* ... 170
7.2.3. Rendu temps réel ... 171
7.2.4. Implémentation des outils d'annotation, mesure, et mise à l'échelle ... 181
7.2.5. Export de vues 2D du projet depuis la scène 3D ... 183
7.3. Implémentation de la méthode de recalage avec l'API WebXR et ARKit ... 184
7.3.1. Preuve de concept ... 184
7.3.2. Résultats préliminaires ... 189
7.3.3. Évaluation des résultats ... 190
7.3.4. Limites et perspectives de la méthode de recalage ... 195
7.4. Bilan ... 196
Conclusion ... 197

Problématique 197
Bilan des contributions 198
Limites techniques et méthodologiques 199
Spatialisation des ressources dans le *viewer* 199
Prise en charge de ressources 2D visuellement éloignées du jeu d'images maîtresses 200
Fusion de données 3D issues de lasergrammétrie et photogrammétrie 200
Méthode de recalage 201
Dérive du suivi AR 201
La représentation des observations 202
Perspectives de recherche 203
Déploiement et évaluation du système 203
Un apport continu en données : vers un suivi multitemporel 204
Traçabilité : un pas vers la compréhension de la relation observateur/objet 204
Potentiels d'utilisation au-delà du domaine du relevé d'art pariétal 206
Bibliographie 207
Table des sigles et abréviations 225
Table des illustrations 226
Annexes 235

Introduction

Contexte général

Le patrimoine est une notion large qui inclut l'ensemble des biens matériels ou immatériels à forte valeur culturelle. Elle peut donc renvoyer à des édifices architecturaux, des vestiges archéologiques, des éléments mobiliers, ou encore des œuvres d'art. Son étude repose principalement sur l'analyse des ressources matérielles et des vestiges existants, de natures extrêmement variées, dont la collecte et la documentation permettent de constituer le corpus indispensable à l'interprétation scientifique. Dans de nombreuses disciplines dont l'architecture et l'archéologie, à la base de ces analyses se trouve la notion de relevé, qui correspond au processus visant à établir, par l'intermédiaire de la représentation, un état des lieux de l'objet étudié, c'est-à-dire de sa forme et des concepts qu'il supporte.

Il s'agit d'une opération d'observation, d'analyse et de synthèse qui consiste autant à extraire des informations portées par l'objet étudié qu'à les transmettre. À ce titre, pour un observateur face à son objet d'étude le relevé constitue un moment privilégié de connaissance qui présente principalement deux visées : mémoriser l'objet étudié dans ses composantes formelles (dimensions, proportions, matériaux, aspect, etc.), mais aussi sélectionner les informations (techniques, relations entre les éléments, concepts, etc.) et les hiérarchiser afin de donner les clés d'interprétation de l'objet. Dans le cas du relevé *classique*, ces opérations sont essentiellement empiriques et reposent principalement sur la transcription géométrique d'éléments significatifs, le dessin, et certaines techniques d'imagerie.

Les progrès de ces dernières décennies dans les domaines de l'informatique et de la métrologie ont favorisé l'essor d'outils numériques de mesure de l'existant. Si le numérique n'a pas fondamentalement remis en question les méthodes de mesure, l'amélioration de leur précision, de l'automatisation, et de la capacité de mémorisation ont, en revanche, constitué une évolution décisive pour de nombreuses disciplines. Dans le domaine du patrimoine, l'introduction des technologies numériques a rendu possible la collecte massive de données 2D et 3D en réponse à des besoins variés d'étude, de suivi, de documentation, d'archivage, ou encore de valorisation. Ces données constituent autant de nouveaux supports précieux pour la compréhension des objets d'étude, en particulier en ce qui concerne leur caractérisation morphologique. Cependant, au-delà de l'enthousiasme suscité par la découverte des très nombreux potentiels de ces nouveaux supports, nous pouvons constater que les données produites lors de relevés numériques demeurent souvent sous-exploitées, faute d'outils adaptés pour faciliter leur manipulation, leur analyse, et leur enrichissement sémantique.

En outre, la capacité à numériser rapidement et massivement les objets patrimoniaux, à travers la reproduction fidèle de leurs attributs géométriques et visuels, introduit un changement de paradigme majeur pour le relevé. Si dans un relevé *classique* l'engagement analytique de l'observateur est mobilisé *in situ* face à l'objet d'étude, dans le cadre d'un relevé numérique la sélection d'éléments signifiants, leur interprétation, et leur hiérarchisation interviennent *ex situ* auprès d'une copie 3D. La dissociation, voire le séquençage des moments d'acquisition, puis d'interprétation et de représentation constitue un changement majeur dans le processus du relevé, qui ne relève plus d'un moment unique d'observation du réel, mais plutôt d'une capacité à trier, écarter et mettre en relation des données virtuelles. Cette rupture entre les environnements réels et virtuels introduit une discontinuité sur le plan de la

connaissance qui se manifeste notamment par un effet de dispersion des données et une singulière difficulté à assurer la corrélation des connaissances au sein d'équipes pluridisciplinaires.

Cas d'étude

Pour ce travail, nous nous intéressons à un cas d'étude particulier qui cristallise les enjeux de ce changement de paradigme : les grottes ornées. Ce vocable réunit les sites d'art pariétal, qui représentent sur le territoire national 186 grottes, abris, ou affleurements rocheux en plein air. Notons que dans la littérature, la désignation d'objets de ce corpus a recours à trois termes. Le terme ***art rupestre*** constitue la définition la plus large, qui désigne indifféremment toute manifestation graphique sur un élément rocheux de l'environnement. Le terme ***art pariétal*** désigne un sous-ensemble d'art rupestre, limité géographiquement à l'Europe et chronologiquement au Paléolithique supérieur. L'expression ***grottes ornées*** constitue une dénomination alternative pour l'art pariétal, généralement plus éloquente pour le grand public. Elle illustre plus clairement l'une des spécificités de l'art pariétal – sa présence en milieux souterrains – mais occulte involontairement les abris sous roche et autres manifestations pourtant comprises dans le même corpus. C'est pourquoi nous privilégierons dans la suite de ce travail la dénomination *art pariétal*.

Ce corpus, constitué de typologies très variées, est en majorité classé (45,5 % des sites) ou inscrit (16 % des sites) au titre des monuments historiques. Les sites d'art pariétal constituent un cas particulier de patrimoine culturel, étant considérés à la fois comme des monuments, des sites archéologiques, et des éléments du paysage naturel *« constituant des écosystèmes complexes en relation étroite avec leur environnement »* (Ministère de la Culture, 2019). Comme le laisse présager ce statut multiple, ils constituent des cas d'étude d'une grande complexité, tant du point de vue morphologique que du point de vue de la diversité des informations qu'ils supportent. Leur étude sollicite donc une grande variété de disciplines (par exemple archéologie, anthropologie, paléogénétique, géologie, climatologie, microbiologie, etc.) et de techniques complémentaires. Si tous les acteurs s'intéressent bien au même objet patrimonial, il subsiste cependant des différences de langages, de méthodes ou d'outils qui rendent le croisement d'informations délicat.

En outre, ces sites sont extrêmement sensibles et ne supportent aucune forme de restauration. Par conséquent, ils sont très peu ouverts au public[1]. Pour les équipes scientifiques qui les étudient, les accès sont également régulés bien qu'autorisés, si bien que ce fonctionnement impose d'alterner les périodes sur le terrain avec des périodes d'étude hors site. Les visites *in situ* constituent à ce titre des expériences précieuses à optimiser, et dans la mesure du possible à sauvegarder et partager. Dans ce contexte, les apports des données numériques 3D sont incontestables, et répondent à des besoins multiples d'études scientifiques, de conservation, d'archivage, ou encore de médiation. Cependant, comme le souligne la littérature, l'étude des grottes ornées ne saurait faire l'économie du processus de découverte apporté par le travail *in situ*. Notons qu'il ne s'agit pas ici d'un fait uniquement lié aux limites des technologies d'acquisition en termes de précision ou d'exhaustivité, mais bien de l'importance de l'engagement cognitif de l'observateur face à son objet d'étude, relégué derrière le dispositif de mesure dans le cas d'une interprétation *ex situ* lors d'un relevé numérique.

[1] En 2015, seuls 15% des sites d'art pariétal sont accessibles au grand public (CNP, 2015).

L'analyse de ce contexte fait émerger la nécessité d'un nouveau paradigme pour le relevé, et plus spécifiquement le relevé d'art pariétal. Il s'agit d'envisager le relevé comme un cycle, un processus dynamique et itératif comprenant des phases *in situ* et *ex situ*, conciliant les apports du relevé numérique en termes de caractérisation géométrique, et ceux du relevé *classique* en termes d'interprétation et d'enrichissement sémantique.

Objectifs

Comme nous l'avons évoqué précédemment, nous pouvons observer que d'un point de vue quantitatif et formel, la capacité computationnelle apportée par les outils informatiques vis-à-vis des ressources numériques 2D ou 3D dépasse largement celle des outils de traitement du relevé *classique*. À l'inverse, ces mêmes outils sont très largement en deçà des capacités cognitives humaines, particulièrement dans des scénarios d'enrichissement de connaissance. Cette dichotomie dans les apports et les limites se manifeste dans la littérature par de récurrentes mises en opposition des méthodes de relevé patrimonial, dont il s'agit souvent d'évaluer laquelle prévaut et dans quel(s) contexte(s). Pour ce travail de thèse, notre posture est contraire. En effet, l'analyse du contexte du relevé d'art pariétal révèle très clairement la nécessité de les considérer comme complémentaires et de les faire converger dans une approche unique afin de bénéficier de leurs apports respectifs. Ainsi, le relevé d'art pariétal se heurte aujourd'hui à deux problèmes majeurs :

- **Un effet de dispersion** des données, informations et connaissances, dû à la diversité des acteurs, outils et méthodes de traitement ;
- **Une rupture** entre l'engagement cognitif des observateurs lors des moments d'étude sur site et l'ensemble des traitements réalisés hors site, ou autrement dit, entre les environnements de travail réels et virtuels.

Notre objectif pour ce travail est de répondre à ces problèmes par la définition d'une approche intégrée permettant la fusion des aspects géométriques, visuels et sémantiques du relevé dans un environnement unique multimodal[2]. Cette problématique place notre recherche à la croisée de deux domaines : **les systèmes d'information patrimoniaux** pour leur capacité à fédérer des données et des acteurs, et **la réalité mixte** pour sa capacité à contextualiser des données numériques sur la réalité et dans notre cas, à garantir la continuité des travaux menés *in situ* et *ex situ*.

L'état de l'art concernant ces deux domaines nous a permis de préciser notre approche et le cadre expérimental. Nous avons ainsi décidé de nous inscrire dans la continuité du travail mené par le laboratoire MAP[3] dans le cadre du développement de la plateforme Aïoli[4]. Il s'agit d'une plateforme dédiée à l'annotation sémantique de ressources numériques pour la documentation collaborative d'objets patrimoniaux. Cependant, si cette plateforme constitue une avancée notable en ce qui concerne le relevé

[2] La notion de modalité peut revêtir plusieurs sens. Le CNRTL en propose une définition générique selon laquelle une modalité correspond aux « conditions particulières selon lesquelles une chose est exécutée, réalisée ». Dans le domaine de la communication Homme-Machine, Nigay et Coutaz (1996) définissent la modalité comme un échange entre un système et son utilisateur, impliquant un niveau matériel faisant appel à la perception, et un niveau représentationnel faisant appel à la cognition. Elles proposent de caractériser les modalités comme un ensemble de couples <p, r> où p désigne un dispositif physique, et r, un système représentationnel. Plus largement, une modalité peut être assimilée à une technique d'interaction. C'est le choix que nous avons opéré ici, le terme « multimodal » renvoyant alors à l'implication de plusieurs modalités.
[3] Modèles et simulations pour l'Architecture et le Patrimoine, UMR 3495 CNRS/MC. www.map.cnrs.fr
[4] Plateforme d'annotation sémantique 3D pour la documentation collaborative d'objets patrimoniaux. www.aioli.cloud

numérique, elle présente encore des limites significatives au regard des enjeux énoncés. En particulier, elle n'autorise pas le travail *in situ* et l'ensemble des ressources manipulées, bien que spatialisées de manière relative, demeure un ensemble décontextualisé et déconnecté de l'objet étudié.

L'approche que nous proposons consiste à étendre à l'objet réel la relation projective établie entre les différentes ressources 2D et 3D relatives à un objet d'étude. Concrètement, il s'agit de considérer que le nuage de points 3D d'un objet d'étude n'est pas qu'un substitut géométrique et visuel de l'objet, mais bien une donnée structurante que l'on peut exploiter à la fois pour fédérer les nombreuses ressources impliquées dans un relevé et comme interface entre l'objet réel et les données numériques. Concernant le cadre expérimental de ce travail, nous avons visé l'élaboration d'une preuve de concept se présentant sous la forme d'une application web. Celle-ci tient Aïoli comme base de départ (architecture logicielle, structuration des données, création de projets, processus de propagation, etc.), étendue de trois préoccupations distinctes :

- **Fédérer** : l'application ne doit pas seulement centraliser l'ensemble des données, informations, et connaissances relatives à l'objet d'étude, elle doit aussi les mettre en relation spatialement au sein d'un espace de visualisation unique temps réel autorisant leur manipulation ;
- **Hybrider** : les données virtuelles doivent être accessibles *in situ* en réalité augmentée afin de permettre aux utilisateurs de les visualiser de manière contextualisée et d'interagir avec un environnement hybride comprenant l'objet réel et les informations numériques associées ;
- **Augmenter** : il s'agit d'exploiter les attributs géométriques, visuels, et sémantiques à disposition pour faciliter le processus d'analyse et de suivi et enrichir la compréhension de l'objet. Cela passe par la définition de modalités de rendu facilitant le processus d'annotation et de description sémantique.

Portées, limites, apports principaux

Ce sujet de thèse se situe dans le cadre de l'élaboration d'un système d'information spatialisé multimodal dédié au relevé d'art pariétal. L'apport de cette thèse est donc principalement d'ordre méthodologique, par l'identification d'une approche permettant de faire converger les activités qui relèvent de la saisie de données spatiales avec celles qui relèvent de l'interprétation à travers des modalités d'enrichissement sémantique. Cependant, les objectifs énoncés se traduisent par des problèmes spécifiques ayant fait l'objet de trois autres contributions distinctes. Celles-ci concernent donc :

- L'identification d'une approche générale pour faire converger les pratiques du relevé *classique* et du relevé numérique ;
- La mise en relation de données spatialisées au sein d'un *viewer* 2D/3D basé sur l'API WebGL ;
- La proposition d'une méthode générique adaptée aux environnements web basée sur l'API WebXR pour l'alignement réel/virtuel semi-automatique de scènes basées sur la réalité ;
- Le développement de stratégies d'augmentation d'attributs géométrico-visuels de l'environnement destinées à faciliter le processus d'interprétation.

En conséquence, ce document est organisé à l'image de ces quatre contributions, c'est-à-dire en deux parties, l'une théorique aboutissant à la définition d'une approche générale, la seconde focalisée sur les trois problèmes spécifiques soulevés par l'approche proposée.

Structure du document

Partie I : Cette première partie dresse le contexte de notre étude et soulève les enjeux et problématiques liées au relevé d'art pariétal. Elle présente un état de l'art des approches concernant les systèmes d'information patrimoniaux et les méthodes d'hybridation, et aboutit sur la définition de notre approche générale, qui constitue notre première contribution.

Chapitre 1 : Ce chapitre a pour objet de présenter le relevé patrimonial puis plus spécifiquement le relevé d'art pariétal, afin d'énoncer les enjeux et problématiques soulevés par cette pratique. Nous montrons en particulier que le relevé d'art pariétal impose un arbitrage difficile entre la justesse géométrique et l'expression graphique des éléments relevés. Face à ce contexte, nous constatons que les outils numériques présentent de nombreux potentiels, qui ne sont pas sans soulever de nouvelles problématiques.

Chapitre 2 : Ce chapitre présente les outils et méthodes impliqués dans le relevé numérique. Cet état de l'art révèle un changement de paradigme majeur soulevant des problématiques de dispersion des données et de rupture entre les environnements réels et virtuels. Nous concluons ce chapitre sur les nécessités de définir une approche intégrée dédiée au relevé d'art pariétal capable de fédérer à la fois les données et les acteurs, et de réconcilier les moments *in situ* et *ex situ*.

Chapitre 3 : Les enjeux soulevés à l'issue des deux premiers chapitres placent notre recherche à la croisée de deux domaines, les systèmes d'information patrimoniaux et la réalité mixte. Ce chapitre dresse donc un état de l'art des approches existantes, qui nous permet de définir le cadre expérimental de notre travail. Nous y motivons notamment le choix de nous inscrire dans la continuité des travaux menés sur la plateforme web Aïoli, et d'exploiter l'API WebXR concernant le développement de modalités d'interactions en réalité augmentée.

Chapitre 4 : À la lumière des états de l'art, ce chapitre discute les problématiques impliquées par le relevé d'art pariétal, et formule notre approche, qui consiste à proposer un système d'information spatialisé multimodal dédié au relevé d'art pariétal. Nous proposons d'exploiter à la fois le nuage de points pour fédérer les ressources 2D et 3D liées à l'objet d'étude, et comme intermédiaire permettant le recalage des données virtuelles sur la réalité. Nous discutons également dans ce chapitre les implications pratiques de cette approche, qui soulève trois problèmes principaux faisant l'objet d'expérimentations.

Partie II : Cette seconde partie présente les méthodes proposées concernant les différents aspects des développements expérimentaux et la preuve de concept obtenue. Celle-ci se présente comme une plateforme web permettant l'annotation sémantique de ressources impliquées dans le relevé d'art pariétal selon différentes modalités – notamment en réalité augmentée – comme annoncé en chapitre 4.

Chapitre 5 : L'approche proposée dans le chapitre 4 nécessite la mise en relation des ressources spatialisées dans un espace de visualisation unique, mais aussi la levée du verrou de l'alignement réel-

virtuel, qui plus est dans un environnement web. Ce chapitre présente les méthodes proposées pour traiter ces problèmes. Concrètement, il s'agit d'un *viewer* hybride 2D/3D, et d'une méthode générique pour le recalage semi-automatique sans marqueurs reposant sur un calcul de résection spatiale, qui constituent deux autres contributions.

Chapitre 6 : À partir des travaux réalisés sur la visualisation et l'hybridation, le chapitre 6 présente des méthodes de traitement des données visant à faciliter l'observation de stigmates et traces anthropiques sur les parois. Ces méthodes se basent sur l'augmentation des attributs géométriques et visuels des ressources spatialisées dans le *viewer*. Elles sont accompagnées de propositions d'outils d'annotation sémantique adaptés aux différentes modalités de visualisation. Elles constituent notre quatrième contribution.

Chapitre 7 : Ce chapitre présente l'implémentation informatique des méthodes proposées dans les chapitres précédents au sein d'une preuve de concept fonctionnelle.

Finalement, nous proposons en conclusion un récapitulatif du travail présenté dans ce mémoire. Après avoir rappelé le contexte et les enjeux de cette étude, nous dressons le bilan des contributions, mais aussi leurs limites et les perspectives de recherche qu'elles soulèvent.

Note concernant l'objet d'expérimentation

Pour des raisons pratiques, les expérimentations menées dans le cadre de ce travail ont été réalisées sur un facsimilé du panneau des chevaux de la grotte Chauvet-Pont d'Arc. Celui-ci nous a aimablement été prêté par le Syndicat Mixte Espace de Restitution de la Grotte Chauvet (SMERGC) que nous remercions. Parmi les différentes options envisagées, ce panneau a été sélectionné pour la diversité des traces qu'il supporte, qui comprennent notamment des peintures, des gravures anthropiques et animales, des fissures, ainsi que des matériaux d'aspects variés. Ce point était essentiel, particulièrement pour la proposition de méthodes d'augmentation d'attributs géométrico-visuels, qui devait se baser sur un cas d'étude représentatif de la diversité des manifestations auxquelles un observateur peut être confronté sur les objets du corpus des grottes ornées.

Partie I

Chapitre 1. Le relevé

1.1. Le relevé du bâti

Par bien des aspects, les grottes ornées et le patrimoine bâti diffèrent. Certains rapprochements peuvent néanmoins être faits, particulièrement en ce qui concerne les approches relatives au relevé ou à la documentation. Qu'il s'agisse des questionnements concernant la posture de l'observateur et de sa subjectivité inévitable, de la position à accorder à l'objet d'étude qui peut être à la fois considéré comme œuvre et témoin (d'un environnement, d'une pratique, d'une époque, etc.), ou encore de la nécessité de faciliter des analyses croisées au sein d'équipes multidisciplinaires, certains questionnements récurrents laissent penser à une convergence d'enjeux et à la possibilité d'y apporter en réponse une approche globale. C'est pourquoi nous nous intéresserons au préalable au relevé du patrimoine bâti.

Dans un premier temps, nous nous attacherons à montrer en quoi l'émergence des pratiques de relevé du bâti découle de la mutation progressive des pratiques architecturales et des modalités de représentation. Nous chercherons à souligner en particulier le lien profond qui lie le relevé aux questions de la représentation et de la mesure de l'existant. Par la suite, nous nous intéresserons aux pratiques du relevé d'art pariétal. Nous tenterons de comprendre comment la nature des sites et des œuvres a influencé les techniques de relevé, que ce soit en termes de mesure ou de codification graphique. Nous nous questionnerons sur les difficultés, problématiques et enjeux du relevé d'art pariétal. Enfin, nous pourrons étudier les pratiques contemporaines du relevé numérique. Nous verrons comment le développement des techniques d'acquisition 3D a provoqué un changement de paradigme important dans le domaine du relevé, en remisant *ex situ* le travail d'analyse.

1.1.1. Historique : de la représentation d'architecture au relevé du bâti

Contrairement à ce que l'on pourrait croire, l'usage systématique de la représentation dans les pratiques de l'architecture est un fait relativement récent, dont on estime, en Occident du moins, qu'il ne date que de la Renaissance. En cause, la position de l'architecte par rapport au chantier. Nous verrons dans cette partie comment la mutation des pratiques architecturales à la Renaissance a engendré des conséquences majeures dans le champ de la représentation d'architecture, et a notamment amorcé les nombreux traités de géométrie descriptive, puis démocratisé l'usage de la perspective. Ces aspects sont essentiels si l'on souhaite interroger la notion de relevé, dont le processus destiné à la compréhension et la figuration de l'existant repose essentiellement sur notre capacité à le représenter, que ce soit en deux ou en trois dimensions.

1.1.1.1. La genèse du tracé planimétrique

Les définitions du terme *représentation* par le CNRTL dans l'acception qui nous intéresse – celle qui est en jeu dans la question de la représentation d'architecture – renvoie tant au fait de représenter, c'est-à-dire au processus, qu'à la manière dont la chose est représentée, c'est-à-dire à son résultat. La notion de représentation devient alors indissociable de l'observateur qui initie la démarche, car c'est son processus cognitif qui entre en jeu dans la caractérisation des éléments essentiels à la figuration du bâti étudié, et la production d'un document communicable. Il devient alors essentiel de souligner la

subjectivité du produit de ce processus, à l'origine d'un problème de traçabilité encore d'actualité malgré l'évolution des pratiques.

> *« […] l'usage, en général, commande le choix du mode de représentation de l'espace et on ne peut étudier la géométrie sous-jacente d'un dessin sans connaître la destination que lui assignait son auteur. »* (Sakarovitch, 1998)

Les exemples de documents graphiques relatifs à l'usage d'une représentation géométrale et antérieurs au XI^e^ siècle sont particulièrement rares. L'un des plus vieux témoignages est cependant daté à environ 2120 avant J.C. Il s'agit de la statue dite de « l'Architecte au plan », représentant un homme assis, une tablette gravée déposée sur les genoux, accompagnée d'instruments de dessin et d'une règle finement graduée (Figure 1). Si la gravure sur la tablette, représentant l'enceinte d'un temple, ne permet pas de statuer sur un usage systématique du géométral à cette période, elle demeure néanmoins la preuve qu'une telle pratique existait bien avant les célèbres traités de la Renaissance (Lo Buglio, 2016 ; Sakarovitch, 1998).

Figure 1 : Statue de Gudea, prince de Lagash dite « statue de l'architecte au plan ». Photographies © RMN-Grand Palais (Musée du Louvre) / René-Gabriel Ojéda.

Différents papyri attestent de l'utilisation de tracé planimétrique en Égypte antique, notamment parmi la collection des papyri de Turin[5], certains étant même accompagnés de légendes sommaires. Un cas remarquable est celui du papyrus de Gui'ab qui présente deux élévations différentes d'un même sanctuaire, tracées en noir sur fond quadrillé rouge, ce qui traduit, selon Sakarovitch, un grand souci de précision ainsi qu'une *« maîtrise de l'échelle et de la correspondance entre les deux vues »*. Bien qu'il soit difficile de statuer sur les rôles spéculatifs, prescriptifs ou descriptifs de ces témoignages, certains indices comme la finesse des tracés ou la présence de détails spécifiques laissent cependant penser qu'il s'agit de représentations de l'existant plutôt que de supports de conception ou de construction. L'Antiquité gréco-romaine laisse, quant à elle, peu de traces de représentations géométrales d'architecture, ce que certains expliquent par la fragilité supposée des supports utilisés (Coulton, 1984),

[5] Les papyri de Turin sont une collection de papyri égyptiens conservés au Musée Égyptologique de Turin, majoritairement rapportés par Bernardino Drovetti au début du XIX^e^ siècle.

et que d'autres interprètent comme une conséquence logique d'une utilisation particulièrement rigoureuse de règles de proportion et de normes de conception (Bundgaard, 1957 ; Will, 1984). Quoi qu'il en soit, les prémices d'une représentation géométrale sont plutôt attribuées au Moyen Âge.

La période romane livre, elle aussi, peu d'éléments de représentations architecturales. En revanche, de nombreux témoignages semblent confirmer l'utilisation d'épures, exécutées *in situ* et en grandeur réelle, afin de guider la construction du bâti. Ce n'est globalement qu'à la période gothique que l'on voit émerger le dessin d'architecture comme médium. L'ampleur des chantiers, leur complexité, la nécessité d'optimiser la planification du travail et de convaincre des financeurs auraient en effet largement contribué à la production d'éléments graphiques de représentation. En France, la première mention d'une construction devant être réalisée à partir d'une représentation graphique date de 1381, lors de la réfection d'un clocher à Toulouse (Mortet, 1900).

Cette pratique n'est pour autant pas encore démocratisée en Europe. En 1391, pour l'érection de la cathédrale de Milan, un conseil d'architectes venant de toute l'Europe est convoqué par les instances dirigeantes de la ville afin de statuer sur la composition de la façade principale, la construction de l'édifice étant pourtant déjà largement entamée. La conception du bâti n'est alors pas encore limitée par le démarrage du chantier. En revanche, la mise en correspondance entre plusieurs représentations géométrales devient plus fréquente, et les exemples se multiplient. Notons particulièrement le célèbre carnet de Villard de Honnecourt, (1200-1250), comportant 250 dessins, dont 74 liés à l'architecture, d'une variété remarquable : plans et élévations mis en relations, accompagnés de tentatives de perspectives.

1.1.1.2. Renaissance italienne : vers une mutation des pratiques architecturales, conséquences sur les pratiques graphiques

On attribue à Brunelleschi l'invention de la perspective au XV^{e} siècle, avec pour objectif de peindre avec la plus grande exactitude possible la réalité. Il est également considéré comme le premier à en avoir fait usage en architecture. Sa démarche marque un changement profond dans le domaine de la représentation, le dessin se définissant alors pour la première fois comme un acte de connaissance.

Au milieu du XV^{e} siècle, dans son traité d'architecture *De re ædificatoria*, Alberti affirme le nouveau statut de l'architecte, résultant d'une dissociation entre les activités de conception et de réalisation. De fait, l'architecte s'éloigne progressivement du chantier, et incidemment, de la figure de l'architecte bâtisseur, pour transiter vers celle de l'architecte concepteur, encore en vigueur de nos jours. La profession d'architecte évolue ainsi progressivement d'un *ars mechanica* vers un *ars liberalis*. En conséquence de cette mutation des pratiques, la production de documents graphiques précis destinés à appuyer la réalisation s'avère de plus en plus souvent nécessaire. Alberti prescrit à ce titre l'usage du dessin comme outil systématique de représentation, recommandant particulièrement l'usage de plans et d'élévations – sur feuilles séparées – mais condamnant l'utilisation de la perspective, alors considérée comme apanage des peintres :

> *« Entre la manière de tracer du peintre et celle de l'architecte, il y a cette différence que le premier s'emploie à faire ressortir les objets du tableau par le moyen des ombres et la diminution des lignes et des angles, tandis que l'architecte évitant l'ombrage, figure les saillies moyennant les tracés du plan, il instruit dans d'autres dessins sur la forme et l'étendue de chaque face et de chaque côté,*

> *en recourant aux angles vrais et à des lignes constantes, à la manière de qui voulût que ses affaires ne fussent pas connues par des apparences, mais bien notées par des raisons et des dimensions certaines. »* Alberti, cité par (Boudon et al., 1974)

Notons qu'Alberti explicite ici le rôle prescriptif du dessin sur lequel l'observateur doit pouvoir mesurer les *vraies* grandeurs nécessaires à la réalisation de l'édifice. Le dessin n'a cependant pas de réelle vocation de médiation : il sert à la compréhension, et non à l'appréhension du bâti en devenir. Quelques années plus tard (entre 1461 et 1464), Antonio Averlino plus connu sous le nom de Filarete rédige *Trattato d'Architettura* dans lequel il distingue trois fonctions du dessin d'architecte. La première correspond au croquis ou dessin d'esquisse, destiné à stimuler la créativité de l'architecte et d'assister la conception. Suivent le dessin en tant qu'instrument de dialogue entre l'architecte et son client puis le dessin à l'échelle comme instrument de la réalisation du projet[6] (Sakarovitch, 1998).

L'intérêt porté à la Renaissance pour l'étude du monde antique favorise alors considérablement ce renouveau des pratiques graphiques. Cela mène à la formalisation des techniques de représentation, d'abord géométrales (*Quattrocento*) puis perspectives (*Cinquecento*), qui entrent progressivement dans les mœurs. À la fin du XV^{e} et du début du XVI^{e}, les architectes italiens commencent en effet à manifester un intérêt pour ce nouveau mode de représentation initialement réservé à la peinture, et des représentations perspectives d'architecture commencent à émerger. Néanmoins, l'utilisation d'un point de fuite surélevé nous permet de supposer que ces éléments graphiques restent encore majoritairement dédiés à la nécessité de faire comprendre le projet, sans s'intéresser aux aspects cognitifs du dessin, suscités par l'adoption de points de vue à hauteur humaine.

Un rapport de Raphaël, adressé au pape Leon X en 1519, précise les préoccupations concernant la nécessité d'adopter des modes de représentations graphiques codifiés pour l'architecture (Choey, Paoli, 2005). Raphaël y reprend le point de vue d'Alberti concernant l'usage des plans et élévations, qu'il complète cependant en abordant le principe des coupes, qu'Alberti n'avait pas évoqué dans ses traités. Si l'auteur s'oppose à l'utilisation seule de la perspective comme outil de transmission et de communication de l'architecture, il demeure cependant moins catégorique qu'Alberti, et ne se positionne pas fondamentalement contre cette dernière, qu'il considère comme un mode de représentation complémentaire aux autres, utile pour *« juger la grâce formelle »* d'un projet (Germann, 1991). Ce témoignage démontre d'une diversification des modes de représentation, où la multiplication des supports vise à offrir une certaine exhaustivité dans la compréhension et l'appréhension du bâti. Ce point de vue acte pour la première fois un panorama des outils de représentation utilisés par l'architecte jusqu'à aujourd'hui encore : plans, coupes, élévations, complétés de représentations perspectives (Viallon, Chatenet, 2008).

À la fin du XVI^{e} siècle, considérant que la perspective centrale de Brunelleschi offre une vision trop ponctuelle de l'espace pour permettre la compréhension globale de certaines architectures (notamment classique ou militaire dont les géométries sont à saisir dans leur plan d'ensemble), certains architectes comme Androuet du Cerceau optent pour une prise de recul importante (Figure 2).

[6] On retrouvera une décomposition similaire des rôles du dessin bien plus tard, dans les théories d'Eugène Ferguson, qui parle en 1992 de :
1. Dessin « croquis de la pensée », supportant l'activité mentale du concepteur ; 2. Dessin « parlant », vecteur de la communication et des discussions avec les acteurs du processus de conception ; 3. Dessin « prescriptif », qui détaille et spécifie l'objet conçu pour les observateurs extérieurs au processus de conception.

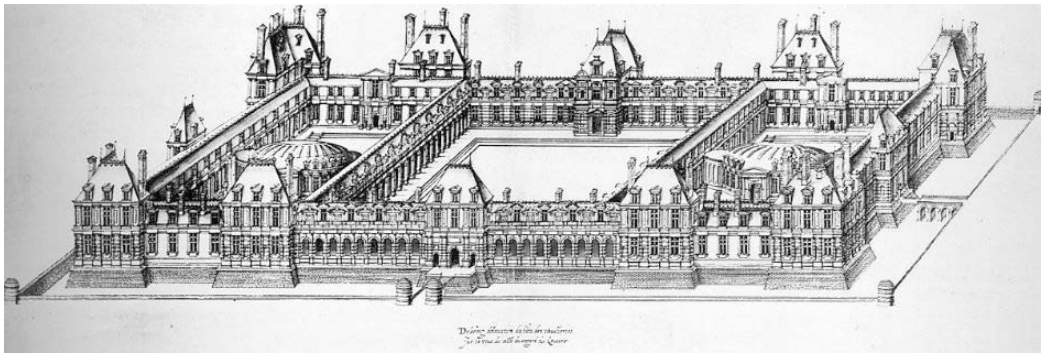


Figure 2 : Perspective cavalière d'un projet pour le palais des Tuileries par Jacques Androuet du Cerceau en 1579. Illustration : (Babelon, 1989).

Celle-ci se traduit par l'adoption d'un point de vue aérien et des points de fuite à l'infini, qui aboutit à ce que l'on appelle encore aujourd'hui la perspective *cavalière*, *militaire*, ou *parallèle*. Cette pratique reste néanmoins longtemps dénigrée des adeptes de la perspective centrale, qui n'y voient qu'un substitut pour ceux répugnant à appliquer des méthodes de construction géométriques savantes (Comar, 1992). Plus tard, la capacité de cette représentation graphique à informer précisément sur les dimensions de l'objet représenté, tout en offrant une vision d'ensemble rendant compte de sa volumétrie s'avère déterminante pour son introduction au domaine de l'ingénierie, dont l'application majeure se trouve dans le champ du dessin technique. Ce n'est toutefois que l'introduction de la géométrie descriptive dans l'enseignement à l'école Polytechnique à la fin du XVIII^{e} siècle par Gaspard Monge, qui marque l'avènement de cette méthode de représentation dans le milieu industriel. En architecture, la perspective parallèle ne devient au moyen privilégié de représentation qu'au début du XX^{e} siècle et prend alors un nouveau nom : l'axonométrie.

Au regard de cet historique, nous pouvons constater que les pratiques graphiques à l'usage en architecture n'ont pas subi de changement majeur depuis cette *révolution* que représente la Renaissance. Le changement de paradigme de l'architecte bâtisseur à l'architecte concepteur ne semble pas avoir été fondamentalement remis en question, et les supports graphiques préconisés par Alberti puis généralisés par Raphaël sont toujours en vigueur de nos jours, à la base de tous les modes de représentation spatiale, assistés ou non par ordinateur (Viallon, Chatenet, 2008). Par ailleurs, nous pouvons noter, à travers l'opposition entre l'usage de la perspective en peinture et son usage en architecture, la naissance d'un dilemme que l'on retrouve aujourd'hui encore dans la pratique du relevé et concerne le choix difficile entre la capacité d'une représentation à *faire comprendre* par le biais de la mesure et de la justesse géométrique, et sa capacité à *faire voir*, par le biais de son expression graphique.

1.1.1.3. De la perspective à la mesure de l'existant : genèse du relevé

L'émergence du relevé d'architecture est intimement liée à l'apparition progressive de la perspective à la Renaissance. La fascination éprouvée à l'époque pour l'architecture antique fait naturellement émerger la nécessité de documenter, décrire précisément, et transmettre par codification graphique le bâti existant. Cela se traduit par un intérêt particulier pour les techniques de mesure et de représentation permettant de mesurer et rendre compte en deux dimensions des monuments en trois dimensions. Dans *Vitruve et le Vitruvianisme*, Georg Germann explique à ce propos les théories exposées par Alberti dans *De re ædificatoria* : *« Pour tout architecte, il est essentiel d'avoir pris les mesures de monuments*

antiques célèbres et d'en avoir établi les relevés, afin de pouvoir se les représenter à l'aide de plans et de maquettes. Il ne s'agit pas pour lui de copier, mais de saisir l'intention artistique de l'architecte de l'époque » (Germann, 1991).

Le relevé apparaît alors déjà comme un processus permettant à l'observateur de comprendre le bâti, à travers une série d'activités comprenant la mesure de l'existant, la sélection d'éléments signifiants, leur interprétation et leur figuration au moyen d'une représentation graphique codifiée.

Queysanne (2000) explique comment les architectes de la Renaissance, et particulièrement Brunelleschi, Alberti et Raphaël, éprouvent le besoin profond de retrouver et comprendre les principes de l'architecture antique à travers des fouilles et relevés de vestiges, afin de proposer et réaliser une architecture nouvelle. La compréhension de ces monuments antiques pose alors un problème majeur : hormis le traité de Vitruve et quelques rares écrits, les sources documentaires sont quasiment inexistantes. Le relevé des monuments devient alors une nécessité, mais également celle des villes. C'est dans ce contexte qu'Alberti propose un dispositif de mesure et de représentation destiné à cartographier la ville de Rome et particulièrement ses points remarquables (Figure 3). Ces points d'intérêts sont identifiés par des coordonnées polaires précises. Ainsi, connaissant la position de l'opérateur du relevé et l'ensemble des mesures polaires, il devient possible de reproduire simplement et sans erreur la cartographie. Le principe décrit évoque vivement le principe des techniques de lasergrammétrie encore exploitées de nos jours[7].

> *« Faites un cercle sur un panneau large d'au moins un bras, et marquez tout autour sur ce cercle des parties égales autant que vous le voudrez ; et plus il y aura de parties, le mieux cela sera, pourvu qu'elles soient bien distinctes et non confuses. […] Munissez-vous d'un fil à plomb, éloignez-vous de cet instrument de deux bras et visez une à une les choses que vous voulez noter, de telle manière que votre regard passe par le fil à plomb, le milieu du centre du cercle et se dirige vers la tour que vous visez. Et le chiffre que vous verrez là où l'extrémité du cercle est coupée vers là où vous visez vous vous en souviendrez en l'inscrivant sur quelque papier. »* (Queysanne, 2000)

Si les techniques de prise de mesure et de représentation diffèrent, les travaux de Brunelleschi sur la perspective vont, eux aussi, dans ce sens : il s'agit bien de comprendre l'édifice, d'en saisir les éléments les plus significatifs, et chercher comment donner à voir en deux dimensions un objet qui en comporte trois, tout en limitant autant que possible la perte d'informations.

Aujourd'hui encore cette ambition du relevé d'architecture demeure. Il existe une ambiguïté dans l'utilisation du terme *relevé* qui, comme pour la représentation, peut se rapporter tant au processus qu'à son résultat. Dans les deux cas, l'observateur est au cœur de la démarche. Cependant, comme nous avons pu le voir, le relevé constitue en lui-même un acte d'analyse permettant à un observateur de comprendre son objet d'étude. En cela, il se présente avant tout comme un processus, dont l'élaboration d'une représentation graphique communicable constitue la finalité. La représentation graphique permet quant à elle de communiquer cette analyse de l'objet, mais aussi d'en dresser un état daté, ce qui peut notamment être exploité ultérieurement pour contrôler son évolution ou planifier sa restauration.

[7] Voir section 2.2.2.1 Les scanners laser, dont les principes de balayage circulaire et de mesure des distances relativement à une station de position et d'orientation connues ne semblent compléter le principe exposé par Alberti que par l'automatisation et la multiplication des hauteurs de mesure.

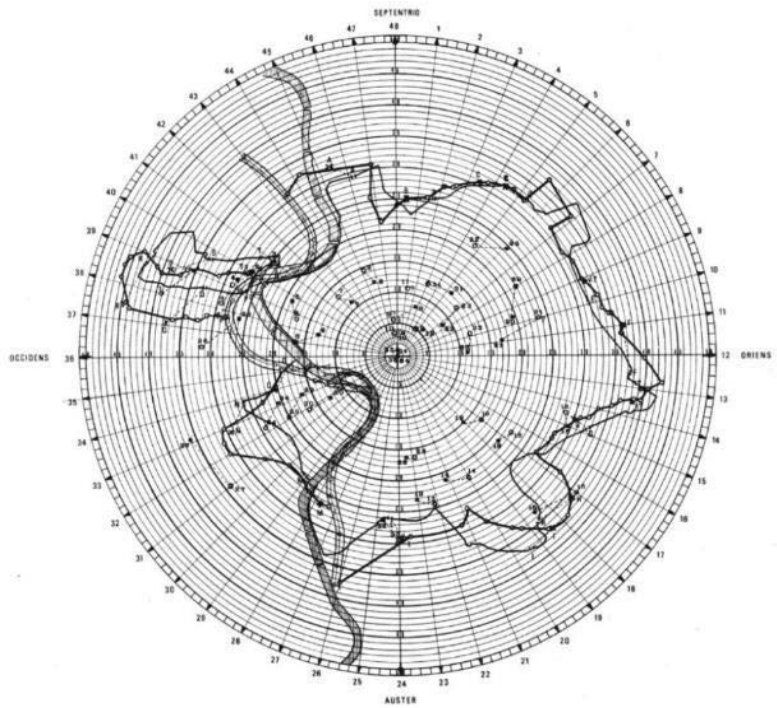


Figure 3 : Cartographie suivant la méthode d'Alberti. Illustration : (Cataldi et al., 2000).

1.1.2. Les moments du relevé : Décrire, Figurer, Représenter

En 1992, Jean-Paul Saint-Aubin définit le relevé comme le processus qui *« conduit à dresser, par l'intermédiaire de la représentation, le constat daté de la forme effective du bâti, avec comme référence constante, l'objet : ses imperfections rendent compte de l'insuffisance des techniques, des outils et des hommes ; ses remaniements racontent son histoire et ses usages ; ses défaillances comme ses déformations en annoncent l'état sanitaire »* (Saint Aubin, 1992).

Il apparaît clairement que le relevé ne se limite pas à la mémorisation des aspects géométriques de l'objet d'étude, mais bien à une sélection d'éléments significatifs au regard d'une préoccupation donnée. Autrement dit, le relevé peut être considéré comme un acte cognitif, un moment privilégié de connaissance du bâti reposant sur l'analyse des relations entre la forme, la géométrie, l'aspect, et le sens architectural. À ce titre, d'après Saint-Aubin, le processus du relevé comprend trois moments interreliés : ***décrire***, ***figurer***, ***représenter***.

Décrire correspond à la saisie la plus exhaustive possible des données significatives de l'objet. Il s'agit d'une opération de mesure orientée vers la satisfaction d'une demande spécifique, d'une problématique guidant l'observation de l'objet. Il s'agit d'une phase relativement objective dans le sens où il s'agit essentiellement d'une opération de mesure d'éléments dont le nombre et la nature dépendent cependant des objectifs fixés.

Figurer vise à condenser l'objet d'étude en un assemblage de représentations simplifiées et à hiérarchiser les informations. Il s'agit d'une opération de traduction graphique qui introduit la nécessité d'établir une codification, une symbolique de rendu visant à compenser la perte d'information. Elle s'accompagne donc de la qualification des éléments figurés. Il s'agit d'une phase sémantique, plus subjective que la précédente.

Représenter, enfin, correspond à l'élaboration d'un document graphique intelligible et communicable. Cela passe par des prises de décisions profondément liées aux moments précédents, notamment le choix des dimensions de la représentation et de la densité d'informations à figurer, mais aussi des modalités de visualisation. Ce moment cristallise le dilemme hérité de la Renaissance, il implique un arbitrage difficile entre la préservation de la vision de l'objet dans ses propriétés perceptives (c'est-à-dire principalement la restitution de ce que l'œil voit), ou dans ses propriétés dimensionnelles et formelles.

Ces trois moments sont profondément liés étant donné que le choix des éléments à mesurer est influencé par l'assignation des éléments observés et leur hiérarchisation au regard de la demande initiale, et incidemment, par l'anticipation des modalités de figuration, et de représentation (Figure 4). Le relevé apparaît donc comme une démarche se situant à l'intersection d'un engagement cognitif et d'un processus de mémorisation objective. Il se caractérise alors principalement par deux éléments majeurs : le système d'analyse défini par l'utilisateur, c'est-à-dire le questionnement qui guide la démarche du relevé, et la précision de l'information acquise, c'est-à-dire sa densité et sa justesse.

Comme nous le verrons dans la section suivante, bien que la nature des objets d'étude diffère, le relevé d'art pariétal constitue une démarche globalement similaire : il s'agit bien d'un processus qui se base sur l'observation du réel pour analyser et rendre compte des relations entre la morphologie, l'aspect, et le sens des éléments signifiants de la paroi.

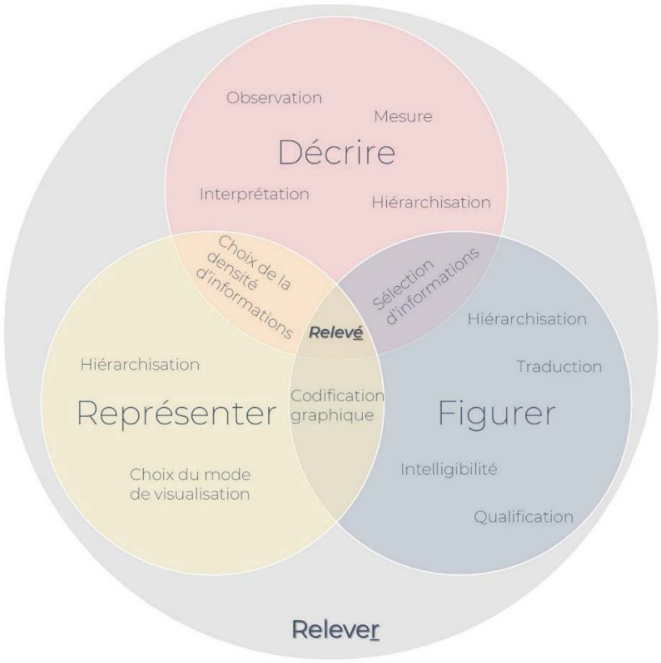


Figure 4 : Schéma présentant les trois moments interreliés du relevé. Illustration : auteur.

1.2. Le relevé d'art pariétal

Traditionnellement, en français le terme *art pariétal* désigne l'art des cavernes et des abris paléolithiques, *art rupestre* étant une dénomination plus large désignant les roches gravées ou peintes des périodes postglaciaires. Notons cependant que cette distinction n'existe pas en anglais, toutes les formes d'art sur support rocheux étant désignées par le terme *rock art*. S'il existe de nombreux sites d'art rupestre répartis dans le monde avec plus de 300 000 sites majoritairement répartis en Afrique et en Asie, les sites d'art pariétal, plus anciens, sont bien moins nombreux et concentrés en

Europe (Geneste, 2006). On dénombre ainsi environ 350 sites d'art pariétal paléolithique. Ceux-ci sont principalement concentrés dans la région franco-cantabrique (nord de l'Espagne et sud de la France), avec 186 sites en France et 130 en Espagne[8].

Historiquement, le relevé d'art pariétal est une discipline plus tardive que le relevé du bâti, les premières grottes ornées préhistoriques ayant été découvertes au XIX^{e} siècle. Comme nous le verrons dans les sections qui suivent, tous deux entretiennent cependant de nombreuses similarités. Comme dans le cas du relevé du patrimoine bâti, le relevé d'art pariétal est une opération qui consiste à la fois à *extraire* des informations de l'objet – une paroi ornée – et à les *transmettre*. Le terme renvoie donc, ici aussi, tant au processus qu'à son résultat. La démarche de l'observateur tient donc un rôle primordial, tout comme les modalités de restitutions choisies.

Les données de l'art pariétal ne sont pas directement accessibles, elles imposent un travail de déchiffrage pour en saisir la complexité. Aussi, le relevé, en tant que processus, constitue pour les préhistoriens une étape fondamentale de la recherche : à l'instar de la fouille pour l'archéologue, le relevé permet en premier lieu d'enregistrer les vestiges, et de servir de base au travail d'interprétation. Le relevé en tant que résultat permet quant à lui d'illustrer un propos scientifique, mais aussi, parce qu'il offre une représentation globale des motifs et de leur support, de dresser un état de la paroi permettant de contrôler son évolution et d'aider sa conservation.

1.2.1. Historique : vers une recherche d'exactitude

L'émergence tardive d'une pratique du relevé d'art pariétal s'explique principalement par la période de scepticisme qui suit sa découverte. En effet, si aujourd'hui l'authenticité des peintures d'Altamira n'est plus remise en question, à la fin du XIX^{e} siècle la richesse et la sophistication de ces œuvres fait douter la communauté scientifique qui peine à se dégager de l'idée d'un homme préhistorique fruste et qui craint de se faire duper par des faux. L'art mobilier pourtant, ne semble pas avoir souffert de ces préjugés puisque, près de trente ans plus tôt, dans les années 1860, les premiers préhistoriens soulignent déjà la qualité des pièces gravées ou sculptées découvertes, dont certaines sont même considérées comme des chefs-d'œuvre (Hurel, 2013).

> *« [l'homme préhistorique] apparaît comme un ouvrier habile qui orne ses outils à ses moments perdus, tel un pâtre désœuvré. L'idée d'une activité artistique en rupture avec la culture technique est incompatible avec la représentation alors dominante de l'homme "antédiluvien" largement déterminé par ses besoins primaires. C'est pourquoi, bien que les fresques d'Altamira mettent en scène une faune semblable à celle de l'art mobilier et soient en parfaite cohérence avec des faits déjà admis, la découverte demeure inacceptable. »* (Fraenkel, 2007)

[8] Cette répartition très localisée explique la prépondérance de sources françaises dans la littérature. Par ailleurs, durant plusieurs décennies, le relevé d'art pariétal constituait un acte individuel et, de fait, concernait une communauté relativement restreinte de chercheurs. Dans ce chapitre, ce contexte se ressent particulièrement au niveau de l'emploi des sources, majoritairement francophones. Notons qu'il impacte bien moins le chapitre suivant, dans la mesure où l'étude d'outils et méthodes modernes mobilise nécessairement une littérature anglophone, dans laquelle la distinction entre l'art rupestre et l'art pariétal n'est pas opérée. Nous y mobilisons donc un champ plus large de références, en justifiant néanmoins ces choix par des similarités constatées, parfois au niveau des objets d'étude mais principalement au niveau des enjeux des outils et méthodes étudiés.

Néanmoins, même si leur authenticité est acceptée, les pièces d'art mobilier ne font pas l'objet de relevés à proprement parler, mais plutôt d'illustrations esthétiques (Piette, 1907 ; Fuentes, 2017). Quant à l'art pariétal, du fait des débats concernant l'historicité des œuvres, leurs illustrations ont avant tout vocation à rendre compte au monde de l'importance de ces découvertes et sont donc réalisées davantage dans l'idée de servir de preuves que de supports d'analyse. L'acceptation des découvertes d'art pariétal intervient au début du XX^{e} siècle, notamment avec la publication d'Émile Cartailhac *Un mea culpa d'un sceptique* (Cartailhac, 1902) qui suit les découvertes des peintures et gravures des grottes des Combarelles et de Font-de-Gaume en 1901 par Henri Breuil, Denis Peyrony, et Louis Capitan. Elle marque un tournant dans l'histoire du relevé d'art pariétal.

Concernant les pratiques du relevé à proprement parler, nous verrons dans les sections suivantes que l'histoire est marquée par deux approches. Dans un premier temps, le relevé se présente comme un acte individuel et se base sur un contact direct avec la paroi. Par la suite, le relevé s'éloigne progressivement des parois jusqu'à s'en détacher complètement pour des raisons de conservation, et devient multidisciplinaire.

1.2.1.1. Le relevé direct

Le relevé est dit *direct* lorsque la technique utilisée implique un contact physique avec la paroi. L'utilisation des méthodes de cette catégorie, aujourd'hui considérées comme destructives, est à l'époque justifiée par la recherche d'exactitude et la volonté d'éviter les problèmes de projection soulevés par les méthodes indirectes.

À la fin du XIX^{e} siècle, Léopold Chiron est le premier à réaliser un relevé par estampage, pour la grotte Chabot (dont il ignore par ailleurs le caractère paléolithique des gravures). Il s'agit d'une technique permettant de prendre l'empreinte en négatif des stigmates d'une paroi en lui appliquant des feuilles préencollées et humidifiées (Dufayet, 2017). Une fois séchée et décollée du moule, l'épreuve rend bien compte des reliefs et des fines incisions. Néanmoins, cette technique s'adapte assez mal au contexte des grottes, où l'humidité ambiante allonge le temps de séchage déjà conséquent. Pour cette raison, la pratique de la prise d'empreinte au plâtre ou à l'argile lui succède, notamment avec les travaux d'Émile Rivière, mais ce procédé est encore plus destructif que l'estampage et laisse des stigmates, qui sont encore visibles aujourd'hui sur les parois qui en ont fait l'objet.

Figure 5 : À gauche : Breuil relevant une figure pariétale. À droite : André Glory relevant les gravures de Lascaux, accompagné d'Alain Roussot. Illustrations : (Groenen, 2018)

À partir du début du XX^{e} siècle, suite à la reconnaissance de l'art pariétal, Henri Breuil, aujourd'hui considéré comme « le pape de la Préhistoire », redéfinit la démarche du relevé. Dans son approche, le dessin n'est plus exploité dans l'idée d'illustrer, mais bien pour enregistrer ses observations relatives aux figures, avec l'objectif de reproduire le plus fidèlement possible les œuvres pariétales. Pour ne pas risquer d'abîmer les parois avec la technique de l'estampage, et à cause des difficultés restreignant l'usage de la photographie (déformations géométriques, manque de recul et obscurité des grottes), il opte pour le relevé par calque direct (Pigeaud, 2004). Il s'agit d'une autre technique de relevé par contact, qui consiste à décalquer la paroi en se servant de supports transparents sur lesquels l'observateur peut reproduire les tracés au crayon de couleur ou à l'encre en s'aidant d'éclairages adaptés (Figure 5 gauche). Les zones les plus fragiles sont relevées par croquis et photographies.

Au milieu des années 1950, l'abbé Glory, disciple de l'abbé Breuil, lui succède dans l'étude des grottes ornées, à commencer par celle de Lascaux. Pour une partie des parois, il opte, lui aussi, pour une technique de relevé direct. Il utilise pour cela une pellicule cellulosique imprégnée de matière caoutchoutée, posée en quelques points sur la paroi, limitant ainsi le contact avec celle-ci (Figure 5 droite).

1.2.1.2. Le relevé indirect : vers un éloignement des parois et une diversification des techniques

Progressivement, la nécessité de préserver ces sites fragiles pousse André Glory à s'éloigner de la paroi et à mettre en place des procédés d'enregistrement alternatifs, sans contact. Il réalise ainsi le relevé des figures peintes de Lascaux en plaçant une plaque de plexiglas semi-rigide à proximité des parois (tenue par deux personnes aux extrémités), sur laquelle il trace la projection des figures observées en transparence. Cependant, le fait de relever les figures à partir d'un support plan transparent pose très rapidement la question de la justesse géométrique de la représentation, et incidemment, des bonnes pratiques quant au choix du mode de projection.

Dans le cas du relevé direct, le problème ne se présentait pas en ces termes, car la méthode de Henri Breuil correspond peu ou prou à une projection sur une surface développable (Figure 6 gauche). Les effets de distorsions étaient alors limités aux éléments situés dans des concavités de la roche, lorsque le support utilisé ne permettait pas de s'adapter convenablement à la paroi. Dans le cas du relevé indirect, le support utilisé ne reproduit pas la morphologie de la paroi. Nous en revenons donc aux fondements de la géométrie projective : le releveur cherche à représenter en deux dimensions un objet qui en comporte trois. Il est d'ailleurs amusant de noter à quel point la méthode proposée par André Glory correspond à la définition de la perspective énoncée par Léonard de Vinci : *« La perspective n'est rien d'autre que de voir un objet qui se trouve derrière un verre lisse et transparent où se reflète tout ce qui se trouve derrière le verre ; tout se rend au point de l'œil par diverses pyramides et ces pyramides se coupent sur le verre »* (Péladan, 1910).

Or, selon les dimensions des panneaux à relever, l'observateur peut être amené à se déplacer, modifiant de fait le centre optique utilisé pour réaliser la projection. De cette manière, les figures relevées peuvent comporter des distorsions liées à ce décalage.

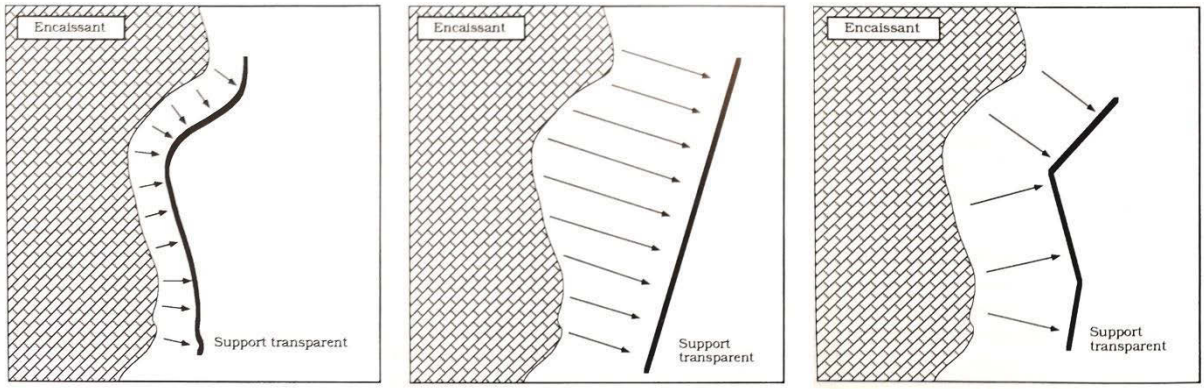


Figure 6 : Les différents modes de projection appliqués au relevé d'art pariétal. À gauche : projection sur une surface développable. Au centre : projection orthographique. À droite : projection polyédrique. Illustration : (Aujoulat, 1993a).

En réponse à ces préoccupations, différents chercheurs ont entrepris d'améliorer la technique d'André Glory en respectant certaines contraintes d'orthogonalité lors des relevés. La première approche consiste à réaliser le relevé suivant une projection orthographique (Figure 6 centre). Le support transparent est placé parallèlement au plan moyen de la paroi. Les différents éléments de la paroi sont relevés au moyen de rayons perpendiculaires au plan du support (point de fuite à l'infini). Cette technique montre cependant rapidement ses limites, lorsque le relief de la paroi est très prononcé. La seconde approche consiste à réaliser une projection polyédrique. La géométrie du support utilisé correspond à la morphologie de la paroi discrétisée en sous-éléments plans (Figure 6 droite). Pour chaque plan de ce support, les éléments sont relevés selon une projection orthographique. Cette méthode est cependant plus complexe à mettre en œuvre, et des problèmes peuvent surgir selon le fractionnement choisi.

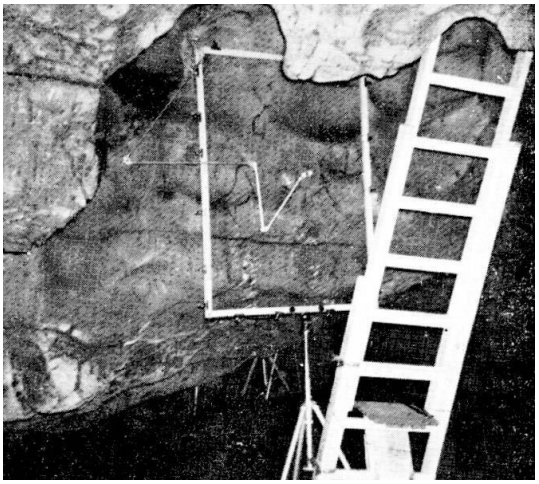


Figure 7 : Vue d'ensemble du système de visée utilisé pour le relevé indirect. Illustration : (Chabredier, 1966).

Dans les deux cas, cette nouvelle contrainte impose de perfectionner la technique d'André Glory, notamment au moyen de différents systèmes de visée. Par exemple, pour les relevés des grottes Chabot, Colombier, et d'Ebbou, Ludovic Chabredier améliore d'une part la structure en montant le cadre sur un châssis tubulaire rigide qui évite le recours à des assistants susceptibles de changer de position au cours du temps en provoquant des erreurs. D'autre part, il augmente la précision de la projection, en utilisant comme système de visée un bras articulé muni d'un viseur optique (Figure 7) afin de réduire les erreurs de parallaxe (Chabredier, 1966).

Dans cette même idée de limiter les déformations liées aux reliefs des parois, Jean Clottes et Denis Vialou développent un procédé de calque indirect dont le film transparent est fixé sur un support métallique articulé épousant approximativement les variations du relief (Aujoulat, 1987). Monique et Claude Archambeau, pour le relevé des gravures de la grotte des Combarelles, proposent quant à eux un procédé permettant d'utiliser les coordonnées cartésiennes (Archambeau, 1982). Pour cela, ils reprennent le principe du dispositif proposé par Chabredier, avec néanmoins un cadre de dimension fixe, fixé sur un trépied, et dont l'horizontalité, la verticalité, et le parallélisme sont réglés à l'aide d'un niveau à bulle et d'une boussole. Le système de visée est remplacé par un feutre lié à un faisceau lumineux.

D'autres variantes de ces méthodes ont été exploitées, avec diverses dimensions de cadre, ou avec la mise en place d'un échafaudage. Bien que sans contact, toutes ces méthodes de relevés sont placées très proches de la paroi. Si bien que dans certains cas, quand le contexte ou la configuration de la grotte l'exigent, une méthode hybride est parfois exploitée, en comprenant très ponctuellement des passages de relevé à caque direct, si le calque touche la paroi au moment du tracé. C'est notamment le cas pour le relevé de la grotte de l'Aldène par Léon Palès au début des années 1970.

La seconde partie du XX^{e} siècle voit émerger l'utilisation croissante de la photographie, non seulement comme illustration des observations, mais aussi comme support d'aide au relevé. En effet, si à la fin du XIX^{e} siècle la photographie est déjà employée dans le cadre des relevés de grottes ornées, elle ne sert d'abord que de « *document susceptible d'apporter la preuve nécessaire à l'authentification de ces manifestations pariétales* » (Aujoulat, 1993b), puis progressivement d'instrument secondant les documents issus de relevés directs à des fins d'illustration. La volonté de s'éloigner des parois aurait pu, par la suite, profiter davantage au développement de cette technique, mais les difficultés de mise en œuvre, particulièrement dans le contexte des grottes (photographie argentique, obscurité, manque de recul, etc.) et de traitement des données contraignent son utilisation, si bien qu'il faut attendre le milieu du XX^{e} siècle pour que ses nombreux potentiels soient plus sérieusement exploités. En 1897, Émile Rivière évoque à ce sujet les grandes difficultés qu'il a eues pour réaliser des photographies dans la grotte de La Mouthe.

> *« Nous ne sommes parvenus à reproduire ces gravures [par photographie], après de nombreux essais d'éclairage, qu'avec une intensité de lumière de 150 bougies environ et des poses de 5 et 6 heures. »* (Rivière, 1897)

La diversification des utilisations de la photographie est finalement consécutive à la commercialisation de films aux distributions spectrales plus étendues et à l'essor de l'informatique qui apporte de nouvelles possibilités de numérisation et de traitement (Groenen, 2018). La photographie ne résout certes pas le problème des déformations dues à la projection géométrique, mais sa capacité de stockage de l'information très au-dessus de toutes les autres techniques alors utilisées, les facilités apportées par le passage de l'argentique au numérique, et l'apport informationnel apporté par l'exploitation de nouvelles bandes spectrales en font un outil à fort potentiel. Précisons cependant qu'à aucun moment la photographie n'a pour vocation de remplacer le relevé, mais seulement de l'accompagner (Lorblanchet, 1993c). D'après Fuentes (2017), l'expérience menée par Jean Gaussens en 1955 à la grotte du Gabillou, constitue la première étude basée sur un relevé indirect sur photographies. La fragilité des parois ne permettant pas une proximité suffisante pour les autres procédés de relevé indirect, il réalise ses relevés *in situ* uniquement sur des calques posés sur des agrandissements de photographies. La proximité de la paroi permet alors de vérifier les observations réalisées sur les photographies et de gérer les problèmes de déformations.

L'exploitation de photographies dans le cadre des relevés ouvre de nouvelles possibilités d'étude et d'analyse. Ainsi, Annette Laming-Emperaire puis de André Leroi-Gourhan utilisent tous deux la photographie dans le cadre d'une approche analytique et statistique des œuvres pariétales, en exploitant les photographies doublées de croquis de lecture (Laming-Emperaire, 1962 ; Leroi-Gourhan, 1965). Aujoulat, (1987) exploite quant à lui la photographie pour proposer un procédé permettant de traduire le relief des parois, basé sur le principe des profils lumineux. Son idée est de matérialiser les profils

successifs d'une paroi par des traces lumineuses projetées sur le plan vertical, et de les photographier. Les données, ensuite saisies sur une tablette graphique et traitées informatiquement permettent d'extraire les coordonnées cartésiennes de chaque profil (Figure 8).

D'autres travaux s'intéressent à l'exploitation des rayonnements infrarouges (IR) et ultraviolets (UV) par la photographie pour l'analyse de peintures pariétales. Notons par exemple le travail de Fernand Windels à Lascaux, qui réalise à la fin des années 1940 des images infrarouges en noir et blanc d'une fresque pour établir une chronologie de la réalisation des différents éléments peints (Windels, 1948). Alexander Marshack, utilise, lui aussi, la photographie infrarouge, et la complète par de la photographie de fluorescence ultraviolette pour analyser les fresques du Salon Noir de la grotte de Niaux (Marshack, 1975).

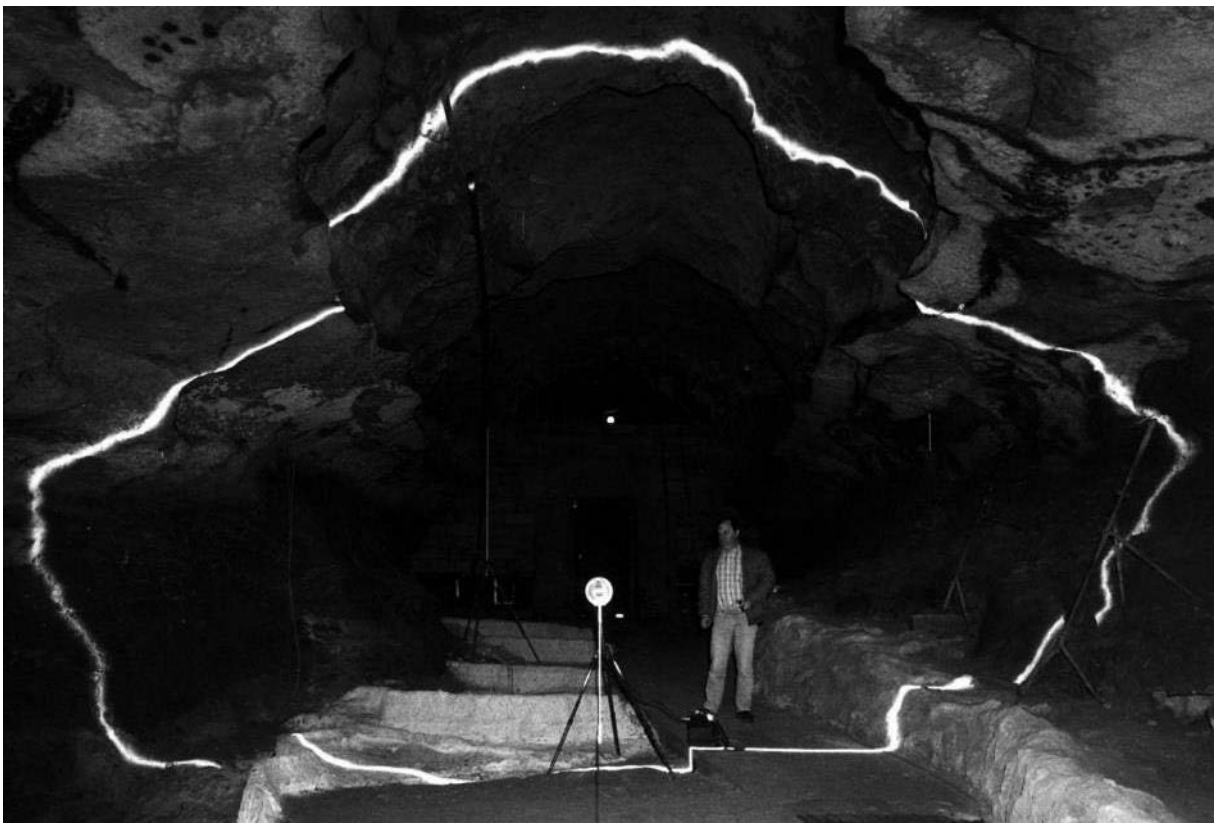


Figure 8 : Enregistrement d'une coupe de la Salle des Taureaux à Lascaux par la méthode des profils lumineux de Norbert Aujoulat. Illustration : © Ministère de la Culture / Centre National de la Préhistoire / Norbert Aujoulat.

Finalement, les travaux de Léon Palès pour l'étude des éléments mobiliers de la grotte de la Marche, constituent une contribution importante dans la mise en place d'un *workflow* multitechniques. En effet, la complexité du site – à travers la grande quantité de gravures, leur superposition, leur diversité technique, et la richesse des motifs – impose l'utilisation conjointe de plusieurs techniques, comprenant la photographie, l'estampage et le calque direct (Aujoulat, 1993b). Dans ce procédé d'analyse et de relevé, l'usage d'agrandissements de photographies permet de réaliser le relevé des pièces de petites dimensions. Transposé et adapté au contexte souterrain, ce processus d'analyse et de relevé constitue les fondements des démarches actuelles.

Cette évolution des pratiques du relevé montre bien que la nature et la complexité des sites étudiés, cumulée à la nécessité de limiter les contacts directs, ont progressivement imposé une diversification des techniques, comprenant différents supports et outils de mesures. La multiplication des outils et

techniques permet dans une certaine mesure de pallier les problèmes causés par la complexité des sites (notamment morphologique). Cependant, cette diversification des approches de relevé soulève de nouvelles interrogations concernant les *bonnes pratiques* à respecter et plus globalement, concernant les finalités du relevé. Nous tâcherons de les expliciter dans la section suivante.

1.2.2. La complexité du relevé d'art pariétal

1.2.2.1. La subjectivité de l'observateur

La contribution de Henri Breuil est assez caractéristique des différentes problématiques posées par la pratique du relevé. Ses travaux, dotés d'une qualité graphique certaine, sont très rapidement associés à un *style Breuil*, du fait des traits utilisés qui contribuent à faire de ses représentations des œuvres d'art à part entière (Figure 9). L'esthétisme de ses travaux est tel que les relevés, dont l'auteur est identifiable en un coup d'œil, en deviennent autosuffisants. Cette tendance ne se limite pas aux relevés de Breuil et s'observe encore sur certains travaux plus contemporains. Fritz et al. (2016) par exemple, désignent même le releveur par le terme « *the artist* ».

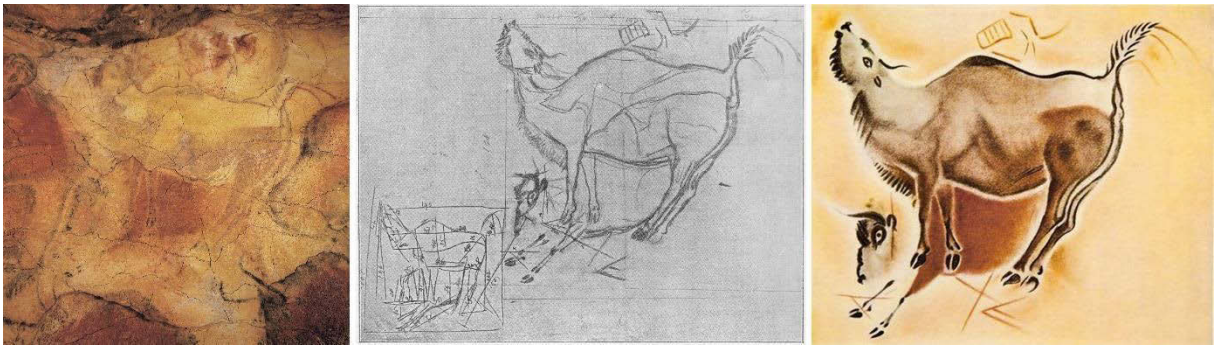


Figure 9 : Relevé du bison mugissant du Grand Plafond de la grotte d'Altamira (Cantabrie) par Henri Breuil. À gauche : paroi de la grotte. Au centre : esquisse au crayon. À droite : reproduction au pastel de la figure relevée. Illustration : (Groenen, 2018).

Par ailleurs, malgré sa recherche de justesse dans le tracé des contours, les relevés de Breuil sont profondément subjectifs. Pigeaud (2004) considère même qu'ils imposent, plus qu'ils ne suggèrent, une interprétation de l'œuvre. L'erreur de lecture du fameux *sanglier d'Altamira*[9] illustre bien le choix personnel de l'auteur de compléter des informations manquantes.

De surcroît, en qualifiant Lascaux de « *chapelle Sixtine de la préhistoire* » (Windels, 1948), Henri Breuil ne confère pas seulement un caractère sacré à la grotte, mais il exerce aussi une certaine appropriation qui implique des codes de lecture hérités de la peinture classique, art dominant de la fin du XIX^{e} siècle[10]. Il est alors difficile d'imaginer comment l'observateur peut ensuite se dégager d'une influence si forte, et conserver une certaine objectivité dans la transcription des phénomènes observés. Or d'après Groenen (1999), dans la mesure où le relevé a pour vocation de constituer un document scientifique, il

[9] En relevant l'une des figures du Grand plafond de la grotte d'Altamira, Breuil interprète une série de fissures naturelles comme le tracé d'un groin, ce qui le conduit à compléter le dessin d'informations supposément manquantes et assigner la figure comme correspondant à un sanglier (Pigeaud, 2004). Une étude ultérieure démontre finalement qu'il s'agit d'un bison.

[10] La légitimité ou non de ce rapport à une connotation moderne de l'art est une préoccupation récurrente dans l'étude des grottes ornées, si bien que (Fritz, Tosello, 2007) sentent nécessaire de préciser « *We use the term art in this paper in order to avoid cumbersome phrases such as "graphic expression". In doing so, we are not implying that modern connotations of art should be applied to the prehistoric world.* » !

se doit d'offrir les garanties de l'objectivité. À ce sujet, Ludovic Chabredier écrit que le relevé exige *« la plus rigoureuse précision, excluant toute part d'interprétation ou de fantaisie […] Il ne s'agit pas aussi d'en faire une œuvre d'art, "plus belle que nature", mais seulement de restituer le caractère et l'intégralité de l'original »* (Chabredier, 1966).

Au-delà du cas des relevés de Breuil, emblématique du fait de son statut de précurseur et de ses nombreuses années d'expertise, la problématique de la juste position de l'observateur face à son objet d'étude se traduit dans le relevé par un arbitrage difficile entre la justesse géométrique des éléments relevés, la transcription d'éléments d'assignation incertaine, et l'expression graphique du document produit. Cette subjectivité soulève également des enjeux en termes de traçabilité, car la connaissance des choix opérés par l'auteur du relevé conditionne ses potentielles exploitations ultérieures.

1.2.2.2. La place de la figure

Une seconde interrogation concerne la place à accorder à la figure. L'idée de réaliser un relevé esthétique place d'office la figure pariétale au centre des préoccupations. Dans son interprétation, Breuil opère une sélection qui exclue complètement le contexte, c'est-à-dire la paroi, le relief, la forme de la roche, ou encore les altérations et autres stigmates. En témoignent par exemple les différences entre son relevé du Mégacéros de Pech-Merle et celui réalisé par Michel Lorblanchet soixante ans plus tard (Figure 10).

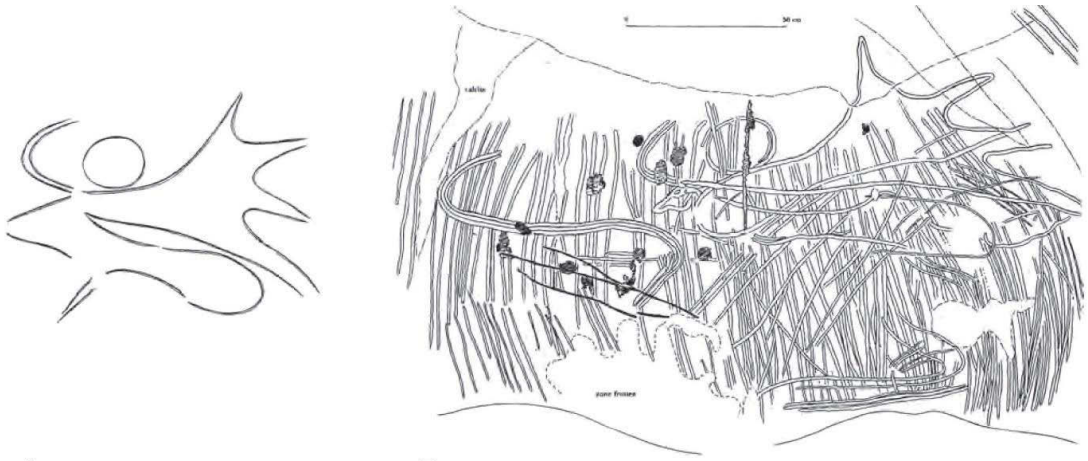


Figure 10 : Deux relevés du Mégacéros gravé de Pech-Merle. A. Celui de Breuil, B. Celui de Lorblanchet. Illustrations : (Pigeaud, 2004).

Il semble clair qu'une sélection qui limite le relevé à la figure ne peut tenir compte des relations qui lient le tracé à la paroi, elle ne peut traduire les éventuels choix opérés par l'homme préhistorique pour *jouer* avec les volumes, elle occulte donc d'emblée un certain nombre d'informations qui, si elles ne sont pas aisément interprétables, contribuent tout de même très certainement à la richesse de l'œuvre elle-même.

> *« Or, ceux-ci [les éléments du contexte pariétal] constituent bien souvent une donnée importante pour la compréhension des œuvres. Le calque d'une représentation d'aurochs de la grotte d'El Castillo (Cantabrie), réalisé par Breuil vers 1910, reproduit avec une précision satisfaisante la silhouette de l'animal, mais il néglige le jeu de fissures qui situe l'animal dans un espace structuré et lui donne l'assise sur laquelle il se déplace. »* (Groenen, 2018) (Figure 11)

Figure 11 : À gauche : Relevé d'un auroch de la grotte d'El Castillo réalisé par Henri Breuil. À droite : l'auroch dans son contexte. Illustration : (Groenen, 2018)

Les choix méthodologiques de Breuil s'expliquent notamment par les théories de l'époque qui, en considérant l'art pariétal comme un art primitif, excluent l'idée que la composition des œuvres sur les parois n'est pas due au hasard et que leur organisation spatiale peut être significative (Sauvet, 1993). Cependant, la question de la place de la figure et de ses limites persiste, aujourd'hui encore.

Alors comment enregistrer l'art pariétal ? Quels éléments méritent d'être relevés ? Il s'agit d'une question fondamentale pour le relevé pariétal, car la place accordée à la figure influence profondément l'approche de la lecture de la paroi. Dans un cas, la figure peut être considérée comme l'œuvre en elle-même, et la paroi en constitue le support voire le décor. Du point de vue du relevé, cette vision se traduit par l'isolement de la figure, et incidemment, l'occultation d'un éventuel travail de composition. Dans l'autre cas, la paroi peut être considérée comme partie prenante de l'œuvre. Du point de vue du relevé, cette vision pose le problème inverse, car le relevé de la paroi dans toutes ses dimensions (géométriques, matérielles, visuelles, etc.) pousse vers une certaine exhaustivité qui pose des problèmes d'intelligibilité dans la représentation. La sélection des éléments relevés et leurs modalités de représentation constituent des enjeux qui rejoignent ceux du relevé du bâti.

1.2.2.3. La contextualisation des éléments relevés et les conventions graphiques

À partir de la seconde moitié du XX^{e} siècle, l'émergence de théories reconnaissant un travail de composition à l'échelle des panneaux ou des grottes entières apporte un renouveau dans l'étude de l'art pariétal (Lorblanchet, 1993b ; Sauvet, Tosello, 1998 ; Brot, 2005). Georges Sauvet souligne alors la nécessité de rendre compte de l'organisation spatiale des figures en cessant de les représenter de manière isolée, et en les contextualisant à l'échelle des panneaux, dont la morphologie doit être restituée. Les panneaux, quant à eux, doivent être situés dans une représentation permettant de *« restituer aux ensembles graphiques leur réalité volumique et leur contexte spatial intégral »* (Sauvet, 1993).

Néanmoins, cette avancée apporte des interrogations qui découlent de la difficulté de reproduire la surface tortueuse des parois de grottes. À ce sujet, les préhistoriens ne s'accordent pas sur les conventions de représentation à adopter, chacune présentant ses atouts et ses limites. Ainsi, certains prennent le parti de représenter la paroi par différentes techniques empruntées au dessin. C'est le cas par exemple pour le relevé d'un panneau de la grotte Huchard par Chabredier et Ayrolles, qui représentent la paroi par des pointillés et des hachures (Figure 12 gauche).

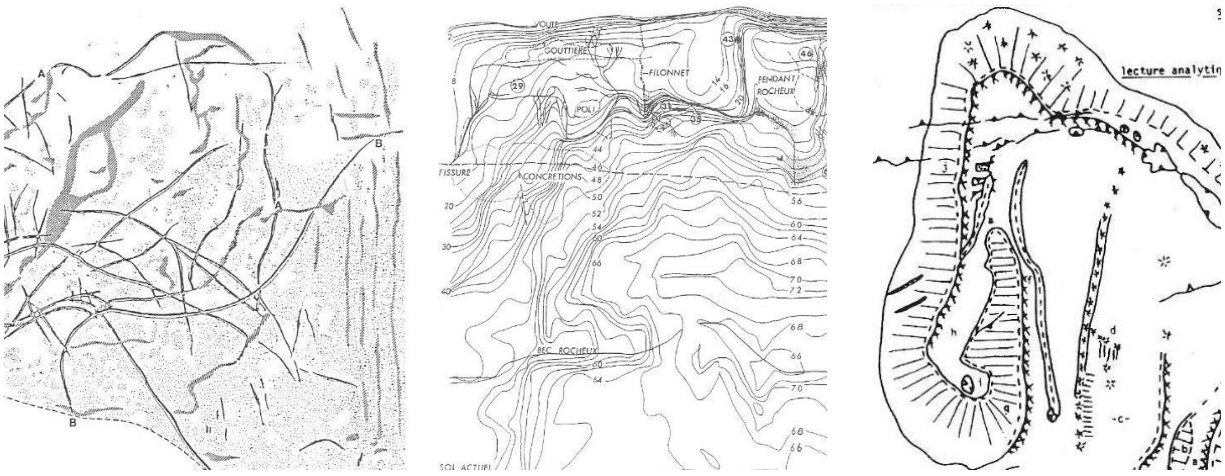


Figure 12 : Différentes approches pour la représentation du relief. À gauche : représentation des concrétions en pointillés, extrait du relevé de la grotte Huchard par L. Chabredier et P. Ayrolles. Illustration : (GRAPP, 1993). Au centre : représentation par courbes de niveau, extrait du relevé de la grotte Carriot par M. Lorblanchet. Illustration : (GRAPP, 1993). À droite, utilisation de conventions graphiques de relevé analytique, extrait du relevé d'un mammouth de la grotte de Domme par B. et G. Delluc. Illustration : (Delluc et al., 1986b).

D'autres choisissent d'adopter un rendu plus schématique comprenant des conventions graphiques prédéfinies. Ces approches presque cartographiques, bien que plus précises et plus complètes, s'avèrent difficilement lisibles du fait de la quantité d'informations. Par exemple, pour le relevé de la frise noire de la grotte de Pech-Merle, Lorblanchet, opte pour une représentation de la paroi par courbes de niveau (Figure 12 centre). Delluc et al. (1986a) proposent quant à eux un ensemble de conventions graphiques de relevé analytique des surfaces rocheuses, destinées notamment à matérialiser les accidents de la paroi (Annexe 1, Figure 12 droite). Nous pourrions aussi citer un exemple en dehors du territoire français avec le relevé du panneau principal de la grotte de Lluera (Fortea Pérez, 1989), qui, dans le même esprit, comporte de nombreuses informations concernant le support, comme le relief ou les altérations.

Figure 13 : Utilisation d'une photographie comme fond graphique du relevé. Relevé d'un fragment de la grotte de Teyjat par P. Paillet, Photographie de J.-C. Domenech. Illustration : (Ministère de la Culture, CNRA, 2017)

Bien que ces rendus soient riches en informations concernant la morphologie de la paroi et ses stigmates, ils demandent de la part de l'observateur une certaine capacité d'abstraction pour se représenter mentalement l'ensemble et dissocier ce qui relève de la paroi et ce qui relève des traces anthropiques. Enfin, pour favoriser l'intelligibilité du document produit, une approche répandue consiste à utiliser une photographie comme *fond* du relevé afin d'illustrer le contexte des figures. Notons cependant que cette approche ne constitue pas une analyse de la paroi à proprement parler (Figure 13).

1.2.2.4. La naissance de deux familles de relevé

Comme l'expliquent Fritz et Tosello (2007), cette complexité des parois et la difficulté de les représenter en 2D ont donné lieu à de nombreux débats concernant la méthodologie du relevé, particulièrement concernant l'usage de la photographie. Si aujourd'hui certains spécialistes l'estiment indispensable pour l'analyse, d'autres condamnent cette méthode et soutiennent que les figures doivent être étudiées sans intermédiaire. Pourtant, les deux ne semblent pas incompatibles.

> *« [C'est] une autre raison pour laquelle la composition de l'art paléolithique est presque toujours ignorée : contrairement à toutes les autres formes d'art visuel que nous connaissons, l'art rupestre est créé sur une surface rugueuse et irrégulière, n'est jamais (ou presque jamais) réalisé sur une vaste surface plane [...] Lorsque vous essayez de reproduire ces images sur papier, soit avec un dessin ou une photographie, vous êtes confrontés à ces difficultés insolubles. Tout est étiré, tordu, déformé dans tous les sens du terme. Nous sommes presque toujours obligés d'isoler une image du groupe dont elle fait partie parce qu'elle occupe une surface concave, convexe ou très irrégulière. Parfois, la photographie ne peut montrer qu'un fragment d'image. Peu à peu, nous nous sommes habitués à penser l'art paléolithique en fragments et nous oublions le tout. »* (Laming-Emperaire, 1962)

De ce point de vue, et tenant compte des moyens de l'époque, il est vrai que la technique de Henri Breuil, qui revient à une projection sur une surface développable de la paroi permet théoriquement un report parfait des figures sans distorsions. Le contact avec le site permet à l'observateur de garder un point de vue global sur le site, tout en réalisant le relevé de ses différents éléments. L'éloignement des parois, pleinement justifié par la nécessité de préserver et conserver les œuvres, prive alors les chercheurs (jusqu'à l'essor de la numérisation 3D) de cette méthode efficace de report. La prise de conscience progressive de la complexité des œuvres et des relations qui les lient à leur site et à leur support complique encore plus la quête d'une méthode *idéale*. Finalement, les préoccupations grandissantes concernant la juste part d'objectivité à adopter semblent alimenter la difficulté à trouver une méthode pleinement satisfaisante, tant du point de vue géométrique que de l'exhaustivité informationnelle et de l'intelligibilité des représentations.

> *« La représentation plane d'une œuvre pariétale, épurée des détails extrinsèques de l'image paléolithique elle-même, devrait d'une part traduire ce que, d'un point donné de la cavité, notre œil a enregistré et, d'autre part, permettre une mesure du tracé matérialisant le dessin. En d'autres termes, le relevé doit exprimer ce que l'organe visuel de l'homme préhistorique a saisi et autoriser l'estimation de son geste. Ces deux aspects sont confondus dans l'éventualité d'une paroi plane, ils divergent lorsque le support s'écarte géométriquement du plan. »* (Aujoulat, 1993a)

Si cet état de fait semble à première vue souligner la coexistence nécessaire de deux démarches différentes (et complémentaires) répondant à ces deux aspects, en réalité il semble avoir guidé l'émergence de deux tendances globales, deux *écoles* de relevé qui perdurent aujourd'hui encore. Ces deux approches, que le CNP désigne comme « *relevé analytique plastique* » et « *relevé analytique technique* » prennent position quant aux questions de la place de l'image dans l'analyse et du traitement des informations menant à l'interprétation. La première considère la figure comme l'élément central conduisant à l'interprétation. La seconde considère le contexte de la paroi comme point de départ et la figure comme la finalité à atteindre.

1.2.2.4.1. Le relevé analytique plastique

Dans la lignée méthodologique de la pratique du relevé de Henri Breuil, l'approche dite du relevé analytique plastique se base sur un enregistrement hiérarchisé d'éléments observés. Il s'agit d'un processus de sélection et traduction graphique des stigmates des parois, avec pour élément central la figure pariétale. L'objectif est de rendre compte de la complexité des faits observés tout en facilitant la lecture et la perception des œuvres. Ainsi, les données caractérisant le support peuvent être enregistrées mais ne sont pas représentées de manière systématique, dans l'idée d'éviter une accumulation d'informations pouvant conduire à une confusion visuelle (CREAP, 2018).

> *« Bien que l'impact d'autres facteurs (comme la morphologie des supports, la présence de griffades) soit également pris en compte, une hiérarchie visuelle favorisant les manifestations anthropiques doit être privilégiée. Cela implique une indispensable et préalable sélection des données. »* (Azéma et al., 2012)

Cette méthode rend donc bien compte de la qualité graphique des œuvres, et, comme dans le cas des travaux de Breuil, les documents qui en résultent peuvent s'avérer très esthétiques, voir par exemple le relevé du panneau des Chevaux par Carole Fritz et Gilles Tosello (Figure 14). Majoritaire parmi les études existantes des œuvres pariétales, elle s'appuie généralement sur la technique du relevé indirect sur des montages photographiques imprimés. Sa chaîne opératoire est généralement composée de cinq étapes (Fuentes, 2017).

1) Prise de connaissance du terrain, première observation des parois : les panneaux à étudier sont choisis puis déchiffrés, afin de repérer et identifier les figures. Celles-ci sont retranscrites avec des croquis de lecture accompagnés de descriptions, et enregistrées dans des fiches de terrain. Il s'agit d'observer, et comprendre le terrain, de le mémoriser avant de l'étudier.

2) Campagne d'acquisition photographique des parois sélectionnées : les photographies sont réalisées en respectant certaines contraintes techniques liées au site, au contexte de l'étude et aux objectifs (par exemple l'orthogonalité et la distance à la paroi, la lumière, le recouvrement, etc.).

3) Traitement informatique des photographies : pour chaque panneau étudié, les images sont traitées et assemblées en une orthophotographie imprimée selon une échelle déterminée en fonction des objectifs et besoins de l'étude.

4) Relevé *in situ* des panneaux : face aux parois, l'observateur retranscrit ses observations sur des calques superposés à l'orthophotographie. Ce travail se fait face aux images originales, l'auteur du relevé peut donc s'appuyer sur les photographies tout en confrontant ses observations à la réalité. Cette étape est considérée comme la plus importante.

5) Traitement, numérisation et mise au propre : en laboratoire, les données relevées sont nettoyées, homogénéisées et mises au propre selon une charte graphique établie. C'est à cette étape que s'opère la majeure partie du travail de hiérarchisation et de sélection (bien que des sélections soient déjà nécessairement opérées lors des étapes *in situ*). L'objectif est d'opérer une réduction des données afin de rendre intelligible le document final.

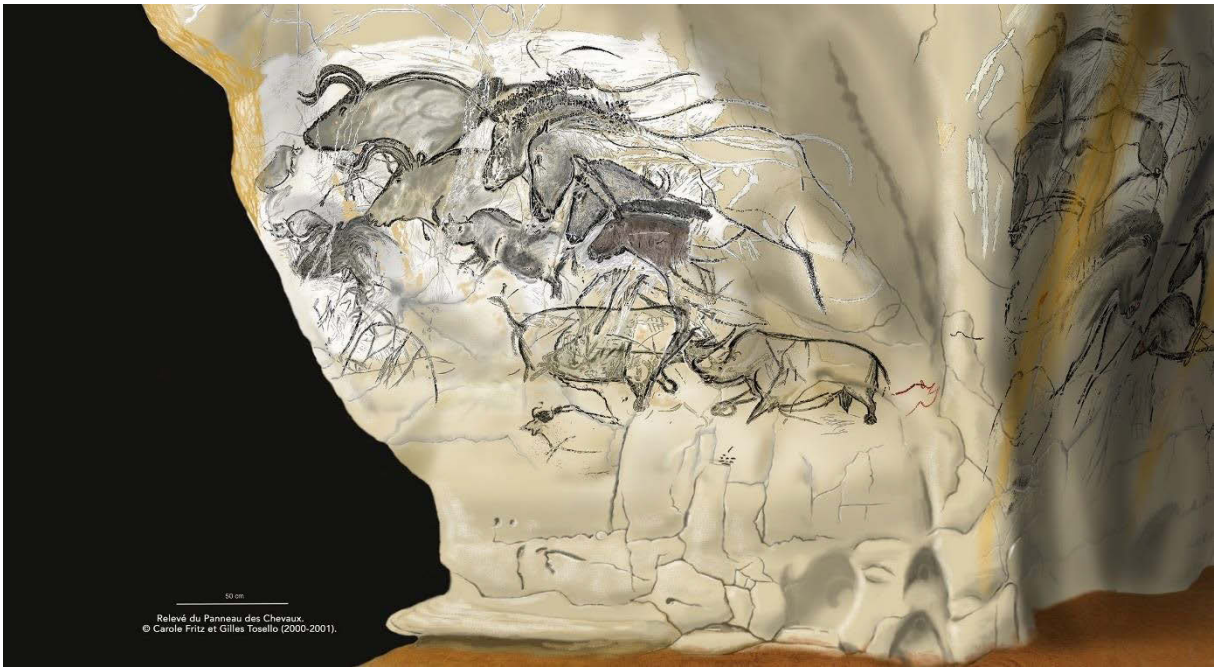


Figure 14 : Relevé analytique plastique du panneau des Chevaux de la grotte Chauvet-Pont d'Arc par Carole Fritz et Gilles Tosello. Illustration : © C.Fritz et G.Tosello.

Cette méthode présente l'avantage d'apporter une grande lisibilité des traces anthropiques, grâce à la sélection préalable des données et à leur hiérarchisation, ainsi qu'au traitement graphique du document, souvent esthétique. En contrepartie, cette recherche de clarification centrée sur l'œuvre pariétale a tendance à reléguer la paroi en un simple support illustratif, un fond plastique, effaçant de ce fait de précieuses informations de contexte. Par exemple, la taphonomie des parois peut apporter d'importantes indications concernant la chronologie des réalisations. Si elle est bien exploitée pour l'analyse de la paroi, elle n'est cependant pas nécessairement retranscrite sur le document final ce qui limite les possibilités d'interrogeabilité ultérieures du document. Par ailleurs, l'absence d'informations concernant la charte graphique adoptée pour ces illustrations confère à l'ensemble des relevés une sensation d'hétérogénéité, qui peut rendre difficile leur lecture pour un public n'ayant pas participé à leur réalisation.

Plus globalement, la subjectivité assumée de la méthode et la sélectivité des informations représentées ne semblent pas permettre d'exploiter les documents produits autrement que pour lire les propos scientifiques de leurs auteurs, établis au moment même de leur réalisation. De ce fait, une fois le relevé achevé, il semble difficile de l'utiliser pour croiser des observations à la lumière de nouvelles données.

1.2.2.4.2. Le relevé analytique technique

Le relevé analytique technique[11] naît de la volonté d'intégrer chaque représentation dans son contexte physique et spatial, considérant que les fissures, les reliefs naturels, les écailles et autres stigmates sont autant d'informations qui participent aux œuvres pariétales, ou du moins, à leur connaissance.

Dans la lignée de l'approche presque cartographique de Michel Lorblanchet dans les années 1970, le relevé technique se distingue du relevé plastique par le statut accordé à la figure. La paroi est au cœur

[11] Ce terme a initialement été proposé par Fritz et Tosello (2007). Les auteurs parlent en revanche de « relevé de lecture » pour se référer au relevé plastique.

de la démarche, la figure elle-même n'apparaît qu'à la fin du relevé comme le résultat du processus d'analyse et non plus comme son entrée. Cette approche se distingue également par la volonté d'apporter une certaine objectivité et exhaustivité dans la démarche du relevé, en s'attachant à la cartographie globale de la paroi et de tous les éléments qui la constituent, sans distinction ni hiérarchie (Figure 15).

Parmi eux, les figures sont donc considérées selon leur nature de vestiges, au même titre que les autres stigmates. L'idée sous-jacente de cette démarche est que la mise en relation des différents registres identifiés et analysés mène à la figure et à sa compréhension dans toute sa complexité. Cette démarche est fondamentalement pluridisciplinaire, et bien qu'elle tende à atteindre un certain degré d'objectivité, il est clair qu'elle implique tout de même une sélection des données et leur interprétation, qui doivent néanmoins faire consensus au sein de l'équipe.

Sa chaîne opératoire est généralement composée comme suit :

1) Visite de la grotte, repérage des parois : ce premier contact permet de se familiariser avec le site et ses parois, d'amorcer des discussions croisées, de comprendre les éléments qui composent les parois (morphologie, nature des roches, identification des types de traces, des techniques, etc.) afin de préciser les problématiques de l'étude et de proposer une stratégie pour le relevé.

2) Campagne d'acquisition photographique des parois et de la cavité : les photographies sont réalisées en respectant certaines contraintes techniques liées au site, au contexte de l'étude et aux objectifs (par exemple l'orthogonalité et la distance à la paroi, le recouvrement, etc.). Les photographies sont complétées par différentes techniques d'imagerie selon les besoins de l'étude (par exemple macrophotographies, imagerie multispectrale, etc.), et accompagnées d'une série de photographies en haute définition en lumière rasante. Ces images sont archivées en vue de servir d'état des lieux de la cavité pour une date donnée.

3) Traitement informatique des photographies : Les images sont traitées et utilisées pour produire des orthophotographies pour chaque paroi. L'équipe détermine les zones à relever et la technique la plus adaptée aux besoins de leur étude : relevé sur orthophotographie ou relevé par vidéoprojection[12].

4) Relevé *in situ* des parois : face aux parois, l'équipe dispose de tablettes numériques tactiles, sur lesquelles sont chargées les photographies en lumière rasante. Ces tablettes sont utilisées comme supports pour la réalisation du relevé. Une charte graphique est adoptée par toute l'équipe, pour être appliquée à l'ensemble des zones relevées. Elle réunit l'ensemble des codes graphiques utilisés pour chaque stigmate. Le relevé est opéré par au moins trois personnes devant la paroi, qui représentent de manière vectorielle les stigmates observés et les regroupent par calques.

Dans un premier temps sont relevés de la manière la plus exhaustive possible les états de surface, anciens ou récents, les apports et enlèvements et les reliefs naturels. Dans un second temps sont relevés les éléments anthropiques récents (post-paléolithiques) comme des coups ou des graffitis. Enfin, l'équipe termine par le relevé des vestiges anthropiques anciens (paléolithiques), qu'ils fassent ou non partie d'une figure. Le relevé est terminé lorsque l'ensemble des éléments identifiés ont été cartographiés et que leur interprétation est communément acceptée par l'équipe.

[12] Voir section 3.3.5.

5) Traitement et mise au propre : en laboratoire, les codes couleur et symboles utilisés sont corrigés et harmonisés. Un calque de prise de note peut être créé, afin de préciser les informations concernant le relevé et les choix opérés (notamment le zonage). L'analyse se poursuit par la mise en relation de différents calques par superposition des calques concernés et masquage des autres. Cette étape permet de produire différents supports pour les publications : le relevé exhaustif, le relevé rendant compte de l'état des parois, ou encore des relevés sélectifs selon le discours.

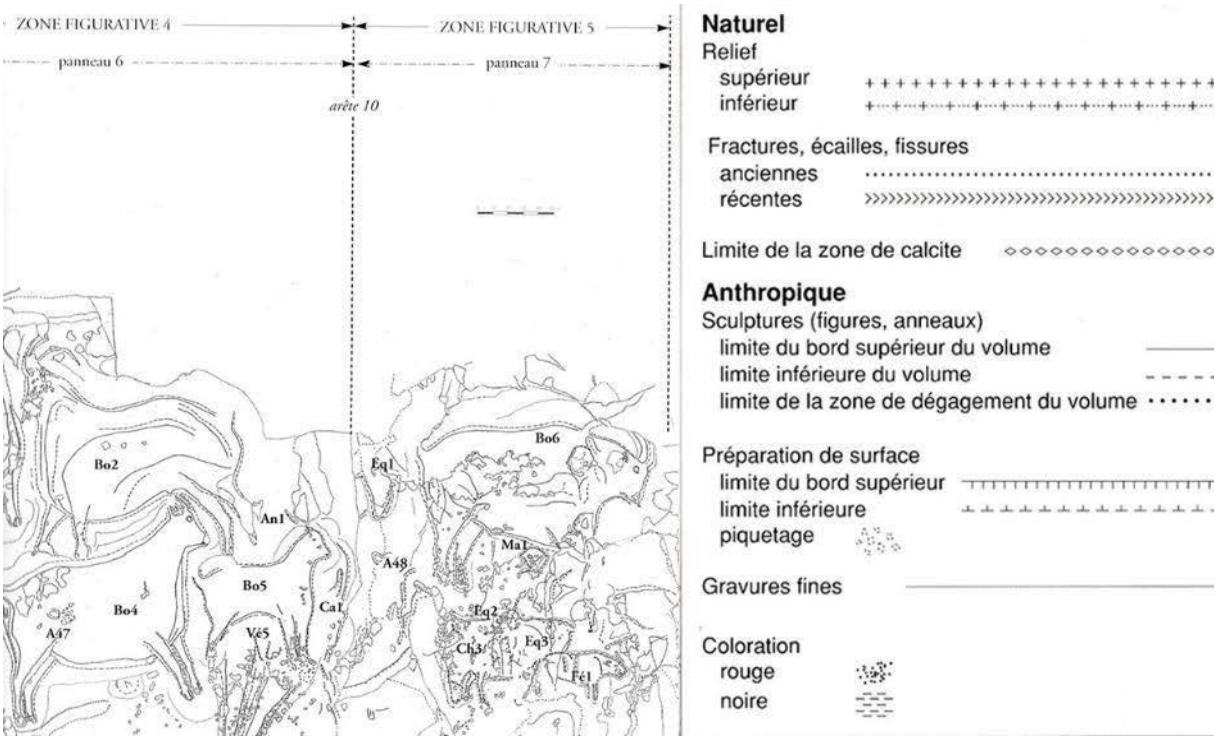


Figure 15 : Relevé analytique technique du panneau des Bouquetins du Roc-aux-Sorciers par Ludmila Iakovleva et Geneviève Pinçon. Illustration : (Fuentes et al., 2019)

La répartition par calques proposée par cette méthode de relevé permet de filtrer les niveaux d'informations (affichage sélectif) ou de les superposer pour croiser des observations de différentes natures. Par exemple, il est possible de n'afficher que les gravures, ou de choisir de les afficher conjointement avec les reliefs naturels, ou encore de décider de superposer le calque présentant les gravures avec l'ensemble des calques qui représentent des dégradations de la paroi. Les représentations qui en résultent sont autant de possibilités d'illustrer un propos scientifique.

Par ailleurs, l'utilisation de tracés vectoriels permet de revenir à tout moment sur les formes et leur assignation. Ainsi, à la lumière de données nouvelles, l'équipe peut décider de corriger certains éléments qui auraient fait l'objet d'une erreur d'appréciation (par exemple, une trace initialement considérée comme un vestige anthropique peut se révéler n'être qu'un élément naturel de la paroi). En termes d'extraction d'informations, cette technique semble donc plus avantageuse que la technique du relevé analytique plastique. En revanche, elle peut s'avérer bien moins lisible et moins efficace en termes graphiques, d'une part du fait de l'accumulation d'informations, et d'autre part car elle ne se s'attache pas à exprimer la *réalité plastique* de l'œuvre étudiée.

Notons néanmoins le cas du relevé de la frise sculptée du Cap-Blanc (Bourdier et al., 2011) pour lequel un calque présentant des ombres a été ajouté pour donner du relief aux volumes illustrés (Figure 16). Si

graphiquement le résultat n'approche pas tout à fait de l'esthétisme d'un relevé analytique plastique, cette expérience démontre tout de même une certaine volonté de faire converger les deux méthodes, afin de bénéficier de leurs qualités respectives.

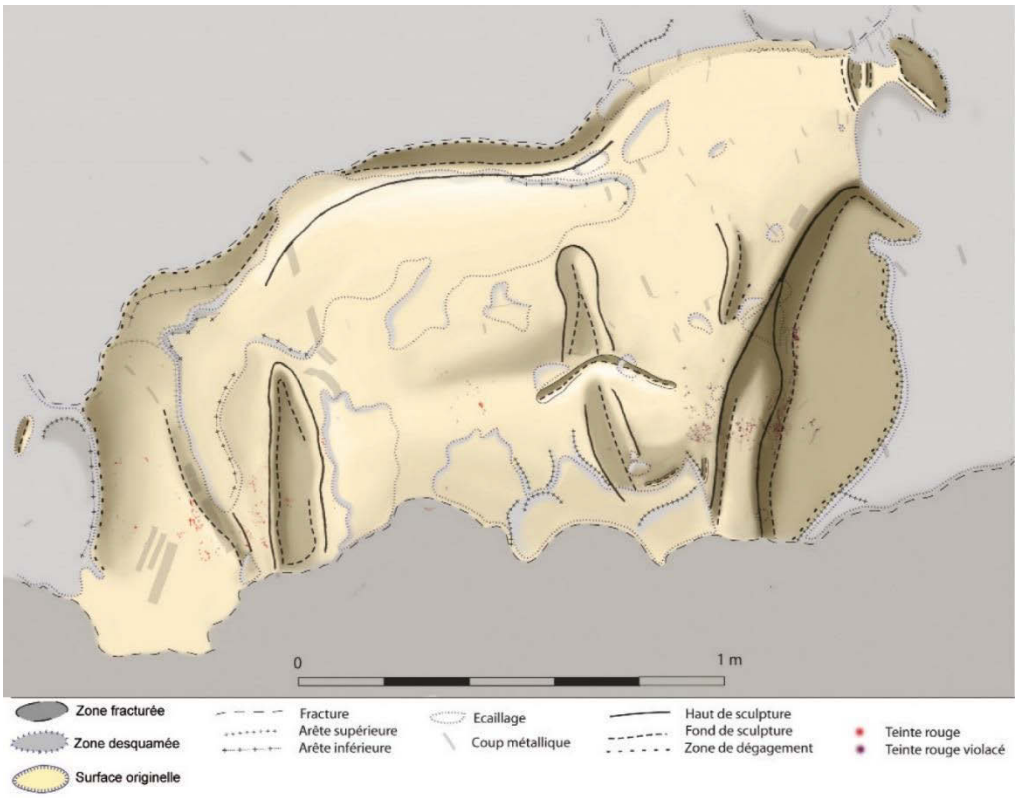


Figure 16 : Relevé analytique technique (ombré) de l'entité graphique n°1 de la frise sculptée du Cap-Blanc, par Aurélie Abgrall, Camille Bourdier, Olivier Huard, Éric Le Brun et Geneviève Pinçon. Illustration : (Bourdier et al., 2011).

1.3. Bilan : Les exigences des méthodes de relevé

Nous avons pu voir que la naissance de deux familles de relevé est profondément liée à la complexité des objets d'étude, dont la traduction graphique en deux dimensions pose de nombreuses difficultés, particulièrement depuis le recul vis-à-vis des parois. En l'absence de méthode satisfaisante pour la saisie du support, la morphologie des parois, souvent tortueuses, impose des arbitrages difficiles entre la justesse géométrique et l'expression graphique des figures relevées.

La reconnaissance d'un travail de composition et incidemment, d'un lien profond entre les figures et leur support pose aussi la question de l'exhaustivité du relevé : où l'observateur doit-il arrêter son intérêt ? Où s'arrête l'œuvre ? Dans le cas du relevé plastique, le parti pris est celui de se concentrer sur la figure, considérant la paroi comme un support illustratif, voire un décor. Dans le cas du relevé technique, il s'agit de relever la plus grande quantité d'éléments possible. Mais la recherche d'exhaustivité dans un relevé pose alors une contrainte supplémentaire, car elle suppose un travail de codification graphique. De fait, s'il était possible de tout relever sans exception, une représentation réaliste serait problématique, car le relevé n'apporterait alors aucun élément de compréhension supplémentaire que l'objet réel lui-même :

> *« On peut en tirer la leçon que, sans réduction, un modèle redoublant totalement la réalité n'apporte aucune connaissance. »* (Boudon, 1987)

Par conséquent, nous pouvons comprendre que la sélection des éléments signifiants est un élément clé du relevé, ainsi que leur hiérarchisation au sein d'une représentation graphique codifiée. En cela, nous pouvons faire le rapprochement entre le relevé d'art pariétal et le relevé du bâti tel que théorisé par Saint-Aubin. Dans les deux cas, le relevé comporte bien trois moments interreliés :

- La **description**, c'est-à-dire le moment de la sélection des éléments et de la mesure. C'est une étape décisive concernant la question de la précision géométrique, mais aussi de l'exhaustivité.
- La **figuration**, c'est-à-dire le moment de l'interprétation des éléments sélectionnés, leur hiérarchisation en vue de leur codification graphique. Il conditionne les possibilités d'enrichissement ultérieur et le potentiel d'interrogation du document résultant.
- La **représentation**, c'est-à-dire la transcription graphique des éléments interprétés en un document intelligible communicable.

Les différentes approches du relevé se distinguent alors entre elles par la pondération accordée à l'importance de ces trois moments. Par exemple, dans le cas du relevé analytique plastique, la figuration et la représentation priment largement sur la description. Dans le cas du relevé analytique technique, la description est au contraire le moment le plus important. Cette pondération a une incidence notable sur les caractéristiques du document produit.

Afin d'offrir une vision synthétique des spécificités des différentes approches du relevé, nous avons décomposé les trois moments de Saint-Aubin en différents aspects intimement liés à leurs définitions et cités de manière récurrente dans l'état de l'art. Cela nous permet de constituer des critères pour représenter schématiquement et comparer les approches et mieux comprendre ce qui les lie ou les distingue. Nous conserverons ces mêmes critères tout au long de ce travail.

Ainsi, nous avons pu constater que, dans le cas du relevé analytique plastique, l'accent mis sur la figuration et la représentation bénéficie à la codification graphique et l'intelligibilité du document, mais nuit à son potentiel d'interrogation et à ses possibilités d'enrichissement ultérieur (Figure 17). Dans le cas du relevé analytique technique, l'accent mis sur la description bénéficie au potentiel d'interrogation, mais nuit nettement à l'intelligibilité du document. Le choix de minimiser la figuration élargit en outre les possibilités d'exploitation ultérieures (enrichissement, modifications).

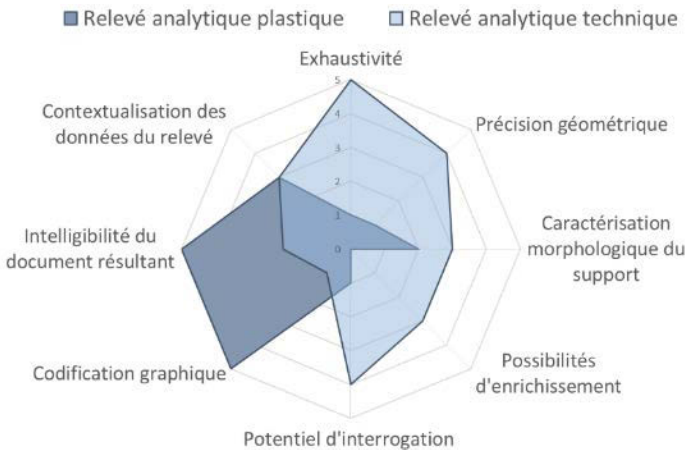


Figure 17 : Comparaison des deux familles de relevé selon différents critères liés aux trois moments de Saint-Aubin. Il apparaît clairement que les deux approches ne répondent pas aux mêmes besoins. Illustration : auteur.

Cet état de l'art nous montre clairement que le relevé demande systématiquement des arbitrages de la part du ou des releveur(s). Cela impose notamment certains efforts en termes de traçabilité pour assurer l'acceptabilité du relevé en tant que document scientifique. Airvaux et al., (1991) rappellent que la finalité du relevé est de rendre intelligible la complexité des œuvres pariétales, et qu'à ce titre, il doit impérativement répondre aux exigences suivantes :

- **Fiabilité** : L'image doit être une projection parfaite de la réalité sans altérations dimensionnelles ;
- **Exhaustivité** : Le relevé doit être le plus exhaustif possible, car l'absence d'un trop grand nombre d'informations peut être déterminante et conduire à une interprétation erronée ;
- **Offrir la possibilité d'un traitement analytique** de façon à isoler les représentations par affinités morphologiques ou directionnelles, et à en faciliter ainsi l'étude ;
- **Permettre une mise en relation entre la morphologie du support** et la disposition des figures sur celui-ci ;
- **Rapidité d'exécution** : Toutes les techniques existantes évoquées précédemment sont très chronophages. D'après les auteurs, le facteur temps est capital dans la mesure où il existe un grand nombre d'œuvres paléolithiques qui doivent être confrontées les unes aux autres pour progresser dans les interprétations. Nous pourrions appuyer ce propos en mentionnant que les nécessités de préservation des grottes ornées imposent aujourd'hui une gestion fine des temps d'accès aux sites, qui sont globalement assez courts.

Concernant l'ensemble de ces critères, qui rejoignent partiellement ceux que nous avons pu établir, les outils numériques ont de nombreux éléments de réponse à apporter. En effet, comme nous le verrons dans les chapitres suivants, le domaine de la numérisation 3D basée sur la réalité constitue pour le relevé une avancée considérable. La capacité à mémoriser l'objet d'étude dans ses aspects formels offre de nombreuses possibilités de traitement et d'extraction d'informations. Les outils de visualisation de données 3D offrent une liberté précieuse en termes de projection géométrique – point qui constituait, comme nous l'avons vu précédemment, l'une des problématiques majeures du relevé d'art pariétal indirect. À propos de la possibilité d'un traitement analytique, nous verrons que les travaux concernant les systèmes d'information constituent des solutions avantageuses, en particulier par leur capacité à mettre en relation des ressources hétérogènes et par le biais des structurations multicouches qui favorisent l'analyse de la corrélation des informations. Enfin, concernant la mise en relation entre la morphologie du support et la disposition des figures, nous verrons que les travaux du domaine de la réalité mixte offrent des perspectives intéressantes qui permettent d'aller bien plus loin dans la contextualisation, en autorisant l'hybridation d'environnements réels et virtuels.

Chapitre 2. Le relevé numérique

2.1. Vers un changement de paradigme

Les progrès réalisés depuis le milieu du XX^{e} siècle dans les domaines de la métrologie et de la vision par ordinateur ont favorisé l'essor des outils numériques de mesure de l'existant. Si le numérique n'a pas fondamentalement remis en question les **méthodes** de mesure, qui se présentent comme les héritages de deux modalités exposées dès la Renaissance basées sur la géométrie projective, l'amélioration de leur précision, du niveau d'automatisation, et de la capacité de mémorisation a en revanche constitué une évolution décisive pour de nombreuses disciplines. Ainsi, dans le domaine du patrimoine culturel, l'introduction des technologies numériques a rendu possible la collecte massive de données 2D et 3D, en réponse à des besoins variés d'étude, de suivi, de documentation, d'archivage, ou encore de valorisation.

En ce qui concerne le relevé – que l'on parle d'art pariétal ou de patrimoine bâti –, cette évolution présente un intérêt majeur : celui de pouvoir exprimer à travers les numérisations 3D la complexité géométrique des objets d'étude, d'en restituer les formes et l'aspect. Dans le cas des grottes ornées, c'est un point particulièrement intéressant, car la morphologie complexe des sites rend toute autre traduction graphique extrêmement fastidieuse (section 1.2.2.3). Pourtant, comme nous l'avons vu précédemment (section 1.2.2.3), la représentation des reliefs permet de rendre compte de la répartition des figures dans le site, de leurs relations avec les reliefs, des liens éventuels entre les différents stigmates ou encore de réaliser le suivi temporel d'un élément qui se dégrade. À ce niveau, la numérisation 3D appliquée au relevé d'art pariétal présente donc un intérêt évident, mais nous verrons par la suite que ses champs d'applications sont bien plus vastes et que ces données présentent de nombreux potentiels (section 2.3).

Cependant, le développement des technologies d'acquisition 3D de ces dernières décennies a clairement placé sa priorité sur la recherche d'une meilleure exhaustivité de la mesure, dans l'idée de permettre la production de copies numériques les plus proches possible des objets réels. D'après David Lo Buglio, ces exigences d'exhaustivité et de précision ont pour effet d'écarter la dimension cognitive de la pratique du relevé.

> *« Ainsi, la production de données géométriques massives, représentées par plusieurs millions de coordonnées, entraîne une surcharge d'informations ; la valeur (culturelle, cognitive, voire paradigmatique) de la représentation architecturale, en d'autres termes sa capacité à constituer un médium de connaissances, s'en trouve progressivement annihilée. »* (Lo Buglio, 2016)

Autrement dit, la sauvegarde massive des attributs géométriques et visuels des objets d'étude décale naturellement le moment où s'opère la sélection d'éléments signifiants. Cette étape intervient face à l'objet d'étude lors d'un relevé *classique*. Dans le cadre d'un relevé numérique, même si l'engagement analytique de l'observateur est encore un peu mobilisé lors du choix des méthodes d'acquisition et de leur mise en œuvre lorsqu'il en est l'opérateur, la majorité du travail de sélection et d'interprétation intervient *ex situ* auprès d'une copie 3D. Ainsi, là où le relevé *classique* peut se définir comme *« un moment où, grâce à l'observation de la forme, on restitue les caractères qui composent le bâti par une démarche qui se sert de l'acte de " dessiner " comme moyen pour " comprendre "*

l'architecture » (De Luca, 2009), le relevé numérique correspond davantage à une séquence d'activités distinctes successives d'acquisition, de mise en relation et de traitement de données virtuelles (Figure 18).

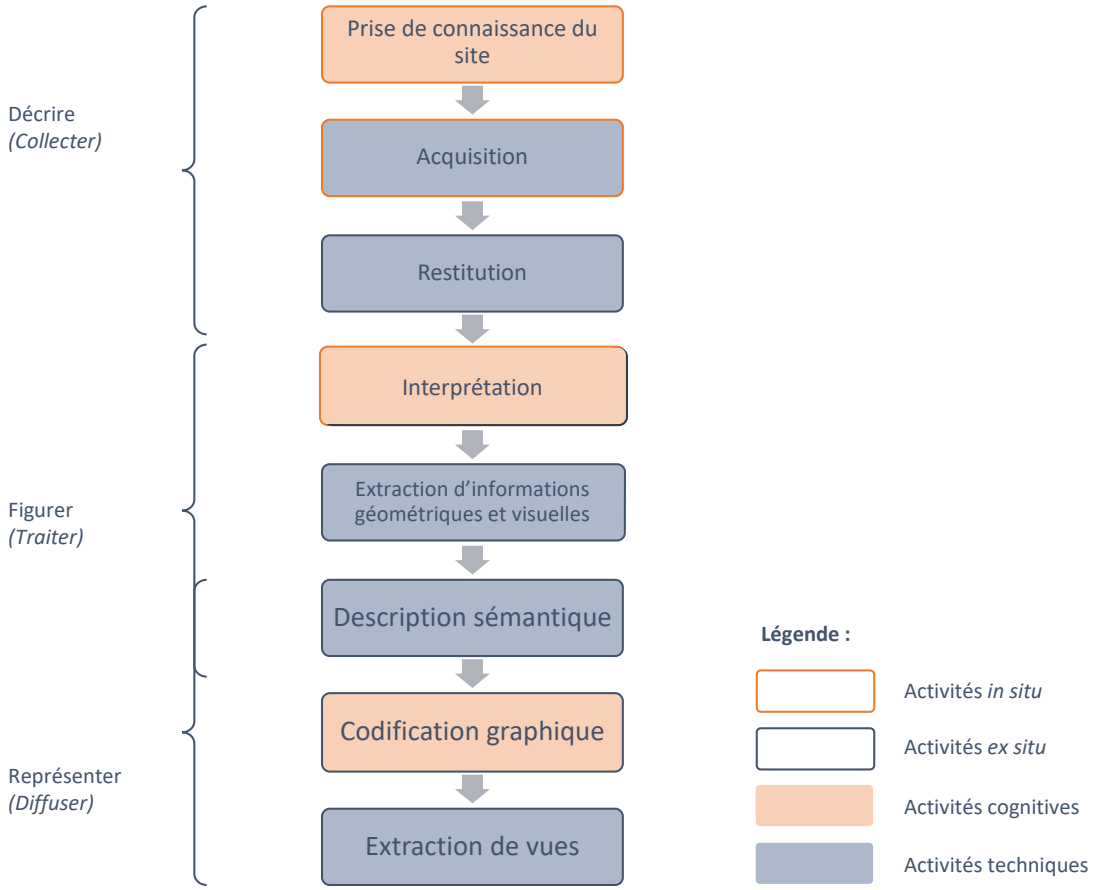


Figure 18 : Schéma des différentes étapes du relevé numérique. Contrairement au relevé *classique*, ce processus est séquentiel. Illustration : auteur.

La discontinuité que cela introduit sur le plan de la connaissance constitue un réel changement de paradigme, qui soulève de nouvelles problématiques. Dans ce chapitre, nous nous attacherons à expliciter les outils et méthodes impliqués dans les différentes étapes du relevé numérique. Nous tirerons un bilan des pratiques du relevé *classique* et numérique afin de comprendre quels en sont les enjeux contemporains, et ce que cela peut impliquer en termes d'outils et de méthodes.

Comme nous le verrons dans ce chapitre, si les données numériques possèdent de nombreux potentiels en termes de caractérisation géométrique et visuelle, leur exploitation requiert cependant une multiplication des procédés de traitement. Ce contexte cause un effet de dispersion des données rendant difficile la corrélation d'observations multidisciplinaires. À ce niveau, nous verrons que l'enrichissement sémantique, c'est-à-dire la possibilité d'intégrer les connaissances des différents acteurs aux données numériques spatialisées, constitue une étape fondamentale mais sous-exploitée faute d'outils adaptés. Enfin, nous verrons que l'éloignement du site d'étude et la décontextualisation des données impliquées par le processus de relevé numérique causent une rupture entre les environnements réels et virtuels qui nuit aux possibilités d'interprétation.

2.2. Vers la numérisation massive de l'existant

Les nouvelles technologies de l'information offrent aujourd'hui de nombreuses possibilités concernant l'acquisition de données concernant l'état d'un objet patrimonial, que ce soit concernant son aspect, sa géométrie, ou encore certains attributs physiques. Depuis plusieurs années, cette étape d'acquisition de données numériques est devenue systématique lors de travaux de suivi, d'étude, ou de conservation d'objets patrimoniaux. En termes de relevé, elle s'inscrit dans le moment de ***description***. Selon les capteurs et méthodes employés, les données acquises peuvent être représentées en 2D ou en 3D. Nous étudierons dans cette section les principales techniques de numérisation impliquées dans le relevé numérique. L'ambition n'est pas de dresser un état de l'art complet des techniques d'acquisition numérique, mais de comprendre la nature et les spécificités des outils mobilisés afin d'en soulever les apports et les limites.

2.2.1. Outils et méthodes d'acquisition 2D

2.2.1.1. La photographie numérique

Les photographies numériques constituent des données incontournables dans le domaine de l'art pariétal, car elles fournissent de nombreuses informations et à plusieurs échelles. Elles sont notamment exploitées comme archives ou témoins lors de constats d'état, comme supports lors de relevés, ou encore comme données d'entrée pour des reconstructions par photogrammétrie.

À la base des appareils photo et caméras numériques, les capteurs photographiques numériques sont des composants électroniques photosensibles qui permettent de convertir des rayonnements électromagnétiques reçus en signaux électriques analogiques. L'intensité de la lumière est captée par des cellules photosensibles nommées photosites. Pour obtenir des informations colorimétriques, celles-ci sont accompagnées de filtres, le plus courant étant le filtre de Bayer, composé à 50% de filtres verts, 25% de rouges et 25% de bleus. Selon les filtres employés, les appareils peuvent permettre de cibler des portions spécifiques du spectre électromagnétique au-delà de la lumière visible, en particulier l'infrarouge et l'ultraviolet (section 2.2.1.3).

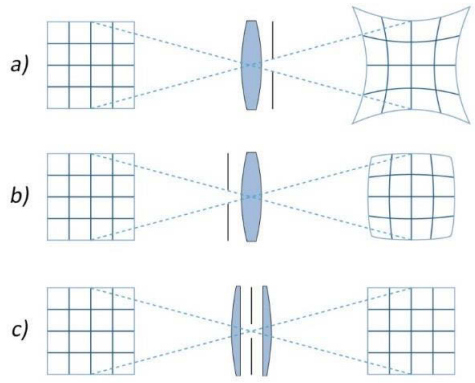


Figure 19 : Les distorsions a) en coussinet, b) en barillet, c) sans distorsion. Illustration : (Panther, 2005).

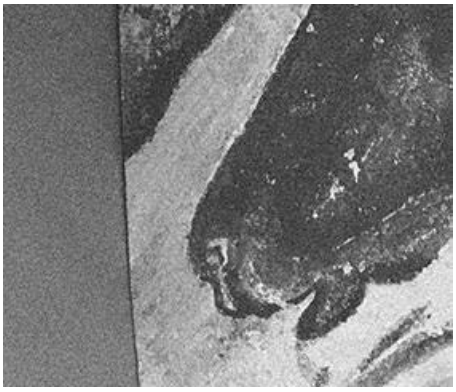


Figure 20 : Bruit numérique sur une photographie (facsimilé du panneau des Chevaux). Illustration : auteur

On distingue trois grandes familles de capteurs (CCD, CMOS[13] et Foveon), qui se différencient principalement par les dispositifs de transfert de charge exploités. Il en existe plusieurs tailles, parmi lesquelles se trouvent le format APS-C (25,1 x 16,7 mm), le plein-format (24 x 36 mm), et le moyen format (49,1 x 36,8 mm). Le choix du capteur et de l'objectif qui l'accompagne a une incidence importante sur la qualité des données acquises mais aussi sur leur potentielle exploitation dans le cas où l'image doit être utilisée comme source pour produire de nouveaux supports d'analyse. Par ailleurs, il est important de tenir compte du fait que les photographies sont soumises à des phénomènes optiques susceptibles de nuire à la précision des mesures obtenues lors de l'utilisation de l'image.

- **Les aberrations géométriques :** Elles surviennent lorsqu'une ou plusieurs conditions de Gauss ne sont plus respectées, notamment du fait de défauts optiques de l'objectif. Elles se manifestent par des phénomènes de distorsion, dont les plus répandus sont les distorsions en barillet et en coussinet (Figure 19). La distorsion d'un objectif est une grandeur mesurable qui s'exprime en pourcentage.

- **Le bruit numérique** : Il entraîne une perte de netteté dans les photographies, particulièrement pour les zones sombres ou peu éclairées (Figure 20). Le bruit numérique peut avoir plusieurs origines comme la température du capteur, des erreurs de transmission de données, la présence de particules sur le capteur, ou encore des origines électroniques dans le cas du bruit de grenaille.

- **Les aberrations chromatiques :** Elles sont liées à la variation des indices de réfraction des lentilles en fonction de la longueur d'onde des signaux qui les traversent. Elles ont pour effet la multiplication des points focaux, plus ou moins éloignés sur l'axe optique, ce qui produit un effet sur les photographies qui se manifeste par des contours irisés, voire flous.

2.2.1.1. Macrophotographie

La macrophotographie correspond à l'ensemble des techniques photographiques permettant de photographier de petits éléments avec des rapports de grandissement[14] supérieurs à 1:1. Cela signifie que, sur ces images, les éléments photographiés ont une dimension supérieure ou égale à leur taille réelle. Les macrophotographies sont particulièrement utiles pour capturer de petits détails. Elles sont notamment utilisées en tracéologie pour analyser les techniques employées sur différents sites présentant des gravures (Zotkina, 2016). Elles peuvent également être exploitées comme images sources en vue d'une reconstruction 3D par photogrammétrie (Plisson, Zotkina, 2015).

2.2.1.2. Photographie en lumière rasante (LR), semi-rasante (SR) ou directe (LD)

Les photographies en lumière rasante ou semi-rasante sont obtenues en procédant à un éclairage de l'objet d'étude à l'aide d'une lampe dont la direction du faisceau lumineux est plus ou moins tangente au plan moyen du sujet. Ces photographies sont particulièrement utiles pour accentuer la perception des reliefs et faciliter la visualisation de gravures. Les photographies en lumière directe constituent quant à elles des supports de lecture sur lesquels les figures des parois sont éclairées de manière homogène (Figure 21). Elles sont habituellement utilisées pour générer les orthophotographies qui

[13] Respectivement *« Charged Coupled Service »* et *« Complementary Metal Oxyde Semi-Conductor ».*
[14] Le rapport de grandissement est le rapport entre la dimension d'un élément sur la photographie et sa dimension réelle.

servent de support de relevé dans le cadre de relevés analytiques plastiques ou techniques, mais aussi comme archives témoignant d'un état des lieux de la cavité à une date donnée.

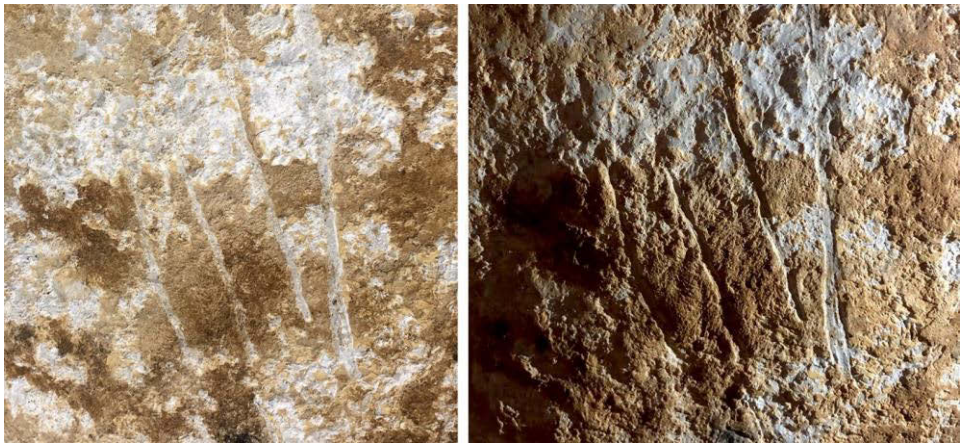


Figure 21 : Détail du facsimilé du panneau des Chevaux (griffure). À gauche : lumière directe. À droite : lumière rasante. Illustration : auteur.

2.2.1.3. Imagerie multispectrale

L'imagerie multispectrale est un procédé utilisé pour observer un objet d'étude à l'aide de plages de longueurs d'onde sélectionnées dans le spectre électromagnétique qui incluent et dépassent les capacités de l'œil humain (Dyer et al., 2013) (Figure 22). En règle générale, outre la lumière visible (VIS), les rayonnements exploités appartiennent à la catégorie des ultraviolets (UV 200-400nm) ou à celle des infrarouges (IR 760-1700nm).

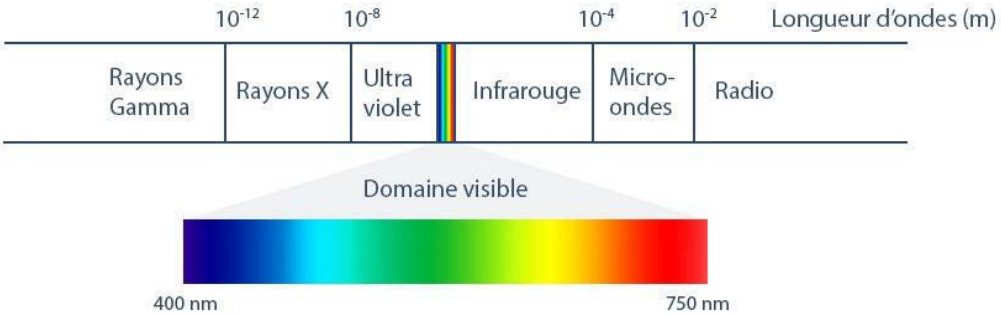


Figure 22 : Principaux domaines du spectre électromagnétique. Le visible n'en représente qu'une faible portion. Illustration : auteur

La façon dont un rayonnement émis pénètre un objet d'étude dépend de sa longueur d'onde et des propriétés optiques des matériaux qui composent l'objet (absorption, émission et réflexion). Ainsi, connaissant le comportement de certains matériaux il est possible de sélectionner les plages de rayonnement les plus appropriées pour un but donné. Par exemple, dans le cas d'un tableau, les longueurs d'onde plus courtes (comme les UV) sont souvent facilement absorbées par les couches extérieures (habituellement les vernis), tandis que les longueurs d'onde plus longues peuvent les traverser et interagir avec la peinture ou d'autres couches inférieures. Ainsi, en sélectionnant des combinaisons particulières d'éclairages et de filtres il est possible de produire des images donnant des informations sur la répartition des matériaux dans l'objet étudié. Dans le cas du relevé d'art pariétal, l'emploi de techniques d'imagerie multispectrale peut notamment permettre de faire apparaître des figures ou des stigmates invisibles à l'œil nu.

Une acquisition multispectrale implique trois composants principaux : 1) Un rayonnement entrant, qui est généré par une source de rayonnement et se déplace vers l'objet, 2) L'objet, qui interagit avec le rayonnement entrant, 3) Un rayonnement sortant qui, suite à l'interaction entre le rayonnement entrant et l'objet, se déplace de l'objet vers l'appareil d'enregistrement. Selon les propriétés du rayonnement entrant et du rayonnement sortant, on distingue plusieurs catégories d'images, qui révèlent différents aspects de l'objet étudié.

2.2.1.3.1. Photographie ultraviolette (UV)

Lorsque l'objet d'étude est soumis à des rayonnements ultraviolets, il est possible de produire des photographies pour enregistrer la réflexion dans les UV de ce rayonnement entrant (UVr), ou bien la fluorescence dans le visible des matériaux soumis à ce rayonnement (UVf). Les deux méthodes sont couramment utilisées comme moyens d'investigation pour la peinture de chevalet, mais des applications à l'étude d'art pariétal ont été mises en œuvre dès les années 1970 (Marshack, 1975). Les photographies UV peuvent être exploitées pour aider à caractériser certains matériaux (pigments, colorants, vernis, etc.) lorsqu'elles sont associées à d'autres techniques comme la spectroscopie de fluorescence et des mesures colorimétriques. Elles permettent également d'évaluer l'état de conservation d'une œuvre ou encore identifier des retouches. Néanmoins, il faut tenir compte du fait que l'émission spectrale des matériaux peut être modifiée par la présence d'autres matériaux environnants, donnant lieu à des phénomènes d'absorption et de diffusion, ce qui peut rendre difficile l'interprétation des images (Cosentino, 2015).

Dans le cas de l'art pariétal, certaines études exploitent le fait que la calcite et les organismes vivants aient une forte réponse fluorescente contrairement aux ocres et au manganèse utilisés comme pigments (Aujoulat, 1993c ; Bahn, Vertut, 1997). Connaissant ce comportement de la matière, l'utilisation de photographies UV permet d'évaluer les dommages causés aux figures par des croissances ou des écoulements de calcite, mais aussi de révéler la peinture sous-jacente pour visualiser les détails de la décoloration. Cette technique a notamment été employée avec succès dans la grotte de Niaux pour révéler le bison du Salon Noir, initialement peu visible car endommagé par l'eau ruisselant le long de la paroi (Fritz, Tosello, 2007) (Figure 23). Dans la grotte ornée Mayenne-Sciences, les photographies UV de Agueda et Denis Vialou ont été utilisées pour mieux visualiser certaines figures afin d'aider à déterminer la nature des animaux représentés (Pigeaud et al., 2004).

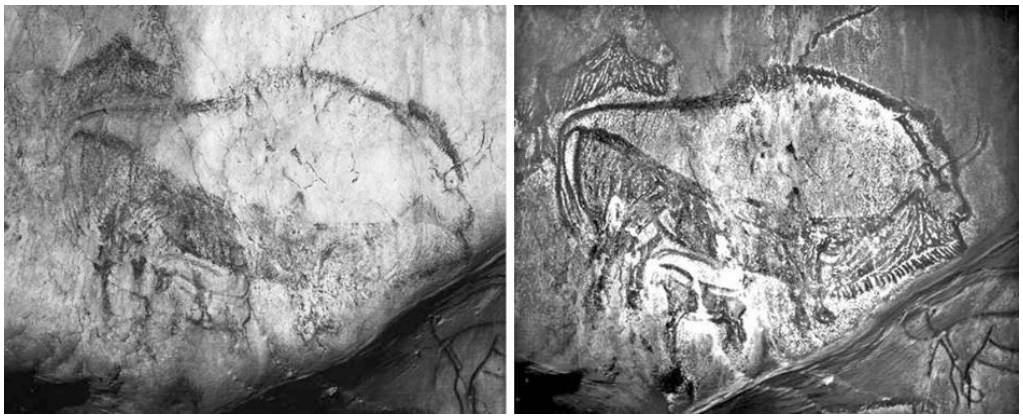


Figure 23 : Bison peint du Salon Noir (Grotte de Niaux, Ariège, France). À gauche : photographie VIS. À droite : photographie UV. Endommagée en 1978, la figure n'est pas facilement visible aujourd'hui, mais en lumière UV de nombreux détails sont à nouveau perceptibles car les traces de manganèse, conservées sur la paroi de la grotte absorbent les rayons UV tandis que la couche de calcite les reflète. Illustration : (Fritz, Tosello, 2007)

2.2.1.3.2. Photographie infrarouge (IR)

La photographie en fluorescence infrarouge (IRf) permet d'enregistrer l'émission infrarouge engendrée par l'exposition d'un objet à un rayonnement de lumière visible. La réflectographie infrarouge (IRr) permet quant à elle d'observer la réflexion infrarouge d'un objet soumis à un rayonnement infrarouge. Les photographies IR sont couramment utilisées dans le cadre d'opérations de conservation ou de restauration d'œuvres d'art, particulièrement dans le cas de peintures pour visualiser les différentes couches de matière (peinture, dessins préparatoires, support, etc.). Elles permettent de mettre en évidence certaines modifications apportées aux œuvres, car des matériaux d'aspects proches en lumière visible peuvent montrer des comportements différents en infrarouge.

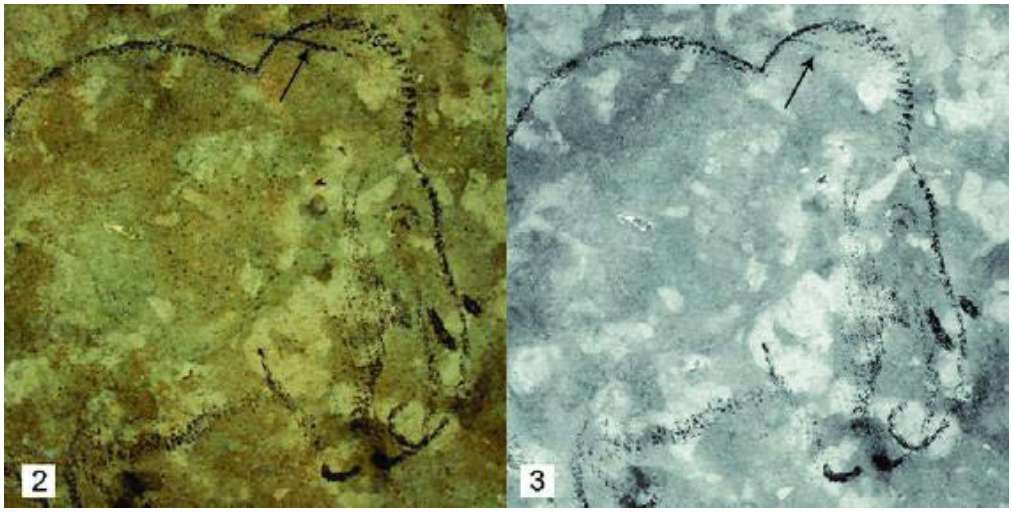


Figure 24 : Mammouth dans la grotte de Rouffignac. À gauche : Photographie VIS. À droite : Photographie IRr. Le trait traversant au niveau de la tête, préalablement identifié comme du noir de carbone par spectrométrie Raman, apparaît plus clair que le reste de la figure, dont les traits sont composés d'oxyde de manganèse. Illustration : (Beck et al., 2014).

Dans le cas de l'art pariétal, ce principe a été exploité pour l'étude de la grotte de Rouffignac. Les photographies IRr ont ainsi permis de différencier sur plusieurs figures de mammouths les tracés réalisés au charbon de bois et ceux réalisés avec de l'oxyde de manganèse (Beck et al., 2014) (Figure 24).

2.2.1.4. Imagerie par transformation de réflectance (RTI)

L'imagerie par transformation de réflectance, plus connue sous le nom de *Reflectance Transformation Imaging* (RTI) est une technique permettant de mettre en évidence les états de surface d'un objet par des jeux d'ombres à partir d'une image dynamique sur laquelle l'utilisateur peut faire varier de manière interactive la direction de l'éclairage (Figure 25).

Les images RTI sont créées à partir de plusieurs photographies numériques prises depuis la même position, chacune avec une source de lumière différente dont la direction du faisceau est connue. Cette série de photographies produit donc des images du même objet avec des ombres et lumières différentes. Ces informations d'éclairage sont ensuite exploitées pour générer un modèle mathématique de la surface de l'objet, permettant de simuler le comportement de l'objet sous des éclairages variés (Malzbender et al., 2001). Pour visualiser ces fichiers, des logiciels de visualisation RTI doivent être utilisés. Ils sont dotés d'une interface utilisateur permettant aux utilisateurs de définir à leur gré la

direction de la lumière, mais aussi de profiter de certaines textures efficaces pour mettre en valeur le relief (Palma et al., 2010).

Figure 25 : Capture d'écran d'une image RTI (facsimilé du panneau des chevaux) visualisée dans le logiciel RTIviewer.

Cette technique est particulièrement intéressante pour visualiser des détails morphologiques très fins comme des gravures, c'est pourquoi elle est couramment employée pour l'étude d'art pariétal (Duffy, 2010 ; Uribe, Wheatley, 2013 ; Díaz-Guardamino et al., 2015 ; Kotoula et al., 2018). Par ailleurs, l'imagerie RTI n'exige pas de matériel spécifique, en dehors de l'appareil photo et des sources de lumière, et n'impose pas de contrainte particulière quant aux bandes spectrales utilisées. Elle peut donc très bien s'appliquer aux différentes techniques de photographie, notamment IR et UV.

2.2.2. Outils et méthodes d'acquisition 3D

Les dispositifs d'acquisition de données 3D sont généralement classés en deux catégories : avec et sans contact. Elles sont majoritairement basées sur l'exploitation du spectre électromagnétique et classées en deux sous-catégories distinctes : les techniques actives et passives (Figure 26). La distinction entre les techniques d'acquisition actives et passives est fonction du rayonnement utilisé pour la prise de mesure (Remondino, 2014).

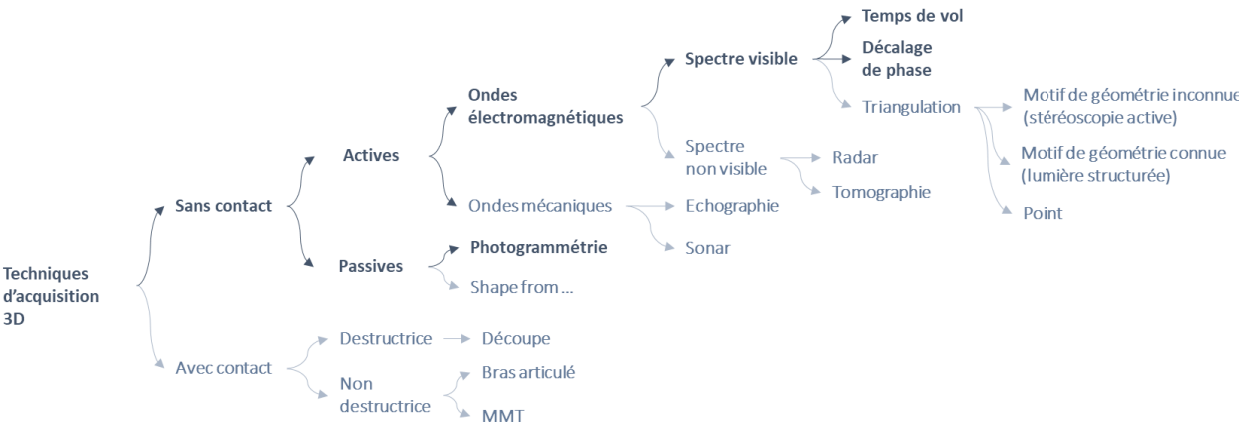


Figure 26 : Taxonomie non exhaustive des techniques d'acquisition 3D. Illustration : auteur, d'après (Jacquot, 2014 ; Loriot, 2009 ; Remondino, 2011).

Les technologies actives se basent majoritairement sur l'utilisation d'ondes électromagnétiques du spectre visible (lasergrammétrie) ou non visible (tomographie, radar, etc.), mais aussi sur l'utilisation d'ondes mécaniques (sonar, échographie, etc.). La numérisation tridimensionnelle active appliquée au relevé d'art pariétal relève principalement de la lasergrammétrie. Contrairement aux techniques actives qui émettent un signal spécifique, les techniques passives se basent sur l'étude du rayonnement ambiant naturel (Jacquot, 2014). Parmi les techniques passives, la photogrammétrie est de loin la plus répandue, nous ne détaillerons donc pas les autres dans ce document. Par ailleurs, dans le cas de l'art pariétal, il est évident que les techniques d'acquisition destructives sont absolument inenvisageables. De même, les techniques d'acquisition avec contact sont peu adaptées du fait de la fragilité et de la valeur historique des monuments que nous étudions, nous ne nous attarderons donc pas sur ces systèmes.

2.2.2.1. Lasergrammétrie

La lasergrammétrie est une technique de numérisation 3D sans contact permettant d'obtenir par balayage un ensemble de coordonnées X Y Z des points de la surface de l'objet étudié. Depuis les premières expérimentations dans le champ du relevé patrimonial, notamment (Robson Brown et al., 2001 ; Barber et al., 2001 ; Boehler et al., 2001), ces dispositifs se sont considérablement améliorés et constituent désormais des outils incontournables pour les relevés d'objets patrimoniaux complexes. Ces dernières années ont ainsi vu se multiplier les campagnes de numérisation par scanner laser, contribuant très largement à la production massive de données 3D. Les scanners laser terrestres sont principalement constitués de quatre éléments : une tête laser, un système de balayage permettant de faire varier l'angle du rayon laser, un système récepteur enregistrant le signal retour après réflexion sur l'objet, et un logiciel de pilotage.

Ils se basent sur l'émission d'un signal laser et sa réception par le système récepteur pour déduire la distance séparant l'émetteur du point d'impact sur l'objet étudié. Le système de balayage fait ensuite varier l'angle horizontal du rayon laser pour réaliser la mesure suivante. Lorsque l'appareil a réalisé un tour complet, une variation de l'angle vertical du rayon laser permet de commencer une nouvelle série de mesures à une hauteur différente (Figure 27).

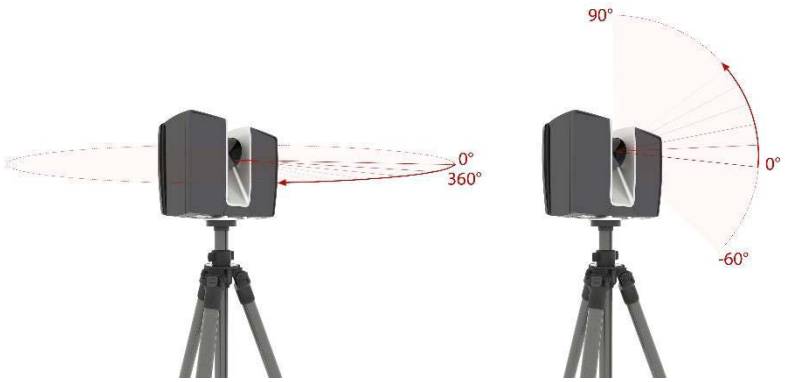


Figure 27 : Principe du système de balayage du scanner laser. Illustration : auteur.

Pour réaliser la mesure des distances, deux technologies existent : les scanners dits *temps de vol* et les scanners à *décalage de phase*. Dans le premier cas, le principe repose sur la mesure du temps de parcours entre l'émission d'une impulsion laser et la réception de l'onde réfléchie par l'objet à acquérir. Dans le second cas, le principe repose sur la comparaison des phases d'un signal émis avec celles du signal reçu

en retour. Dans les deux cas, les coordonnées sont obtenues en coordonnées sphériques. Elles sont généralement converties postérieurement en coordonnées cartésiennes, et peuvent être accompagnées de valeurs d'intensité[15], et éventuellement d'attributs colorimétriques RGB si le scanner est équipé d'un appareil photographique calibré. Ces coordonnées sont relatives, et mais sont souvent converties en coordonnées absolues par géoréférencement lors du post-traitement. Les données résultant d'une acquisition par scanner laser se présentent sous la forme de nuages de points denses (Figure 28).

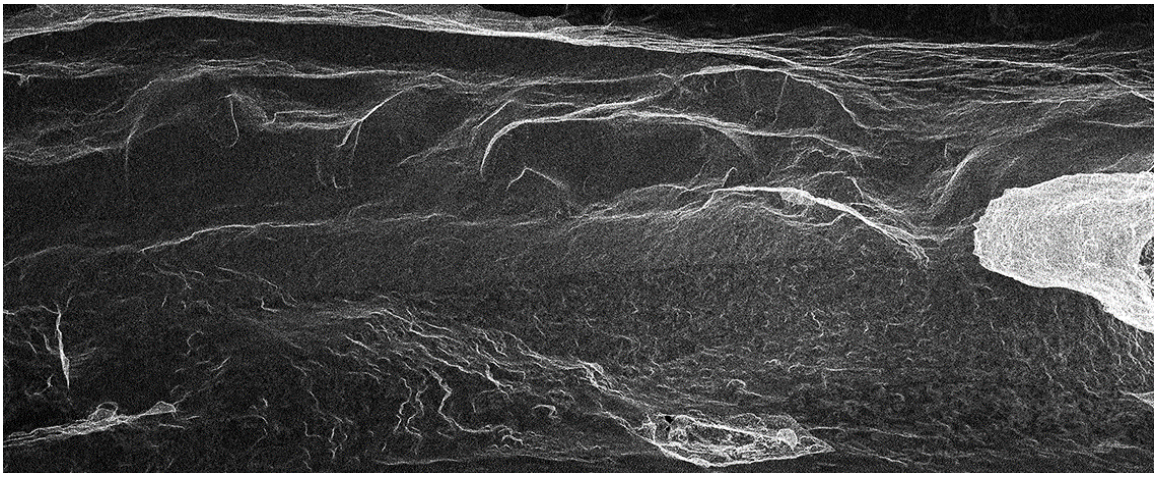


Figure 28 : Capture d'écran du nuage de points de la frise sculptée de l'abri du Cap-Blanc. Données : © AGP/CNP.

La lasergrammétrie est particulièrement intéressante pour sa capacité à acquérir rapidement de très grandes quantités de données avec une résolution spatiale de l'ordre de quelques millimètres. En contrepartie, elle exige un travail important de post-traitement, car la numérisation d'un objet patrimonial impose très souvent de réaliser plusieurs stations, qui doivent donc être ensuite assemblées en un seul fichier représentant l'objet dans sa globalité : on parle de consolidation. Le fait que cette technique délivre pour chaque station des données de résolution homogène constitue cependant une limite, car les objets patrimoniaux disposent souvent d'une grande variété de niveaux de détails. Cette méthode a donc naturellement tendance à figer la copie virtuelle de l'objet étudié alors même que son étude implique, une pluralité d'acteurs aux besoins variés, de nombreux allers-retours, et différents niveaux de lecture liés à sa complexité.

2.2.2.2. Photogrammétrie

Les premiers travaux de photogrammétrie, au milieu du XIX^{e} siècle, visaient déjà le relevé architectural d'édifices remarquables (Héno, Chandelier, 2014). Dans la deuxième moitié du XX^{e} siècle, de nombreux travaux de relevé ont été initiés par l'IGN, comprenant notamment des cathédrales françaises emblématiques (Strasbourg, Paris, Rouen), ainsi que la grotte de Lascaux (fin des années 1960). Les chambres métriques d'autrefois ont progressivement été remplacées par les appareils de photographie modernes, et les processus de calculs ont été automatisés, rendant de fait, la photogrammétrie aujourd'hui accessible à tous pour un investissement globalement faible. Aujourd'hui, la photogrammétrie peut proposer des résultats dont la qualité est proche de ceux obtenus par

[15] L'intensité correspond à l'intensité du signal reçu après réflexion du rayon sur la surface de l'objet d'étude, c'est-à-dire à la quantité de lumière reçue. Elle dépend de plusieurs paramètres : l'angle d'incidence du rayon émis, la distance entre l'objet et le scanner, la réflectance de l'objet (et donc de la longueur d'onde du laser) et de la puissance émise (Landes, Grussenmeyer, 2011).

lasergrammétrie (Remondino et al., 2012). À la différence de cette dernière, la photogrammétrie offre cependant une grande flexibilité quant aux niveaux de détails, ce qui permet de produire des numérisations 3D multirésolutions. Les avancées en matière de photogrammétrie incrémentale offrent également une certaine évolutivité particulièrement intéressante dans le cas du suivi d'objet patrimonial.

2.2.2.2.1. Génération de coordonnées 3D à partir de photographies

La photogrammétrie repose sur les principes de la perspective énoncés par Brunelleschi dès le XV^{e} siècle. On peut en effet considérer la prise de vue photographique comme un cas classique de projection géométrique d'un espace en trois dimensions sur un plan en deux dimensions. Pour se placer dans ce cas de figure, on réalise une approximation en assimilant l'objectif de l'appareil photo, qui comporte en réalité un groupe de lentilles, comme une unique lentille mince convergente. On assimile également le centre du diaphragme comme le centre optique de la lentille. Ainsi, les lois optiques affirmant que « *tout rayon incident passant par le centre optique d'une lentille n'est pas dévié par la lentille* », permettent de schématiser la prise de vue photographique de la façon suivante (Figure 29) :

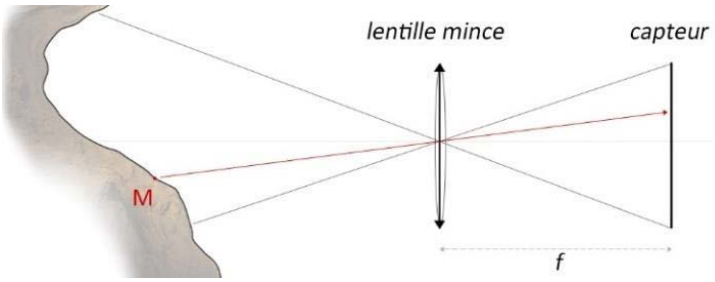


Figure 29 : Approximation du fonctionnement d'un appareil photo. Illustration : auteur, d'après (Héno, Chandelier, 2014).

Pour générer des coordonnées 3D à partir de photographies, il s'agit alors de considérer que la position dans l'espace d'un point d'une photographie, est située sur la droite passant par le centre optique de l'appareil photographique et ce point (Figure 30 A). À partir des poses de plusieurs images du même objet, il est donc théoriquement possible de calculer la position 3D des différents points (Figure 30 B).

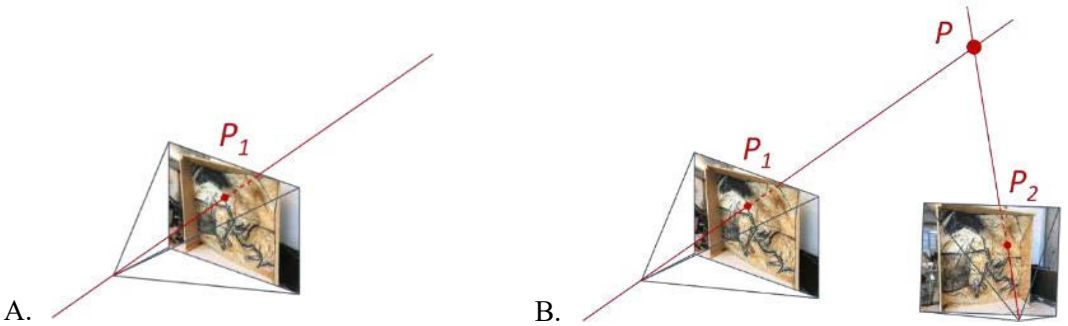


Figure 30 : Position 3D d'un point apparaissant sur une (A) et deux (B) photographies. Illustration : auteur, d'après (De Luca, 2009).

En réalité, comme nous l'avons vu précédemment (section 2.2.1.1) les systèmes optiques ne sont pas parfaits et peuvent notamment causer des distorsions. Le processus repose donc en premier lieu sur la capacité à valider les conditions de la projection perspective en tenant compte des paramètres internes de l'appareil photographique et des éventuelles aberrations géométriques : on parle de calibration. Le reste du processus repose sur la mise en relation de différentes photographies puis sur l'orientation externe des caméras, c'est-à-dire la détermination de leur position et orientation dans l'espace-objet (De Luca, 2009). Si autrefois le processus comprenait des étapes manuelles (notamment la mise

en correspondance d'images), les progrès réalisés ces dernières décennies dans le domaine de la vision par ordinateur permettent aujourd'hui l'automatisation de l'ensemble de la chaîne de traitement. Le processus général est détaillé par Pierrot-Deseilligny et al. (2011).

Les données générées sont stockées sous la forme de cartes de profondeurs qui sont généralement exploitées par la suite pour produire des nuages de points denses. Notons toutefois que la photogrammétrie calcule des positions relatives dans un repère arbitraire. Pour exprimer les coordonnées dans un repère global, il est nécessaire de passer par le géoréférencement. Il s'agit d'une étape manuelle, nécessaire dans le cas de prises de mesures à l'échelle.

Au cours du processus, plusieurs erreurs peuvent provenir des différentes étapes et affecter la mesure. Aujourd'hui, l'écueil le plus important provient de l'appariement des points homologues, qui peut notamment livrer des incohérences dues à la scène photographiée ou à la qualité des prises de vues. La numérisation par photogrammétrie peut s'avérer délicate dans le cas de matériaux réfléchissants ou transparents, d'éléments à la colorimétrie uniforme, ou encore dans le cas de détails trop répétitifs. Les autres sources d'erreurs ne sont pas tant liées à la nature de l'objet photographié qu'à des choix de l'opérateur, qui se manifestent particulièrement à travers la qualité des prises de vue qui doivent être réalisées en tenant compte de contraintes spécifiques (recouvrement des images, bruit, luminosité, etc.).

2.2.2.2.2. Photogrammétrie incrémentale

La photogrammétrie incrémentale permet de spatialiser une ou plusieurs nouvelles images autour d'un jeu de données principal. Pour cela, plusieurs méthodes sont possibles : par résection spatiale (Shih, Faig, 1988) ou par ajustement de faisceaux (Jacobsen, 1998) (section 5.3.3). Les progrès concernant l'automatisation des chaînes de traitement (Pamart, Morlet, et al., 2019) constituent une avancée importante dans la mesure où ils permettent non seulement d'optimiser le processus de spatialisation de nouvelles images sur un jeu de données principal, mais aussi parce qu'ils permettent d'envisager le nuage de points 3D comme une entité dynamique évolutive susceptible d'être enrichie au fil du temps. Ce principe de spatialisation d'images autour d'un modèle de référence constitue par ailleurs un moyen simple et efficace pour mettre en relation des données 2D hétérogènes autour d'une représentation numérique 3D (Snavely et al., 2006 ; Manuel, 2016).

2.2.2.3. Fusion de données 3D multimodales

Les sections précédentes montrent bien que les différentes méthodes d'acquisition de données 2D et 3D présentent des atouts et limites variés répondant à des besoins opérationnels différents. Dans le cas du relevé patrimonial, la complexité des objets d'étude requiert souvent l'utilisation conjointe de différentes méthodes. Par exemple, dans le cas des grottes ornées, il est souvent nécessaire de réaliser une numérisation globale du site par lasergrammétrie, et d'exploiter la photogrammétrie pour obtenir une résolution plus fine sur certains panneaux spécifiques. Un même panneau peut également nécessiter des études à différentes échelles nécessitant des numérisations à différentes résolutions.

Pour cette raison, des travaux de recherche se sont intéressés durant ces dernières décennies à différents aspects de la problématique de la fusion de données 3D multimodale, notamment l'hybridation lasergrammétrie/photogrammétrie (El-Hakim et al., 2004 ; Rönnholm et al., 2007 ; Remondino, Rizzi, 2010), la fusion multirésolution (Bonora et al., 2005 ; Gruen et al., 2005 ; Guidi et al., 2009) ou encore

l'acquisition multitemporelle (Vallet et al., 2012 ; Peteler et al., 2015 ; Dimitriou et al., 2016). Si ces pratiques sont devenues courantes dans les domaines du patrimoine et de l'art pariétal, elles exigent cependant des chaînes de traitement complexes, parfois empiriques, posant des problèmes d'interopérabilité et de reproductibilité.

2.2.3. Bilan des méthodes d'acquisition de données numériques

Nous avons balayé dans cette section les méthodes de numérisation 2D et 3D les plus répandues dans domaine du relevé numérique. L'état de l'art révèle qu'elles présentent des intérêts et limites différents en termes de résolution, de vitesse, de contraintes opérationnelles, de post-traitement, etc. Dans le cas de l'art pariétal, les différents niveaux informationnels des parois rendent la photogrammétrie plus appropriée pour l'étude des panneaux, principalement du fait de sa capacité à proposer des résolutions hétérogènes adaptées aux besoins (qui dépendent notamment de la morphologie des parois, des zones d'intérêt, des dimensions du plus petit élément à acquérir, des disciplines impliquées, etc.). Par ailleurs, la photogrammétrie présente un avantage notable pour les pariétalistes, car elle permet de s'affranchir du besoin d'intermédiaires pour l'acquisition de données brutes, ce qui permet aux acteurs du relevé d'intervenir à chaque étape du processus (Figure 18).

La lasergrammétrie est privilégiée pour la numérisation des sites dans leur ensemble et rendre compte de leur volumétrie, car la complexité morphologique des sites rendrait cette opération trop fastidieuse en photogrammétrie (quantité de données et temps de traitement). Dans les deux cas, la quantité de données brutes et les chaînes de traitement nécessaires pour l'obtention des numérisations 3D demandent de plus en plus de puissance matérielle, ce qui nous conduit à penser que le déport des calculs vers des solutions *cloud* constitue une avancée intéressante.

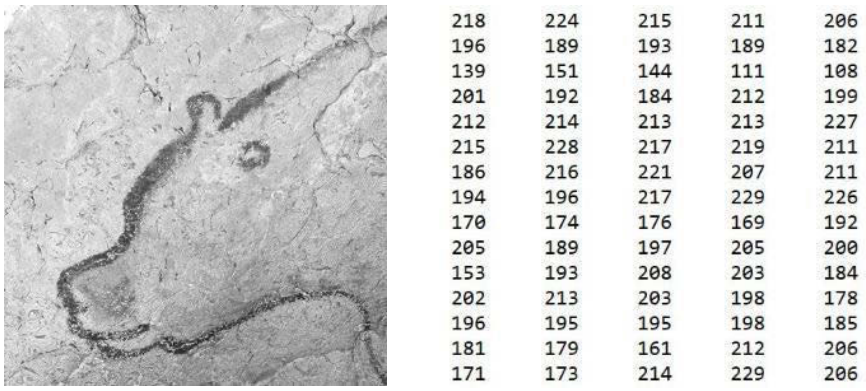
Enfin, cet état de l'art nous montre que toutes ces méthodes d'acquisition constituent des projections d'une réalité 3D continue dans des représentations informatiques discrètes 2D. C'est évident dans le cas des photographies, mais cela concerne aussi les méthodes d'acquisition 3D : dans le cas de la photogrammétrie, les coordonnées 3D proviennent de cartes de profondeur, elles-mêmes 2D et discrétisées en données pixellaires. Dans le cas de la lasergrammétrie, les mesures des stations sont stockées dans des images panoramiques équirectangulaires, donc également sous la forme de données pixellaires. D'ailleurs, cette discontinuité demeure visible après le traitement des données brutes, étant donné que les fichiers résultant de ces processus de numérisation sont des nuages de points.

On parle alors de ***fossé sensoriel***, c'est-à-dire d'un fossé séparant les objets du monde réel et l'information contenue dans une description informatique dérivée de l'enregistrement de la scène (Smeulders et al.,2000). Il semble alors clair que l'engagement cognitif sollicité pour l'interprétation de l'objet d'étude par un observateur face à celui-ci n'est pas le même que face à une représentation numérique de ce même objet. Autrement dit, lors d'un relevé, la copie numérique ne saurait prétendre remplacer l'original.

2.3. Extraction d'informations géométriques et visuelles

Le terme de ***fossé sémantique*** illustre bien les enjeux soulevés par l'usage exclusif de données numériques lors d'un relevé. Cette notion provient du domaine de la vision par ordinateur, et se définit

par *« le manque de concordance entre les informations que l'on peut extraire de données visuelles et l'interprétation qu'ont ces mêmes données pour un utilisateur dans une situation déterminée »* (Smeulders et al., 2000 ; Boucher, Le, 2005). Plus simplement, il s'agit de considérer le fait que les données numériques acquises par les techniques énoncées précédemment sont en réalité un ensemble de données discrètes concernant l'objet étudié, et que la façon de les représenter revêt plus ou moins de sens pour son destinataire – par exemple un observateur dans le cas du relevé, ou un ordinateur dans le cas d'approches automatiques (Figure 31).



218	224	215	211	206
196	189	193	189	182
139	151	144	111	108
201	192	184	212	199
212	214	213	213	227
215	228	217	219	211
186	216	221	207	211
194	196	217	229	226
170	174	176	169	192
205	189	197	205	200
153	193	208	203	184
202	213	203	198	178
196	195	195	198	185
181	179	161	212	206
171	173	214	229	206

Figure 31 : Deux représentations différentes d'une même image. À partir de l'image gauche, l'être humain est capable de reconnaître immédiatement le contenu d'une image numérique comme une représentation logique et continue, alors que l'image droite, représentant les mêmes données d'intensité, est bien moins intelligible. Photographie © CNP/MC.

Quelles que soient les technologies de numérisation employées, les données résultantes nécessitent donc la plupart du temps des opérations de traitement ou simplement certaines manipulations à l'aide d'outils spécifiques pour permettre aux différents acteurs de les exploiter et d'extraire les informations qu'elles contiennent, avant d'y associer un sens.

Dans le cas des images numériques, il peut s'agir d'opérations de retouche, par exemple pour réaliser des transformations géométriques (comme redresser les verticales), des modifications des couleurs et des valeurs (comme l'ajustement des niveaux), des recadrages ou sélections partielles, ou encore l'utilisation de méthodes de traitement d'images plus spécifiques comme l'application de filtres ou d'effets variés. Nous verrons également que certaines informations peuvent être extraites de manière automatique pour faciliter certaines chaînes de traitement ou l'interrogation ultérieure des données. Dans le cas des données 3D, la volonté de les exploiter dans des logiciels tiers peut impliquer notamment des étapes de maillage, de texturage, ou encore d'optimisation. L'interrogation des données passe par la manipulation des nuages de points ou modèles 3D, à travers différentes stratégies qui dépendent des objectifs de l'étude. En termes de relevé, cette étape s'inscrit dans le moment de ***figuration***.

2.3.1. Traitement des données 2D

L'exploitation des données 2D brutes d'acquisition passe par des traitements permettant de mettre en évidence davantage d'informations. Il existe un grand nombre de méthodes de traitement d'images utilisées en fonction du type d'informations attendu. Une image numérique peut par exemple permettre d'obtenir des informations concernant la luminance, le contraste, les couleurs, la netteté, mais aussi permettre, de détecter la présence ou non d'objets dans l'image à l'aide d'algorithmes de segmentation, de classifier les pixels par familles ou encore d'extraire des contours. Cette liste n'est absolument pas

exhaustive, et elle s'allonge si l'on considère que les images produites par différents capteurs exploitent l'encodage colorimétrique pour contenir des informations qui concernent des aspects non-visuels de l'objet d'étude : c'est par exemple le cas des photographies issues de caméras thermiques. Ces traitements peuvent être manuels et issus de procédés empiriques, semi-automatiques, ou automatiques.

2.3.1.1. Traitements manuels, retouches

Les traitements manuels visent à modifier ou corriger l'aspect visuel des images à travers la manipulation des couleurs ou des valeurs pour faire ressortir certains attributs (Figure 32). Les images sont principalement traitées à l'aide de logiciels spécialisés (par exemple Photoshop, Gimp, ou ImageJ) en jouant sur la luminosité, le contraste, le seuil. La séparation des trois canaux de couleurs (rouge, vert, bleu) est également exploitée dans le cas des techniques de composition d'images en fausses couleurs, qui consistent à affecter aux bandes spectrales de l'image des couleurs ne correspondant pas aux trois couleurs primaires. Cette technique est particulièrement employée lors d'acquisition de données multispectrales.

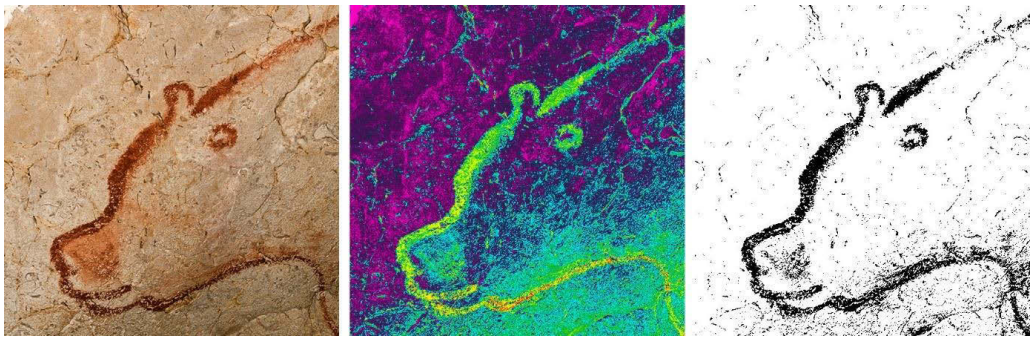


Figure 32 : Exemples de traitements manuels d'une image. À gauche : image originale (photographie © CNP/MC). Au centre, traitement en fausses couleurs. À droite : seuillage. Illustration : Auteur.

Néanmoins, les traitements manuels sont souvent des procédés empiriques, qui peuvent être considérés comme incompatibles avec la recherche d'objectivité qui caractérise certaines méthodes de relevé. Grimaud (2015) explique, en parlant du relevé d'objets d'époques plus tardives (mégalithes datant du néolithique), que *« l'objectif des relevés va constamment osciller entre la recherche de " l'objectivité absolue " à travers l'automatisation des méthodes et la réduction de l'intervention de l'opérateur, et une certaine subjectivité afin de retrouver le tracé originel. »* Ce constat semble partagé au relevé d'art pariétal (Cerrillo-Cuenca et al., 2014).

2.3.1.2. Calibrage des couleurs

Lors de prises de vues photographiques, il est presque impossible de garantir que les conditions lumineuses restent strictement identiques. Sur les images, cela se traduit par des différences d'intensité et de couleurs (qui seront également présentes sur les textures des données 3D résultantes). Cela peut poser un problème dans la mesure où l'extraction d'informations à partir d'une image passe par l'interprétation des données qu'elle comprend : les couleurs dans le cas qui nous intéresse. Ainsi, sans calibrage il sera impossible de comparer rigoureusement deux photographies, par exemple en vue de suivre l'évolution d'une dégradation, car elles n'auront pas la même référence. L'objectif du calibrage

des couleurs est de remédier à ce problème en corrigeant les couleurs des différentes photographies afin d'approcher un modèle de référence.

2.3.1.3. Decorrelation stretching

Très couramment utilisé comme méthode d'accentuation visuelle, le *decorrelation stretching*[16] est un processus permettant d'améliorer (étirer) les différences de couleur trouvées dans une image numérique. La méthode utilisée pour ce faire inclut l'élimination de la corrélation intercanaux trouvée dans les pixels d'entrée, d'où l'expression *étirage de décorrélation*. Initialement l'algorithme avait pour but de faciliter la lecture d'images multispectrales (Alley, 1999). Dans le domaine de l'art pariétal, cette méthode est très prisée du fait de son efficacité pour faciliter la visualisation des peintures. Elle est habituellement exploitée via le *plugin* Dstretch pour ImageJ, conçu par Jon Harman en 2005 (Le Quellec et al., 2013).

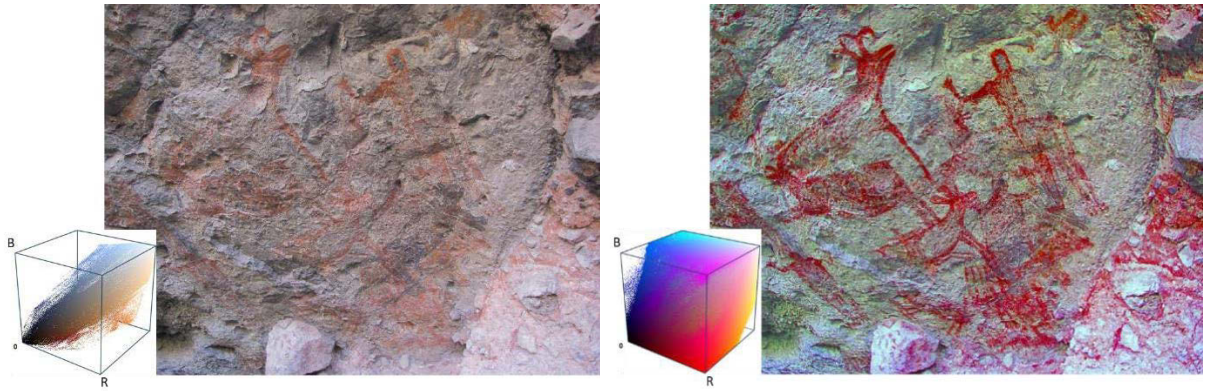


Figure 33 : À gauche : image d'entrée (Cueva Raton, Californie) et répartition des pixels dans l'espace colorimétrique. À droite : image résultant de son traitement par *decorrelation stretching*. Illustration : d'après Jon Harman.

Une image numérique comprend, pour chaque pixel, un triplet de valeurs permettant d'en définir la couleur (le plus souvent R, G, B). Chaque triplet peut être considéré comme les coordonnées du pixel dans un espace orthonormé à trois dimensions. À partir de cet ensemble de données, le traitement consiste à d'identifier de manière ordonnée les axes qui portent le plus d'informations, par une Analyse en Composantes Principales (PCA). Ces axes sont utilisés pour *étirer* l'ensemble de données, afin qu'elles remplissent l'espace colorimétrique, ce qui a pour effet d'accentuer considérablement les écarts entre familles de couleurs présentes sur l'image. En résulte une image aux couleurs non réalistes mais sur laquelle les éléments peints sont bien plus visibles que sur l'image initiale (Figure 33). Cette image est souvent utilisée comme support ou aide au relevé de figures pariétales. Selon les espaces colorimétriques utilisés, les effets du traitement sont différents. Ainsi, certains espaces colorimétriques sont à privilégier si l'on souhaite améliorer la visualisation d'éléments rouges, d'autres sont plus efficaces pour les jaunes, etc.

Cette méthode est privilégiée aux traitements manuels d'une part parce qu'elle est plus efficace, et d'autre part pour sa répétabilité. Cependant, les utilisateurs peuvent agir sur certains paramètres comme l'échelle utilisée pour l'étirage ou la zone de l'image sur laquelle est appliquée la PCA, voire même proposer des espaces colorimétriques personnalisés. La répétabilité des traitements dépend alors de la façon dont ces choix sont documentés.

[16] Ou « étirage de décorrélation » en français, mais la traduction n'est jamais utilisée.

2.3.1.4. Segmentation

Si l'humain sait intuitivement séparer et identifier des éléments dans une image, la mise en place des algorithmes permettant d'informatiser ce processus est loin d'être triviale et reste encore aujourd'hui l'un des thèmes les plus courants en traitement d'images. Le but de la segmentation est de simplifier ou de transformer une image d'entrée en une représentation significative plus facile à analyser, en la partitionnant en groupe de pixels homogènes selon certains critères. Les pixels ainsi regroupés forment des régions. Cette méthode peut viser différents objectifs mais dans le cas du relevé numérique il s'agit principalement d'extraire des informations de l'image pour en proposer un découpage susceptible d'être utilisé pour l'enrichissement sémantique (section 2.4). Ainsi, l'attribution d'un label à chaque zone caractéristique permettra d'assigner un sens spécifique aux pixels qui la composent. Dans le cas de l'art rupestre, la segmentation est utilisée dans les travaux de reconnaissance automatique de pétroglyphes (Figure 34).

Figure 34 : Méthode proposée par Zeppelzauer et al. (2016) pour la segmentation interactive de pétroglyphes. Illustration : (Zeppelzauer et al., 2016).

Il existe de nombreuses méthodes de segmentation d'images, permettant de faciliter le processus d'extraction d'informations. Elles sont généralement regroupées en quatre catégories : les segmentations par régions (*region-based*), par contours (*edge-based*), par seuillage (*threshold*), et les segmentations combinant les méthodes précitées. Cependant, dans le cas d'objets patrimoniaux leur utilisation est souvent délicate du fait de l'unicité des objets, et de leur complexité, notamment morphologique. Ainsi, chaque cas d'étude suppose le développement de stratégies adaptées à ses spécificités, par exemple par le traitement préalable des images d'entrées, ou l'exploitation conjointe d'informations colorimétriques et géométriques (Manuel, 2016).

Nous avons étudié dans cette section différents traitements visant à extraire des informations à partir d'images 2D. Si ces approches sont souvent appliquées à des photographies, elles ne se limitent pas à ces seules ressources. De nombreuses autres images 2D peuvent en effet être produites à partir des acquisitions 3D, comme nous le verrons dans la section suivante. Nous étudierons également différentes méthodes d'extraction d'informations basées sur le traitement de données 3D, employées lors de relevés numériques d'art pariétal pour faciliter le travail d'interprétation.

2.3.2. Traitement des données 3D

La construction et l'utilisation de données 3D permettent d'adopter de nouveaux points de vue concernant l'objet d'étude, et de réaliser des opérations complexes voire impossibles dans la réalité. Par la possibilité d'étendre les perceptions de l'observateur, ces données tendent à mettre en évidence des propriétés parfois inattendues de l'objet d'étude et ainsi à aider la formation d'hypothèses de recherche.

Les traitements usuels de données 3D s'intéressent à ce titre à la mise en évidence d'informations géométrico-visuelles. Il existe un grand nombre de méthodes, sollicitées selon le type d'informations attendu et en prévision de leur utilisation future au sein du processus de relevé.

2.3.2.1. Extraction de vues par projection géométrique

En traitant des techniques de relevé *classique*, nous avons pu voir que la distance aux parois imposée par la nécessité de les préserver soulevait des questions importantes en termes de justesse géométrique. Plus précisément, il s'agit de limiter un maximum les déformations liées au système de projection utilisé, ce qui peut s'avérer compliqué lorsque les surfaces des parois sont tortueuses. Le premier apport de la numérisation 3D est de faciliter le relevé en permettant aux releveurs de choisir librement le système projectif qui leur convient le mieux pour une activité donnée. En effet, à partir d'un modèle 3D ou d'un nuage de points, il est aisé d'extraire des vues spécifiques en maîtrisant le système de projection. Les plus courantes sont la projection perspective à un, deux ou trois points de fuite, la projection parallèle, et les projections panoramiques cubiques, cylindriques et sphériques (De Luca, 2009).

- **Projection perspective :**

La **projection perspective** est une projection linéaire où les points de l'espace sont projetés sur un plan image. Les objets ainsi représentés apparaissent d'autant plus petits qu'ils sont éloignés, un objet situé à l'infini étant assimilable à un point. On distingue trois cas, liés au nombre de points de fuite. La projection parallèle est un cas particulier de projection perspective, où les lignes de visée ou rayons sont parallèles entre elles. La projection est dite orthographique si les rayons sont perpendiculaires au plan de l'image. La **projection parallèle** est particulièrement utilisée pour la représentation du bâti, notamment dans le cas de la production d'orthophotographies et d'élévations.

Figure 35 : À gauche : projection perspective. À droite : projection parallèle. Facsimilé du panneau des Chevaux de la grotte Chauvet Pont-d'Arc. Illustration : auteur.

- **Projections panoramiques :**

La **projection cubique** consiste à projeter l'espace de la scène sur les six faces d'un cube de centre *C*. Chaque point P_n de l'espace est alors représenté par le point construit par l'intersection de la droite (CP_n) avec une face du cube (Figure 36 gauche). Chaque face du cube correspond ainsi à une projection perspective dont le champ est de 90° x 90° et de centre optique *C*. La projection résultante correspond à une représentation développée des six faces du cube. Ce mode de projection est couramment exploité pour la création d'applications de visualisation d'images panoramiques interactives.

La **projection cylindrique** est une transformation géométrique qui consiste à projeter les points de l'espace sur la surface d'un cylindre droit dont la hauteur et le périmètre de la section correspondent

respectivement à la hauteur de l'image panoramique générée et à sa largeur. Si est *C* le milieu de la hauteur du cylindre, chaque point P_n de l'espace est représenté par le point construit par l'intersection de la droite (CP_n) avec la surface cylindrique (Figure 36 centre). L'espace représenté est donc distribué sur 360° sur l'axe horizontal. Il est généralement réduit sur l'axe vertical du fait des déformations causées aux limites.

La **projection sphérique** consiste à projeter tout point de l'espace sur une sphère. Si *C* est le centre de la sphère, chaque point P_n de l'espace est alors représenté par le point construit par l'intersection de la droite (CP_n) et la surface de la sphère. L'espace représenté est distribué sur 360° sur l'axe horizontal et 180° sur l'axe vertical (Figure 36 droite). Ce mode de projection est notamment exploité en lasergrammétrie pour stocker les mesures de chaque station sous la forme d'images distinctes.

Figure 36 : Les modes de projection panoramique. À gauche : Projection en faces de cube. Au centre : Projection cylindrique. À droite : Projection sphérique. Illustration : auteur, à partir d'une image panoramique de l'abri du Cap-Blanc © CNP/AGP.

Concernant la production d'images par l'utilisation conjointe des numérisations 3D et de ces méthodes de projection, l'orthophotographie est sans aucun doute le support le plus utilisé. L'image est créée selon un plan de référence choisi par l'utilisateur (dans le cas de l'art pariétal, il s'agit le plus souvent du plan moyen de la portion de paroi étudiée), l'axe de prise de vue étant normal à ce plan. L'intérêt majeur de l'orthophotographie est de corriger les déformations dues aux reliefs, les erreurs de parallaxe des photographies, et les distorsions des objectifs. Dans le cas où le plan de référence est horizontal, l'orthophotographie se présente comme une image superposable à une carte. Ces images tendent donc à aplanir l'objet d'étude et permettent d'éviter les effets de perspective.

Lorsqu'elles sont mises à l'échelle, elles peuvent être utilisées pour réaliser des mesures. Pour ces raisons, l'orthophotographie est un support extrêmement prisé dans l'étude d'art pariétal, et est exploitée comme support de relevé, que ce soit dans le cadre de relevés analytiques plastiques ou de relevés analytiques techniques. Dans le cas du relevé technique, cette image est même parfois reprojetée sur la paroi à l'aide d'un vidéoprojecteur positionné suivant l'axe utilisé pour la projection orthographique, et sert de support cartographique aux observateurs munis de tablettes graphiques tactiles (Pinçon et al., 2018).

Outre les orthophotographies, tous types de vues géométrales tels qu'utilisés en architecture peuvent être extraits des données 3D. Il est ainsi possible de produire des images rendant compte des dispositions, des formes, et des dimensions de l'objet étudié sans tenir compte de la perspective, par le biais de plans, d'élévations, ou de coupes. Dans le cas de l'art pariétal, ces données sont produites pour illustrer des informations contextuelles concernant la morphologie de la grotte et sa spatialité, et également servir de support à certaines informations concernant la localisation des panneaux ornés ou des mesures réalisées.

Enfin, les données 3D sont couramment utilisées pour produire des images (captures d'écran) suivant des points de vue impossibles ou difficiles en réalité, à des fins d'illustration. Globalement bien que cela soit difficile à quantifier, la production de représentations 2D semble demeurer l'utilisation principale des données 3D. Ce fait peut s'expliquer d'une part par les paradigmes de représentation hérités de la Renaissance, mais également comme un choix par défaut, en l'absence d'outils adaptés pour leur interprétation et leur enrichissement. Comme nous le verrons par la suite, il demeure aujourd'hui plus simple de travailler sur de l'image 2D que sur de la donnée 3D, bien que cela implique souvent une perte d'informations que l'on ne récupère qu'en multipliant les points de vue traités.

2.3.2.2. Prise de mesures

Une autre utilisation incontournable lors de relevé numérique est la prise de mesures à partir de points caractéristiques. Bien que techniquement l'élaboration d'un nuage de points 3D constitue déjà le résultat d'une série de mesures d'éléments caractéristiques, **la quantité d'informations contenue est trop importante pour être réellement intelligible pour un observateur car les points mesurés ne sont pas nécessairement tous signifiants par rapport à son système d'analyse**.

L'extraction de mesures métriques à partir des données 3D acquises constitue alors un moyen pour l'observateur de comprendre son objet d'étude par sa réduction à une série d'éléments signifiants au regard d'une préoccupation donnée. Ces mesures peuvent être de différentes natures : extraction des coordonnées d'un point, mesure d'angles, de distances, ou encore de différences de niveaux. Dans tous les cas, leur calcul repose sur la sélection préalable de points d'intérêt et l'extraction de leurs coordonnées, et nécessite que les données d'entrée soient à l'échelle voire géolocalisées.

Figure 37 : Le *web-viewer* 3DHop, utilisé pour le projet 3D Cuneiform. Illustration : (O'brien et al., 2019).

Cette opération repose sur la libre manipulation des données 3D à l'aide de logiciels de visualisation tels que CloudCompare, Meshlab ou encore Blender. Cependant, comme le notent Thivet et al. (2019), depuis la démocratisation du HTML5 et de l'API WebGL, en termes de visualisation 3D les usages se tournent progressivement vers des solutions web, par exemple Sketchfab, Autodesk A360 ou 3DHop (Figure 37), pour des raisons de praticité et flexibilité.

2.3.2.3. Maillage, texturage, décimation

Comme nous l'avons vu précédemment, les données résultant des opérations de numérisation 3D se présentent sous la forme de nuages de points. En l'absence de logiciel spécifiquement dédié au relevé

patrimonial, les travaux de traitement des données sont réalisés à l'aide de logiciels empruntés au domaine de l'infographie 3D qui exploitent majoritairement des modèles 3D maillés (Figure 38). Différentes approches permettent de passer du nuage de points au maillage. La polygonisation ou triangulation 3D (*meshing*) permet de créer un maillage non structuré à l'aide d'algorithmes comme *Delaunay* (Chew, 1993), ou *screened Poisson* (Kazhdan, Hoppe, 2013). À l'issue de cette opération, le maillage 3D obtenu comporte des sommets, des arêtes et des faces, et représente l'ensemble des objets scannés de manière indifférenciée. Pour les identifier, il faut recourir à la segmentation.

Dans certains cas, le maillage est obtenu par modélisation 3D manuelle en utilisant le nuage de points comme repère. Dans ces cas, les objets distincts sont bien différenciés sur le modèle 3D résultant, mais cette méthode est plus longue et surtout adaptée à des objets de géométrie relativement simples (éléments architecturaux réguliers, éléments plans et orthogonaux, surfaces de révolution, etc.). Dans le cas du relevé d'art pariétal, la modélisation manuelle est exclue du fait de la complexité morphologique des objets étudiés.

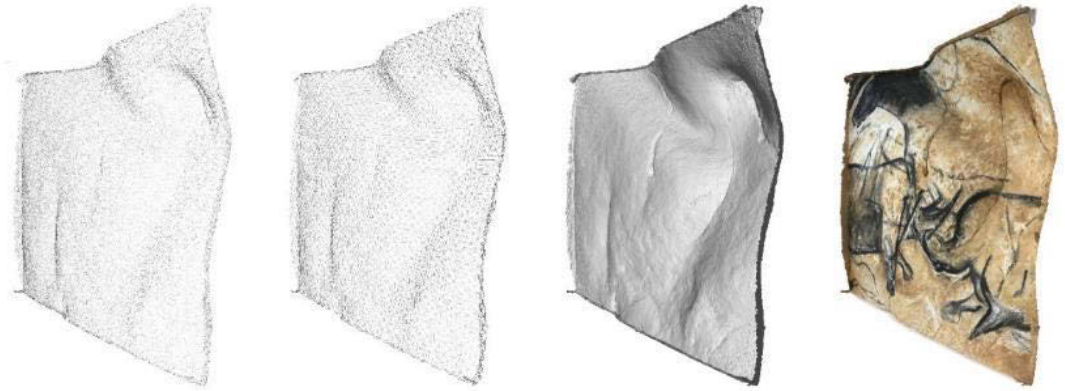


Figure 38 : Maillage et texturage d'un nuage de points du facsimilé. Illustration : auteur.

Différentes étapes de postproduction suivent la génération du maillage. En premier lieu, ce dernier peut être *nettoyé*. Cette opération vise à détecter et boucher les trous éventuels, à déplacer ou supprimer les nœuds qui provoquent des incohérences (faces qui s'intersectent), ou encore à supprimer les éléments déconnectés et les doublons. Le type de maillage peut également être modifié en passant d'un maillage triangulaire à un maillage par quadrilatères (Tarini et al., 2010), par hexagones (Ray et al., 2017) ou par d'autres types de polygones (Owen, 2000). On parle alors de *simplification*.

La décimation est une étape de postproduction visant à réduire le nombre de faces d'un maillage (ou le nombre de points d'un nuage) en minimisant autant que possible les modifications de l'aspect de la surface. Cette étape est souvent incontournable pour alléger les fichiers et permettre une manipulation plus fluide. Néanmoins, elle s'accompagne nécessairement d'une perte d'informations et donc de précision. Pour cette raison, le travail de traitement réalisé sur les données 3D doit être bien documenté car il conditionne leurs exploitations ultérieures.

2.3.2.4. Enrichissement visuel

La manipulation de maillages 3D très denses ou tout simplement de fichiers 3D volumineux peut s'avérer contraignante, car les performances des machines sont souvent insuffisantes pour assurer une navigation fluide tout en assurant un niveau de qualité d'affichage très élevé. Nous avons vu qu'une

technique pour faciliter la manipulation est de décimer le modèle 3D pour l'alléger, mais cela implique une perte de précision non négligeable. Pour compenser cette perte, une technique est d'utiliser des méthodes d'enrichissement visuel par cartes de textures pour simuler graphiquement des détails géométriques ou visuels obtenus à partir de modèles très résolus, sur des fichiers plus légers sans ajouter de polygones à la géométrie des objets (Figure 39). Il s'agit donc d'améliorer la qualité de rendu graphique en limitant la consommation des ressources de calcul et d'affichage.

Visualisées en tant qu'images, ces cartes procurent également des informations intéressantes concernant l'objet d'étude, car elles correspondent à la projection en 2D de certains de ses attributs géométriques ou visuels. Cela permet d'en apprécier visuellement les caractéristiques, mais aussi de bénéficier des méthodes usuelles de traitement d'images pour manipuler de l'information 3D. Dans le cas de l'art rupestre, les méthodes d'enrichissement visuel permettent notamment de faciliter la visualisation des gravures par le biais des simulations lumineuses en lumière rasante (Cassen, Robin, 2010). L'analyse des normales d'un objet tridimensionnel peut permettre de faire ressortir d'autres informations importantes, elles facilitent par exemple la détection de surfaces planes ou parallèles par la recherche d'éléments dont les directions de normales sont identiques. De même, elles permettent de détecter des éléments saillants comme les ruptures de pente.

Dans le cadre de l'étude de l'abri sous roche de Marsoulas, le traitement des textures du modèle 3D a permis de nettoyer virtuellement les parois des graffitis du XIX^{e} siècle pour faciliter la visualisation des traces anthropiques anciennes (Fritz et al., 2010). Zeppelzauer et al. (2016) présentent une approche de segmentation 3D automatique permettant d'extraire les gravures rupestres des surfaces 3D reconstruites. Pour simplifier et accélérer le calcul, ils exploitent notamment les cartes de profondeur pour stocker la donnée 3D d'entrée sous forme d'images 2D, et supporter les opérations de segmentation et classification.

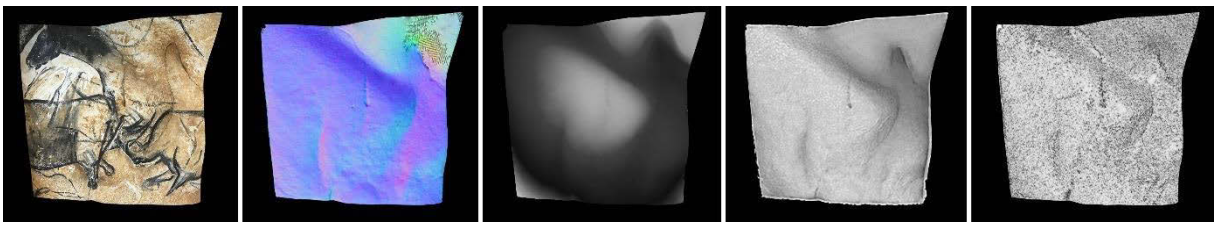


Figure 39 : Exemples de cartes de texture caractérisant des attributs géométriques ou visuels de l'objet d'étude. De gauche à droite : couleurs, normales, profondeur, occlusion ambiante, rugosité.

2.3.2.5. Accentuation des attributs géométriques par l'utilisation de *shaders*

Les cartes énoncées précédemment sont essentiellement utilisées dans le cadre de rendus physiques réalistes (PBR), avec le but de simuler de manière la plus réaliste possible le comportement des objets 3D vis-à-vis de certains paramètres, le plus souvent la lumière. Cette méthode de rendu est particulièrement utile pour manipuler la donnée 3D d'une façon qui serait extrêmement complexe voire impossible en réalité.

Une autre méthode de rendu existe cependant et trouve des applications intéressantes dans le domaine de l'archéologie, il s'agit du rendu non-photoréaliste (NPR) aussi appelé *rendu expressif*. Celui-ci s'intéresse à la mise en valeur de caractéristiques des scènes 3D représentées dans le but de faciliter la lisibilité des images produites sans s'attacher à garantir leur fidélité au réel. Il permet également de

représenter des hypothèses de restitution sans biaiser leur interprétation par un rendu trop réaliste (Vergne et al., 2007). Dans cette optique, certains *shaders* permettent d'influencer l'aspect des objets 3D et trouvent des usages directs dans l'extraction d'informations lors de relevés d'art pariétal. Les *shaders* sont des programmes informatiques permettant de déterminer, à partir de paramètres, les propriétés d'une scène 3D (aspects des surfaces, lumière, etc.) pour en permettre le rendu.

- **Occlusion ambiante**

Dans la réalité, les objets réfléchissent une partie de la lumière qu'ils reçoivent, et deviennent, ainsi, des sources de lumière indirecte pour les éléments environnants. Des techniques de rendus comme l'illumination globale visent à tenir compte de ces effets, mais cela exige des calculs importants, ce qui est contraignant en termes de temps et de ressources. La technique de l'occlusion ambiante a été développée pour tenir compte des effets de la lumière ambiante et contribuer au réalisme de l'éclairement de la scène 3D étudiée, tout en simplifiant les calculs (Shanmugam, Arikan, 2007).

Figure 40 : Détail du facsimilé, rendu avec l'occlusion ambiante.

Concrètement, la technique consiste à évaluer la quantité de lumière ambiante qui parvient à chaque point de la scène, afin de simuler l'assombrissement des zones les plus difficiles d'accès. Visuellement, l'effet rend sur les différents éléments les ombres douces créées par les obstacles à la lumière indirecte (Figure 40). Si initialement cette approche visait à améliorer le photoréalisme du rendu, elle est également utilisée comme technique de visualisation pour apprécier visuellement le relief des objets 3D. Tout comme le *Radiance Scaling*, l'occlusion ambiante est particulièrement intéressante pour visualiser de légères variations de surfaces, par exemple des gravures ou des bas-reliefs (Novaković et al., 2017).

- **Eye-Dome Lighting**

Figure 41 : Nuage de points d'un détail du facsimilé, rendu avec le *shader* EDL.

Bien que plus performante que l'illumination globale, l'occlusion ambiante demeure coûteuse car elle nécessite souvent d'importants précalculs. La technique du *Eye-Dome Lighting* ou EDL (Boucheny et al., 2011) est une alternative destinée à la visualisation scientifique. Non photoréaliste, elle s'inspire de l'occlusion ambiante en la rendant dépendante du point de vue de l'observateur pour simplifier et accélérer le calcul. Elle peut être appliquée aux nuages de points 3D et est combinable avec d'autres techniques de rendu dans l'idée de faciliter la compréhension de scènes complexes par le renforcement des relations spatiales et des reliefs (Figure 41).

- **Exaggerated shading**

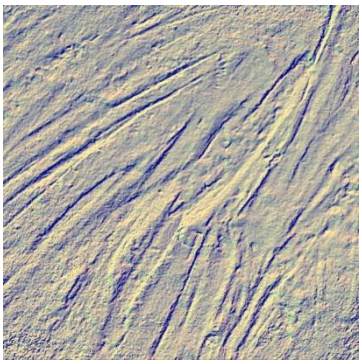


Figure 42 : Détail du facsimilé rendu à l'aide de XShade.

Proposé par Rusinkiewicz et al. (2006) le Xshade ou *Exaggerated Shading* (shading exagéré) est une technique de rendu expressif inspiré des techniques de relief cartographiques. Elle se base sur l'ajustement dynamique de la position de la source lumineuse pour ombrer différentes zones de la surface de l'objet. Son intérêt majeur est de révéler des détails de la surface de l'objet indépendamment de leur orientation (Figure 42). Il est notamment utilisé par Carrero-Pazos et al. (2018) pour faciliter la visualisation et le relevé des fines gravures sur le site d'art rupestre *Laxe da Forneiriña II* (âge du Bronze, Galice).

- **Radiance scaling**

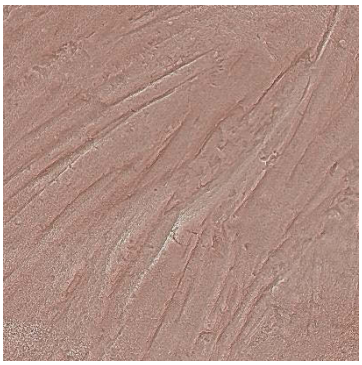


Figure 43 : Détail du facsimilé rendu à l'aide du *shader* Radiance Scaling.

Depuis son implémentation dans le logiciel *Meshlab*, le *shader Radiance Scaling* (Vergne et al., 2010) est devenu très populaire pour la mise en valeur d'éléments gravés (Figure 43). Il s'agit d'une technique de rendu expressif basée sur l'ajustement de l'intensité de la lumière réfléchie en fonction de la courbure de la surface de l'objet 3D étudié et des caractéristiques de son matériau. Elle est par exemple utilisée par Porter et al. (2016), Vilas-Estevez et al. (2017), Carrero-Pazos et al. (2018) pour la visualisation de gravures fines. L'intérêt majeur de cette technique est qu'elle améliore la visualisation des concavités et convexités, notamment lorsqu'elle est utilisée avec des *Lit Spheres* différentes pour les zones concaves et convexes[17].

Il existe de nombreux autres *shaders* exploitables dans différents contextes, selon que l'utilisateur souhaite obtenir un rendu réaliste ou au contraire accentuer certaines caractéristiques de l'objet manipulé. L'industrie vidéoludique et plus globalement le domaine de l'infographie en fournissent de nombreux exemples. Concernant les techniques d'accentuation du relief que nous n'aurions pas énoncées ici, le lecteur pourra se référer notamment à Ammann et al. (2012) ou Grimaud (2015).

2.3.2.6. Simulations lumineuses

Facilité par l'utilisation des techniques d'enrichissement visuel énoncées précédemment, un usage important de la numérisation 3D pour l'extraction d'information est la simulation des éclairages sur les objets d'étude. De fait, les éclairages rasants sont particulièrement efficaces pour révéler certains détails comme les gravures et ne sont pas forcément envisageables dans la réalité. Mais d'un point de vue archéologique, la lumière est aussi considérée comme un facteur fondamental, notamment pour la compréhension des sociétés passées. Dans la perception des œuvres pariétales, elle est considérée comme un paramètre essentiel, notamment pour les sculptures qui jouent avec les effets d'ombre. Une simulation d'éclairage a été menée pour étudier les sculptures de l'abri du Cap-Blanc (Devlin et

[17] Les *Lit Spheres*, ou *MatCaps*, sont des images représentant des sphères, utilisées pour définir des matériaux et leurs propriétés physiques. Cette technique, basée sur (Sloan et al., 2001) a été pensée pour permettre de créer simplement des styles très variés.

al., 2003). L'objectif de la simulation était de vérifier si l'éclairage du site par une lampe à graisse peut laisser supposer un travail concernant le mouvement sur les sculptures. Une simulation plus récente a été réalisée pour la grotte ornée de Ardales en Espagne (Hoffmeister, 2017). L'auteur s'appuie sur une numérisation 3D obtenue par lasergrammétrie et sur l'emplacement des vestiges des lampes à graisses pour simuler l'éclairage des parois et évaluer l'incidence de chaque source lumineuse sur la visibilité des représentations. D'autres simulations ont été menées dans le cadre d'études de sites d'époques plus tardives, par exemple pour la grotte Fugoppe située au nord du Japon (Masuda et al., 2006), ou encore pour tester des hypothèses archéoastronomiques concernant l'orientation solaire de mégalithes néolithiques (Lescop, 2015).

2.3.2.7. Simulations physiques

Un autre atout des données 3D est de se prêter à un grand nombre de manipulations physiques proscrites dans la réalité. Les objets numérisés peuvent notamment être virtuellement déplacés et supporter de multiples transformations géométriques, pour réaliser des comparaisons de blocs ou des simulations d'assemblage. Ces techniques ont été sollicitées à plusieurs reprises au Roc-aux-Sorciers et à la Chaire-à-Calvin. Une anastylose virtuelle a été menée sur les vestiges de la voûte ornée effondrée, afin de tester virtuellement les raccords (Figure 44), évalués par le degré d'adhérence par rapport à un raccord idéal (Pinçon et al., 2010). Dans l'abri Bourdois, Bourdier et al. (2015) ont eu recours à la comparaison de blocs pour confronter différentes représentations animales et assister l'interprétation de figures sculptées partiellement détruites.

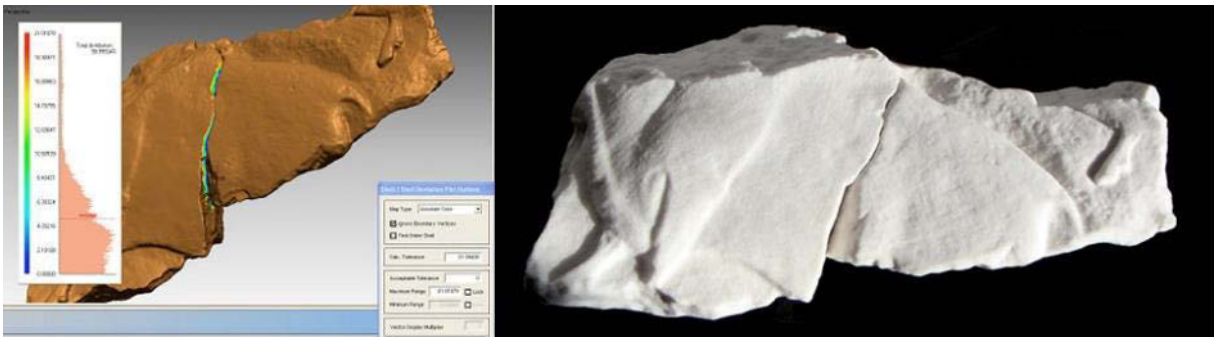


Figure 44 : Simulation de raccord réalisée avec des blocs du Roc-aux-Sorciers. À gauche : modèles 3D virtuels. À droite : impressions 3D. Illustration : (Bourdier et al., 2015). ©CNP / DRAC Poitou-Charentes / AGP.

Des numérisations 3D ont également été exploitées pour mener des simulations de dynamique des fluides, notamment à la grotte de Lascaux, où un simulateur a été établi afin de prédire l'influence de différents scénarios d'intervention sur les conditions thermiques et hydriques de la grotte, et faciliter sa conservation (Lacanette et al., 2009). D'autres simulations ont été menées dans le cadre du projet CarMoThap, dans le but d'étudier l'influence des différents combustibles sur les caractéristiques des altérations et des résidus observés sur les parois (Lacanette et al., 2017).

2.3.2.8. Suivi temporel

Chaque numérisation de l'objet d'étude permet de témoigner de son état de conservation pour une date donnée. Ainsi, si ce dernier a fait l'objet de plusieurs numérisations, il devient possible d'exploiter les données 3D pour étudier son évolution dans le temps et révéler par comparaison les dégradations,

déplacements, et autres modifications survenues (Peteler et al., 2015). Pour cela, les données 3D doivent partager le même référentiel. Il s'agit alors de produire une carte des écarts, c'est-à-dire une représentation permettant de visualiser la distance séparant deux modèles 3D[18] par l'utilisation d'un dégradé de couleurs. Cette approche est particulièrement intéressante pour la conservation, car elle permet de quantifier les changements survenus sur l'objet d'étude tout en offrant une visualisation relativement intuitive de ces informations. Dans le cas de l'art pariétal, cette approche est notamment exploitée par Ruiz (2016) (Figure 45).

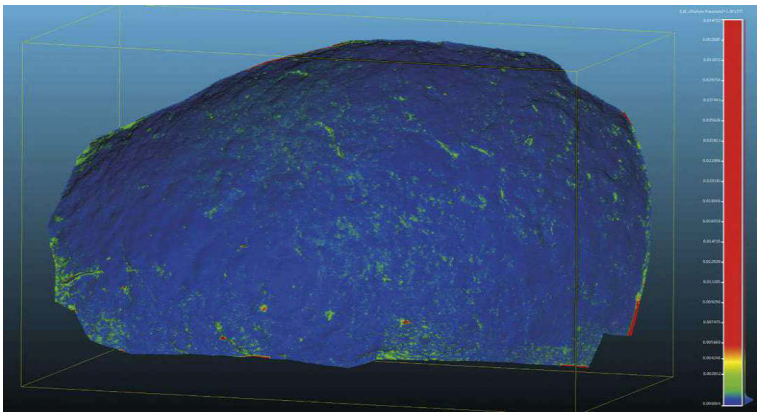


Figure 45 : Comparaison de deux modèles 3D produits à deux mois d'intervalle. L'image révèle des écarts correspondant à des pertes de matière. Abrigo del Buen Aire (Espagne). Illustration : (Ruiz, 2016)

2.3.3. Bilan des méthodes d'extraction d'informations

Cet état de l'art des méthodes de traitement pour l'extraction d'informations montre bien à quel point les données numériques et les capacités computationnelles de l'informatique peuvent bénéficier à l'étude de l'art pariétal. Nous avons pu constater qu'avec de bons outils d'analyse, les attributs géométriques et visuels stockés dans des copies numériques des objets d'étude pouvaient largement contribuer à la production de connaissances dépassant le seul champ de l'objet en lui-même.

Cependant, cette section nous révèle aussi que les pratiques actuelles mobilisent une très grande variété d'outils logiciels et de compétences, ce qui cause d'importants problèmes d'interopérabilité, de traçabilité, et surtout, de dispersion des données. Nous pouvons également observer que tous ces processus sont informatisés et réalisés *ex situ*, principalement du fait des ressources matérielles requises par ces calculs exigeants. D'autre part, les données sont très majoritairement manipulées de manière isolée de leur contexte. Cette situation cause une rupture significative entre les environnements d'étude réels et virtuels, qui n'est aujourd'hui compensée que par la multiplication des allers-retours sur site/hors site des différents acteurs.

Dans ce contexte, la possibilité de croiser des regards au sein d'équipes pluridisciplinaires repose essentiellement sur notre capacité à intégrer les connaissances des différents acteurs au sein des données numériques. Nous étudierons dans la section suivante les approches existantes d'enrichissement sémantique. Nous pourrons observer que, contrairement aux phases d'acquisition et d'interprétation, l'enrichissement sémantique est souvent négligé dans le processus de relevé numérique, faute d'outil approprié.

[18] Cette opération peut être réalisée par le biais de logiciels tels que CloudCompare, Meshlab, Geomagic, etc. La distance entre les deux jeux de données 3D est calculée à l'aide d'outils mathématiques tels que la distance de *Hausdorff*.

2.4. Enrichissement sémantique de données numériques

La mise en relation des données géométriques et visuelles avec les connaissances propres à l'observateur est primordiale d'une part pour donner lieu à des formalisations sémantiques permettant de mettre en relation les observations avec leurs supports, et d'autre part, pour identifier de potentielles corrélations entre les observations et générer de nouvelles connaissances par le croisement des regards issus d'équipes pluridisciplinaires. Nous nous intéressons ici aux méthodes et solutions permettant l'enrichissement sémantique des données traitées lors de relevés.

Le terme *sémantique* est emprunté à la linguistique, et plus particulièrement à la *lexicologie*. Cette dernière, qui se définit comme l'étude scientifique du lexique ou « *science du mot* » (Ullmann, 1959), comporte deux subdivisons selon qu'on s'intéresse aux mots ou aux sens. L'aspect formel des mots est étudié par le domaine de la *morphologie*, et les significations par le domaine de la *sémantique* (CNRTL, 2012). Par extension à d'autres domaines disciplinaires que l'étude du langage, le terme *sémantique* renvoie donc à l'étude des signifiés, c'est-à-dire aux sens que portent les éléments étudiés. Nous pouvons donc observer une dualité entre le signifié et le signifiant, autrement dit entre le fond et la forme, tous deux complémentaires.

Comme nous l'avons vu précédemment, en informatique, l'expression *fossé sémantique* désigne *« le manque de concordance entre les informations que les machines peuvent extraire depuis les documents numériques, et les interprétations qu'ont ces mêmes données pour un utilisateur dans une situation déterminée »* (Smeulders et al., 2000). Selon Boucher et Le (2005), il s'agit donc d'un problème de liaison entre traitements de bas niveau (traitement des propriétés objectives de l'objet dirigé par les données) et de haut niveau (traitement qui correspond à l'interprétation selon des attentes spécifiques, dirigé par les concepts). Conformément au paradoxe de Moravec, l'informatique peut se montrer très efficace concernant le premier aspect et peiner sur le second, au contraire de l'humain pour qui le traitement de haut niveau s'avèrera plus intuitif et plus rapide.

Dans le cas qui nous intéresse, celui du relevé, l'enrichissement sémantique correspond à la mise en correspondance entre le fond et la forme, c'est-à-dire entre les caractéristiques géométrico-visuelles des données issues des procédures d'acquisition 2D et 3D et l'assignation établie par un acteur selon un contexte et un domaine de connaissance donnés. Cette phase se situe à l'interface des moments de ***figuration*** et de ***représentation***.

Dans le cas du relevé d'art pariétal, le processus d'interprétation résulte traditionnellement en une représentation de l'objet étudié à l'aide d'une charte graphique ou de codes de représentation permettant de transcrire symboliquement l'information sémantique. Comme nous le verrons dans cette partie, selon les cas les documents résultants sont inégaux en termes d'interrogeabilité, les plus avancés d'entre eux pour cet aspect étant les cartographies vectorielles. Néanmoins, même dans ces cas, l'assignation des tracés dessinés par les releveurs demeure implicite et ne s'exprime qu'à travers la charte graphique choisie, lorsqu'elle est communiquée. Nous verrons que des avancées réalisées dans les secteurs du web sémantique commencent à s'étendre au domaine du patrimoine à travers des solutions d'annotation 2D ou 3D laissant entrevoir de nouveaux scénarios d'analyse et de corrélation, à l'interface entre la vision par ordinateur et l'analyse géométrique.

2.4.1. Approches 2D

2.4.1.1. Cartographies 2D

Nous avons vu précédemment que le relevé d'art pariétal comprend deux *familles* en termes de production de documents synthétiques, les relevés plastiques et les relevés techniques (section 1.2.2.4). En termes informatiques, la différence majeure entre les deux approches est que la première vise la production d'images *bitmap* quand la seconde vise la production d'images *vectorielles*. En pratique, il s'agit bien de processus d'enrichissement basés sur un support d'entrée et comportant deux phases, l'identification de régions spatiales et leur assignation.

Cependant, il semble que dans les deux cas, nous ne pouvons pas réellement parler d'enrichissement sémantique au sens informatique du terme, d'une part parce que l'information sémantique est implicite (lorsqu'elle est communiquée, elle l'est uniquement à travers une légende spécifiant les choix d'encodage graphique), et d'autre part parce qu'il s'agit de fichiers uniques indépendants, sans systèmes de gestion de bases de données, qui ne permettent pas d'informatiser l'analyse de la corrélation des observations sur la base des attributs sémantiques ou géométriques des sélections : le lien entre la donnée et la connaissance n'est pas concrètement établi. Nous ne cherchons pas ici à dire que ces documents ne permettent pas d'extraire des connaissances et d'aboutir à des résultats semblables à ce que l'on pourrait obtenir aux termes d'approches issues du web sémantique, il s'agit plutôt de souligner que ces résultats dépendent alors pleinement de la capacité de l'utilisateur des données à réaliser des croisements d'informations uniquement sur la base de processus cognitifs.

2.4.1.2. Annotation sémantique d'images

Depuis l'introduction de la notion de web sémantique (W3C, 2019), dans les différents domaines de l'informatique de nombreuses recherches s'intéressent aux modalités d'enrichissement sémantique, particulièrement à travers la création d'outils d'annotation. Le principe de l'annotation sémantique repose sur la mise en relation entre un document et des informations complémentaires, afin d'établir des relations entre les informations liées aux données plutôt qu'entre les données elles-mêmes. Autrement dit, il s'agit de passer d'une structure de données à une structure de connaissances, afin de permettre efficacement la recherche, le partage, et le croisement d'informations.

Des données enrichies sémantiquement permettent de réaliser des recherches par proximité sémantique plutôt que formelle. Dans le cas de la recherche d'images par exemple, cela permet de passer d'une recherche classique de type : *« je cherche des images dont l'histogramme ressemble à cette référence »* (critère quantitatif) à : *« je cherche des images de parois présentant comme traces anthropiques anciennes des représentations animales sculptées »* (critère sémantique incluant des relations entre plusieurs concepts). L'annotation sémantique n'est donc pas une fin en soi, mais une étape intermédiaire permettant de générer des métadonnées descriptives destinées à des utilisations à travers des traitements ultérieurs (Kiryakov et al., 2004).

L'enrichissement sémantique repose sur la mise en relation entre tout ou partie d'un document (signifiant) et des informations complémentaires issues de son interprétation (signifié). Elle comprend donc deux aspects : l'annotation (définition ou sélection spatiale des aspects formels concernés par la

description) et la description sémantique (assignation des éléments sélectionnés). Une annotation peut être (Figure 46) :

- **Globale** : la description est associée à l'ensemble du document, elle en décrit le contenu. Ce fonctionnement présente l'avantage de permettre un traitement par lot (c'est-à-dire de décrire simultanément plusieurs images) mais s'adapte mal à la description de scènes complexes.
- **Locale** : la description concerne des zones spécifiques du document (points, courbes, polygones, etc.). Ces méthodes sont plus précises mais présentent l'inconvénient d'être très chronophages, car elles ne permettent d'annoter qu'une ressource à la fois. Dans le cas du relevé cela peut s'avérer problématique car les contextes d'étude imposent souvent la manipulation de grandes quantités de données.

Figure 46 : Types d'annotations sémantiques d'images 2D. De gauche à droite : globale, ponctuelle, rectangulaire (boîte englobante), région (groupe de pixels). Photographie : Rhinocéros ZD1-20 du panneau des chevaux. ©CNP. Illustration : auteur, d'après (Tylecek, Fisher, 2018).

Concernant la description sémantique, Andrews et al. (2012) distinguent quatre modèles, qui correspondent à différentes méthodes pour associer une description aux éléments annotés :

- **Par Tags** (ou *étiquettes*) : Un tag est un mot-clé assigné à une ressource. Habituellement choisis de façon personnelle par l'auteur, les tags ne sont pas structurés, et l'ordinateur et les autres utilisateurs ne connaissent pas le sens que l'annotateur voulait lui donner sauf si le terme utilisé est sans équivoque. Les tags représentent un modèle d'annotation minimal du point de vue de la complexité structurelle et, par conséquent, ne peuvent permettre qu'un nombre limité de services principalement axés sur la navigation (par exemple, récupérer toutes les ressources dotées d'un tag spécifique). Parce qu'ils ne décrivent qu'implicitement les propriétés des ressources, les tags sont sujets à l'ambiguïté dans l'interprétation de ces propriétés. Par exemple, le tag "Rémy" attaché à une photo ne précise pas si Rémy est une personne sur la photo ou s'il en est l'auteur.
- **Par Attributs** : Un attribut est une paire d'éléments comprenant le nom de l'attribut et sa valeur, par exemple : `("date de l'observation", "22/10/1964")` ou `("auteur", "Julie")`. Ce modèle est plus intéressant que le modèle par tags car il permet de définir les types de données utilisées comme attributs, ce qui limite les ambiguïtés lors des requêtes. Pour reprendre l'exemple précédent, une requête du type `[auteur = Julie]` ne renverrait que des images réalisées par Julie et non les images d'autres auteurs sur lesquelles Julie apparaît. Contrairement aux tags, les attributs définissent donc explicitement les propriétés des ressources décrites et, par conséquent, permettent une annotation et un langage de requête plus riches. Notons cependant que les attributs ne permettent pas de caractériser les interrelations qui existent entre les ressources.
- **Par Relations** : Une relation est une paire d'éléments comprenant le nom de la relation et une ressource associée. Il s'agit d'une extension du modèle d'annotation par attributs permettant de

relier différentes ressources entre elles. Ce modèle est intéressant pour la navigation et propice à la sérendipité. Néanmoins, il ne décrit pas directement la ressource mais ses liens éventuels avec d'autres documents.

- **Par Ontologies** : Selon Borst (1997), une ontologie se définit comme *« une spécification explicite et formelle d'une conceptualisation partagée »*, c'est-à-dire un modèle conceptuel formel qui caractérise les phénomènes intervenant dans un domaine (concepts, relations, contraintes, règles, etc.), incluant ainsi leurs définitions et leurs contraintes. Ce modèle d'annotation permet de décrire et d'interconnecter des ressources en les qualifiant par des concepts et en définissant les relations, propriétés et restrictions qui les lient. Ce modèle est le plus sophistiqué et le plus complet. Néanmoins, il est plus complexe à mettre en place, notamment parce qu'il exige un consensus concernant les définitions.

Enfin, les techniques existantes pour l'annotation d'images sont divisées en trois catégories :

- **Annotations manuelles :** les régions sont tracées manuellement sur chaque image devant faire l'objet d'annotations, et les descriptions sont choisies par l'utilisateur (définies librement ou sélectionnées parmi des classes préétablies).
- **Annotations automatiques :** ces méthodes incluent la sélection automatique des régions d'intérêt sur les images, et leur description. Globalement, il s'agit d'apprendre automatiquement les concepts sémantiques à partir d'un grand nombre d'échantillons d'images, et de les utiliser ensuite pour décrire de nouvelles images.
- **Annotations semi-automatiques :** Ces méthodes sont des combinaisons de méthodes manuelles et de méthodes automatiques (Zomahoun, 2015). Elles consistent à faire intervenir les utilisateurs pour valider les décisions de la machine, afin de profiter des performances des méthodes automatiques tout en garantissant la pertinence des descriptions réalisées.

Aujourd'hui, le principal écueil est sans doute le temps nécessité par les méthodes d'annotation dont l'automatisation passe par l'exploitation d'algorithmes de segmentation peu adaptés à la complexité des cas d'études patrimoniaux, et encore moins aux grottes ornées dont les morphologies tortueuses limitent les approches basées sur la détection de *patterns* ou de discontinuités. Par ailleurs, il existe une grande diversité d'objets patrimoniaux, ce qui rend les approches générales difficiles à élaborer car elles nécessitent alors une très grande quantité de données de références, qui doivent être suffisamment représentatives. Les objets patrimoniaux ne font pas (ou très rarement) l'objet de productions en série, et présentent des variations morphologiques minimales qui rendent l'identification particulièrement délicate : deux objets peuvent par exemple avoir une allure globale similaire mais s'avérer très différents lorsque l'on s'intéresse au détail. Il peut également être complexe de distinguer ce qui relève du désordre ou de l'intention artistique.

Pour ces raisons, dans le domaine du patrimoine peu de travaux exploitent des méthodes purement automatiques. Notons également que les approches semi-automatiques, bien que plus intéressantes du point de vue de la sémantique présentent les mêmes limites que les approches automatiques en ce qui concerne la segmentation, puisque les limitations des algorithmes de reconnaissance de formes restent les mêmes. Dans le cas de l'art rupestre, les techniques d'annotation sémantique sont majoritairement

exploitées dans le cadre d'études de pétroglyphes nécessitant leur classification et comparaison (Papaleo et al., 2011 ; Zhu et al., 2011 ; Seidl, Breiteneder, 2011 ; Seidl et al., 2015) (Figure 47).

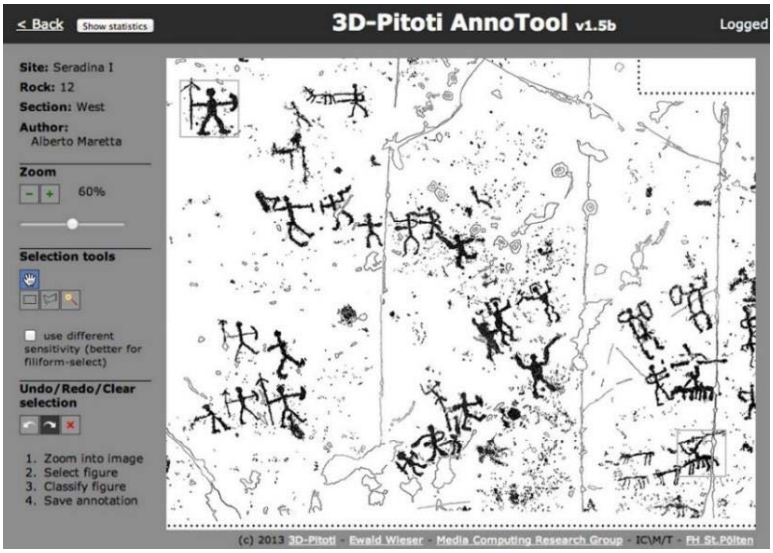


Figure 47 : Outil d'annotation proposé par Seidl et al. (2015) pour constituer le jeu de données de référence. Illustration : (Seidl et al., 2015).

Bien qu'extérieures au corpus des grottes ornées du fait de l'époque plus tardive des sujets d'étude, ces méthodes semblent s'inscrire conceptuellement dans la lignée de l'approche structuraliste d'Annette Laming-Emperaire et André Leroi-Gourhan. Dans le cas spécifique du relevé d'art pariétal paléolithique néanmoins, l'annotation sémantique d'images ne semble pas une pratique très répandue lors de relevés numériques, ou du moins, est très peu documentée. Notons cependant le travail en cours de Stoleru (2019) s'intéressant à la définition d'une ontologie de domaine dédiée à la description des traces anthropiques qui pourrait à l'avenir être intégré dans un outil dédié.

2.4.2. Approches 3D

Comme pour les approches 2D, des données 3D enrichies sémantiquement permettent d'envisager des recherches exploitant les propriétés des données mais aussi les connaissances qu'elles véhiculent. Conceptuellement ces approches sont assez semblables. Ainsi, les annotations sémantiques de données 3D peuvent être globales ou locales, et les modèles de descriptions (par tags, attributs, relations, ou ontologies) sont les mêmes. À l'instar des approches 2D, les approches 3D soulèvent donc des questions semblables concernant les points *critiques* de l'annotation, c'est-à-dire l'efficacité et la précision des outils permettant la sélection de régions, et les modèles de description à privilégier pour garantir une richesse sémantique tout en assurant l'interopérabilité des descripteurs.

Dans le domaine du patrimoine, les méthodes d'annotation 3D sont exploitées depuis plusieurs années et nécessitent souvent le passage à un modèle surfacique ou maillé de l'objet étudié, c'est par exemple le cas de Havemann et al. (2009), Manferdini et Remondino (2010), Serna et al. (2012), ou encore Yu et al. (2013), Yu et Hunter (2013). Elles s'inscrivent majoritairement dans des cadres de gestions de collections numériques, pour lesquels l'enrichissement sémantique sert efficacement les opérations de classifications et de navigation.

Bien que l'apport de l'enrichissement sémantique soit manifeste, l'utilisation de modèles maillés segmentés est problématique dans la mesure où cela tend à figer la représentation de l'objet d'étude. Or,

les acteurs de la conservation ou de la documentation du patrimoine ne pourraient se contenter d'un modèle 3D unique car les besoins en termes de précision varient selon les contextes d'études et les problématiques abordées. Différents travaux visent à surmonter ce problème, par le biais de méthodes de transfert d'annotations à différents modèles 3D d'un même objet d'étude (Scalas et al., 2017), ou par la prise en charge de nuages de points provenant d'acquisitions photogrammétriques (De Nunzio, 2018).

En ce qui concerne l'utilisation des méthodes d'annotation 3D dans le cadre de relevés d'art rupestre, nous pouvons citer en premier lieu les travaux de Barbier et al. (2017). Pour faciliter l'étude de l'art mégalithique du Néolithique, ces derniers se sont intéressés au développement d'une application d'annotation sémantique en réalité augmentée. Les auteurs proposent ainsi une application dédiée au casque Microsoft Hololens, permettant aux utilisateurs de charger un modèle 3D de l'objet d'étude et le visualiser en superposition avec leur environnement réel (Figure 76, section 3.3.5). L'utilisateur dispose d'outils spécifiquement pensés pour faciliter l'observation de gravures, notamment une source de lumière déplaçable projetant des ombres sur l'objet. La sélection de régions est manuelle et exploite les fonctionnalités de reconnaissance gestuelle du casque. Cependant, la description des annotations ne correspond qu'à un texte libre, dicté par l'utilisateur et transcrit par l'application à l'aide d'outils de reconnaissance vocale. L'application n'est pas encore complètement opérationnelle mais semble prometteuse. Notons le fait que même si toutes les interactions s'effectuent en réalité augmentée, la copie numérique de l'objet d'étude demeure décontextualisée, les auteurs ayant pris le parti d'un environnement collaboratif exclusivement *ex situ*.

Figure 48 : Relevé sur modèle 3D sur Blender avec l'outil *Grease Pencil*. Illustration : © O. Fuentes / J.Lepelé / CNP.

Enfin, en ce qui concerne l'art pariétal, Fuentes et al. (2019) détaille une chaîne de traitement permettant de tracer des observations sur un modèle 3D de son objet d'étude. Pour cela, il utilise le logiciel Blender, dans lequel il charge le modèle 3D maillé et texturé. Les observations sont tracées sur l'objet à l'aide de l'outil *grease pencil* (crayon gras) qui permet de dessiner directement sur la surface des objets de la scène (Figure 48). Les tracés sont structurés sous forme de calques qui peuvent être nommés librement par l'utilisateur. Suite à ces opérations, l'utilisateur doit transformer les tracés en éléments vectoriels sous forme de courbes de Bézier pouvant être éditées. Cette proposition constitue une évolution intéressante des outils de traitement des acteurs du relevé d'art pariétal, elle correspond d'ailleurs globalement à la transcription 3D du processus de relevé analytique technique. Cependant, bien que cette méthode puisse s'apparenter aux outils d'annotations présentés précédemment, elle ne bénéficie pas de leurs avantages en termes de sémantique, d'une part parce que la description se limite aux intitulés

de calques et de tracés et à l'éventuelle charte graphique, d'autre part parce qu'il s'agit du détournement d'un logiciel d'infographie ne permettant de bénéficier efficacement de l'information sémantique, par exemple pour établir un système de requête permettant la recherche avancée, ou pour faciliter la corrélation d'observations et d'éléments géométrico-visuels.

Finalement, l'état de l'art révèle que les initiatives en termes d'annotation 3D concernent principalement des contextes de gestions de collections muséales. Dans le cadre du relevé d'art rupestre, les expérimentations de Barbier et al. (2017) et Fuentes et al. (2019) démontrent bien l'apport du geste de l'annotation des données 3D dans la compréhension de l'objet représenté, mais elles n'exploitent pas le potentiel de descriptions sémantiques explicites et interopérables. Enfin, il apparaît clairement qu'à l'exception des travaux exploratoires de Scalas et al. (2017) qui tentent de surmonter ce problème, les méthodes d'annotation 3D se limitent à des représentations statiques non évolutives des objets d'étude, ce qui entrave considérablement les possibilités dans le cadre du relevé, car comme nous avons pu le voir précédemment, celui-ci constitue un processus cyclique nécessitant un apport informationnel continu.

2.4.3. Approches hybrides 2D/3D

Les supports 2D et 3D générés lors de relevés présentent tous des avantages au regard d'une utilisation donnée. Dans la majorité des cas, il est inenvisageable de privilégier exclusivement un support au détriment d'un autre, car les données se complètent. Si l'on s'en tenait aux méthodes d'enrichissement présentées précédemment, cela imposerait aux releveurs de répéter les opérations d'annotation pour chaque type de support, ce qui s'avèrerait fastidieux. Pour faciliter l'utilisation combinée de ces données, certains travaux s'intéressent à la définition d'approches hybrides 2D/3D pour l'enrichissement sémantique. Ces approches imposent de connaître la position relative des données (images, nuages de points et/ou modèles 3D) les unes par rapport aux autres, pour établir une relation projective permettant de transférer des annotations sémantiques réalisées sur l'un des supports aux autres éléments de la scène. Lorsqu'une relation projective est établie entre des ressources 2D et 3D, deux approches sont possibles : partir de l'espace 3D pour projeter une annotation vers les ressources 2D, ou l'inverse.

La projection 3D → 2D est exploitée par Busayarat (2010), qui considère le modèle 3D de l'objet d'étude comme un moyen privilégié d'accès aux informations patrimoniales. L'application web Nubes Imago repose sur la construction d'une représentation 3D complète de l'édifice à partir d'un relevé terrain. Le modèle est ensuite segmenté et chaque élément décrit par des attributs sémantiques. Les ressources iconographiques 2D représentant le même édifice sont ensuite référencées spatialement autour du modèle 3D. Le modèle peut alors être utilisé comme source pour stocker et propager par projection géométrique les attributs sémantiques sur l'ensemble des ressources 2D. Ce principe est exploité par De Luca et al. (2010) pour l'analyse spatio-temporelle d'édifices. Leur approche repose sur l'utilisation de différentes maquettes 3D représentant des états temporels successifs. Cependant, dans les deux cas la méthode exige un important travail de préparation et de traitement des données 3D. Ce n'est pas envisageable dans le cas de l'art pariétal, principalement parce que la segmentation des objets d'étude n'est pas rigide, elle dépend des intérêts de l'observateur et peut évoluer au cours du temps selon les connaissances de la communauté qui l'étudie.

Grilli et al. (2018) proposent une méthode 2D → 3D. Il s'agit d'une approche de classification de modèles 3D d'objets patrimoniaux basée sur la segmentation de leurs textures par des méthodes d'apprentissage supervisé. Le processus part d'un nuage de points 3D coloré ou d'un modèle 3D maillé. Dans un premier temps, des orthophotographies et des cartes UV sont créées. Un modèle supervisé est entraîné sur un ensemble d'images annotées manuellement. Les orthophotographies et textures sont segmentées et classifiées à l'aide de ce modèle, puis le résultat de cette opération est reprojeté dans l'espace objet 3D.

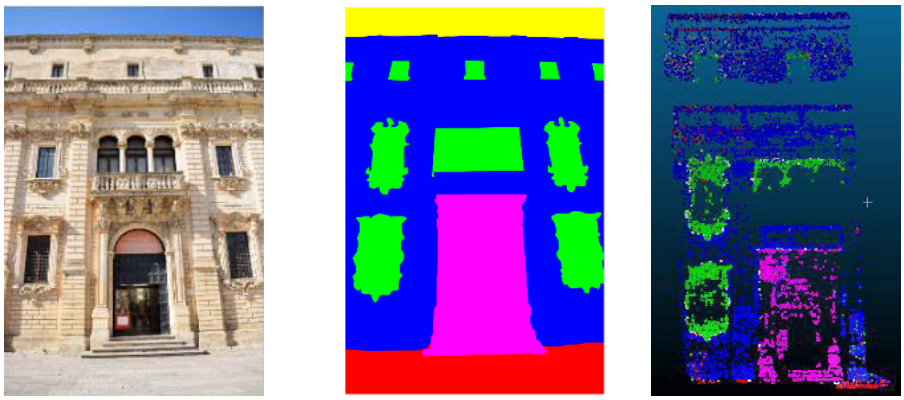


Figure 49 : Résultats expérimentaux de l'approche proposée par Stathopoulou et Remondino (2019). Gauche : Image originale. Centre : Image segmentée sémantiquement. Droite : Nuage de points épars après transfert des labels. Illustration : (Stathopoulou, Remondino, 2019).

Stathopoulou et Remondino (2019) ne considèrent pas l'enrichissement sémantique comme un post-traitement mais comme une étape du processus d'acquisition. Leur approche est d'introduire des méthodes de segmentation et classification automatique directement au sein du processus de photogrammétrie. Dans cette approche, l'annotation sémantique d'images est utilisée pour optimiser la sélection des points d'intérêt lors de l'appariement (il s'agit d'éviter les erreurs de correspondance en vérifiant que les points homologues sont bien associés au même tag). Elle est également employée pour générer des masques permettant d'exclure des zones lors de la génération du nuage de points dense (par exemple le ciel), et enfin, pour transférer les tags détectés sur les images sur le nuage de points résultant (Figure 49).

Cette approche est prometteuse, particulièrement en ce qui concerne la qualité des données 3D générées. En revanche, concernant l'annotation sémantique du nuage de points, elle semble pour l'heure insuffisante. Seules cinq classes sont utilisées pour segmenter les scènes, sans possibilité d'édition ou complétion manuelle des descriptions. Dans l'immédiat, nous pouvons imaginer que cette approche offre des perspectives intéressantes concernant les édifices architecturaux, mais nous devons émettre des réserves quant à la compatibilité avec notre sujet d'étude, l'art pariétal.

Dans sa thèse, Manuel (2016) propose également de profiter de la relation projective établie lors du processus de reconstruction photogrammétrique. Sa méthode d'indexation, a donné naissance à la plateforme *cloud* Aïoli, dédiée à l'annotation sémantique 2D et 3D pour la documentation collaborative d'objets patrimoniaux (section 3.1.7). L'idée est de mémoriser les correspondances liant les pixels des images aux points du nuage de points généré. La relation bijective ainsi établie permet de connaître, pour chaque point 3D, la liste des images sur lesquelles le point apparaît et sa position sur celles-ci. Inversement, pour chaque point d'une image, il est possible d'obtenir les points 3D correspondants.

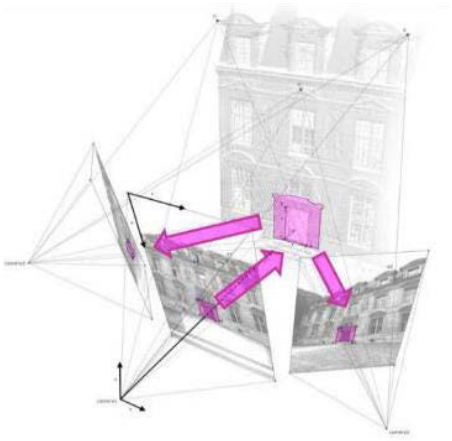


Figure 50 : Approche de propagation 2D/3D d'annotations, proposée par Manuel (2016). Illustration : (Manuel, 2016).

Grâce à ce processus, des annotations réalisées sur n'importe quelle image de l'objet d'étude peuvent être automatiquement projetées sur toutes les autres images, avec de surcroît l'avantage de gérer les occlusions (Figure 50). Actuellement, cette méthode n'est exploitée que pour l'annotation à partir des supports 2D, mais sur le plan théorique, rien ne s'oppose à ce que le nuage de points 3D soit, lui aussi, utilisé comme entrée du processus d'annotation. Les régions propagées sont structurées sous forme de calques et associées à des descripteurs géométriques (calculés automatiquement sur la base des données d'entrée) et sémantiques (définis librement par les utilisateurs). Les descriptions sémantiques peuvent être réalisées sous forme d'***attributs*** ou de ***relations***.

Par la suite, Messaoudi (2017) a montré dans sa thèse que ce fonctionnement pouvait supporter également l'utilisation d'ontologies de domaines. Menées dans le cadre d'une étude sur la conservation des structures maçonnées, les expérimentations ont permis de réaliser des annotations structurées et décrites selon une ontologie de domaine : les calques contenant les annotations correspondaient aux classes thématiques de l'ontologie, et les descriptions sémantiques étaient guidées par les concepts de l'ontologie. À partir de ces données, l'auteur a alors pu étudier les relations spatiales entre les concepts, et notamment comparer si sémantiquement, deux concepts de classes thématiques différentes se superposaient spatialement (Figure 51).

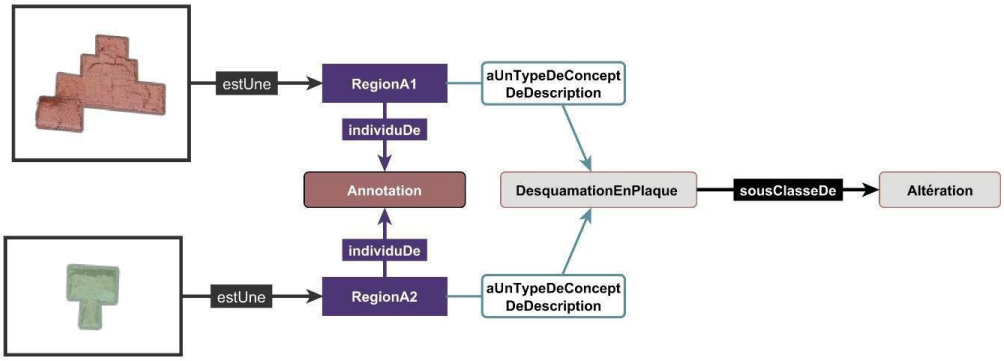


Figure 51 : Exemple issu de l'approche proposée par Messaoudi (2017), deux instances d'une annotation liées au même concept. La description basée sur une ontologie permet d'établir un modèle de corrélation multidimensionnel. Illustration : (Messaoudi, 2017).

Ces expérimentations montrent qu'il est possible de relier au sein d'un seul modèle de connaissances les dimensions sémantiques et morphologiques, ce qui rend possible l'accès à de nouveaux niveaux d'analyse et de gestion d'informations, particulièrement dans des cadres de travail collaboratifs.

Dans le cas de l'art pariétal, nous pouvons constater une certaine volonté d'aller dans le sens d'approches hybrides, à travers différents travaux s'intéressant au transfert de relevés sur les modèles 3D de l'objet d'étude. Robert et al. (2014) proposent une méthode pour mettre en relation les données archéologiques et analytiques et les données 3D, avec l'objectif d'y associer un système d'interrogation. Dans un premier temps, les auteurs se sont intéressés à la définition d'un protocole pour le recalage des relevés

analytiques 2D sur les modèles 3D correspondants, qu'ils soient issus de lasergrammétrie de photogrammétrie (Figure 52). Le recalage passe par l'identification des points de correspondances entre le relevé 2D et le modèle 3D, puis par une étape de triangulation pour optimiser la projection.

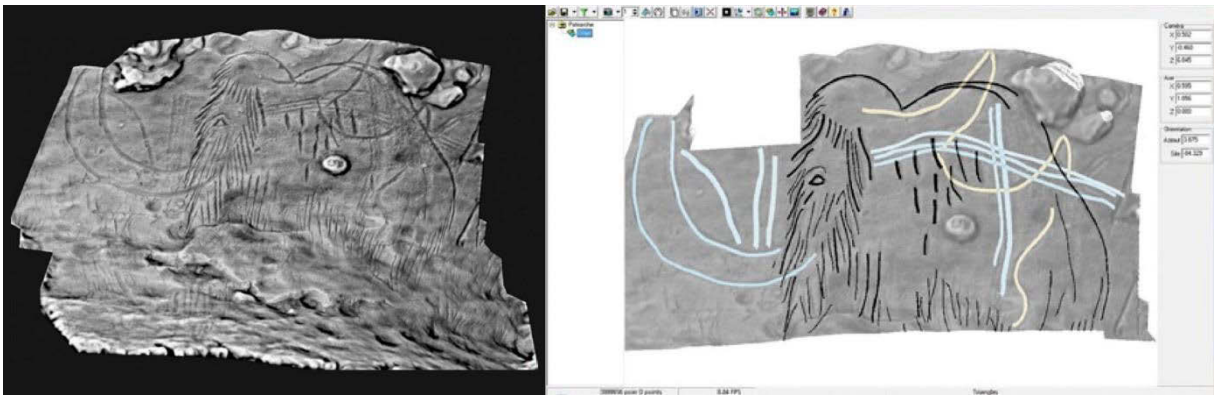


Figure 52 : Modèle 3D du mammouth du Patriarche (grotte de Rouffignac), et relevé infographique correspondant. Illustration : (Robert et al., 2014).

Pour la grotte Blanchard, les auteurs ont développé un système d'information nommé Cumulus dans le but de centraliser les données archéologiques et analytiques (relevés graphiques, fiches d'études des objets, microanalyses, etc.). Chaque ressource est référencée sur le modèle 3D de la grotte par un URI liée à un point 3D de l'objet. Les utilisateurs peuvent alors comparer la localisation spatiale des différentes données archéologiques recensées et accéder instantanément aux informations qui y sont associées (Figure 53).

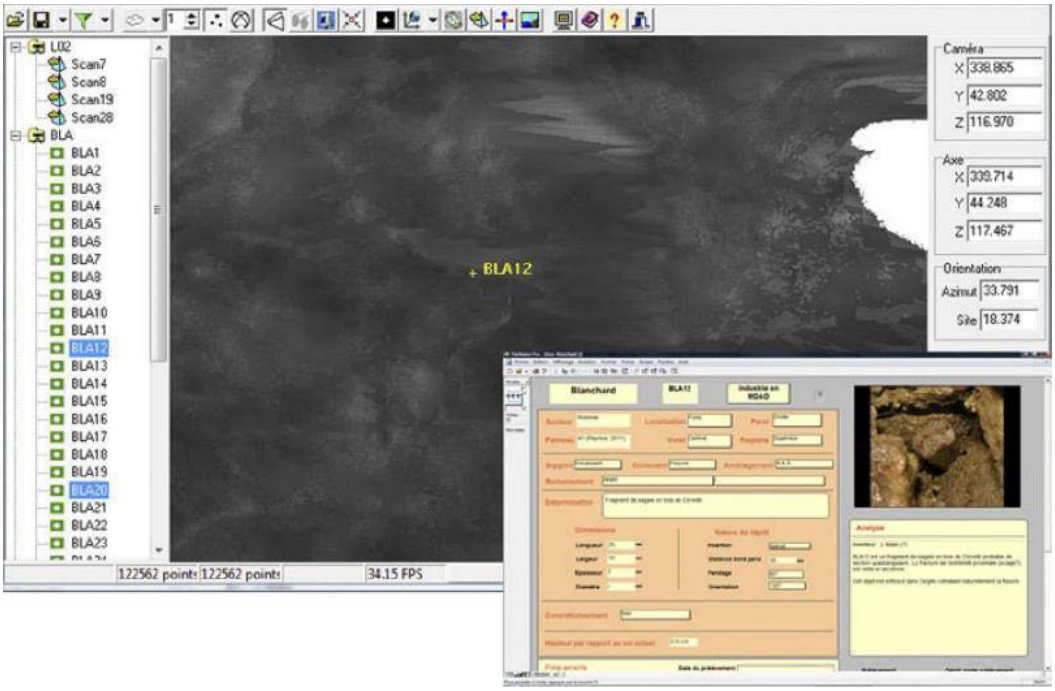


Figure 53 : Logiciel Cumulus, proposé par Robert et al. (2014) pour l'intégration et la consultation de données archéologiques et analytiques à partir du modèle 3D. Illustration : (Robert et al., 2014).

Le système est encore expérimental et les auteurs concluent sur l'intérêt d'un système d'informations archéologiques en trois dimensions, *« qui constitue à la fois un support pérenne pour la conservation du site, un support de recherche (par exemple pour des analyses spatiales), et un support d'archives 3D pour l'ensemble des données analytiques obtenues dans la grotte ».* Notons cependant que l'approche

hérite des limites des techniques de relevés analytiques, qui ne permettent pas d'établir informatiquement des relations sémantiques entre les régions identifiées. Ainsi, malgré leur projection précise sur le modèle 3D, ces données constituent une limite au système d'interrogation avancée que les auteurs entendent mettre en place.

Dans le cadre de l'étude de la grotte de Marsoulas, Fritz et al. (2016) proposent un *workflow* permettant de reprojeter les relevés réalisés en 2D sur des modèles 3D haute-résolution des panneaux concernés. La méthode proposée implique la création d'un modèle 3D par photogrammétrie, à partir duquel sont extraites les orthophotographies destinées à servir de supports aux relevés analytiques 2D. Ces images sont extraites sous différents éclairages, pour faciliter la lecture des reliefs. Le releveur retourne alors dans la grotte, charge les orthophotographies sur une tablette à l'aide du logiciel Photoshop, chaque éclairage faisant l'objet d'un calque. Face à la paroi réelle, le releveur réalise son relevé sur tablette, aidé des différents calques d'éclairage. Enfin, le document créé est renvoyé *ex situ* pour être traité et reprojeté sur le modèle 3D de la paroi. La projection des tracés sur le modèle 3D est utilisée pour étudier la façon dont la morphologie de la grotte a été exploitée comme dispositif artistique.

Ce protocole hybride 2D/3D exploite bien la capacité des données 3D à supporter les différents modes de projection géométrique, et illustre en outre une certaine volonté d'allier les pratiques de relevé *in situ* avec les capacités de traitement apportées par l'informatique. Cependant, les tracés sont projetés sous la forme de textures *bitmap*, les géométries sous-jacentes sont donc discrétisées ce qui constitue une perte d'un point de vue géométrique. Par ailleurs, il semble que les structures d'annotations établies lors du relevé *in situ* par le biais des calques ne soient pas conservées lors de la projection 2D/3D. Enfin, la charte graphique employée pour la codification des observations n'étant pas citée dans l'article, nous pouvons supposer qu'elle constitue un document annexe du modèle 3D généré à l'issue du processus, et qu'elle n'est donc pas encore, à ce stade de l'étude, exploitable pour l'interrogation des données résultantes.

2.4.4. Bilan des méthodes d'enrichissement sémantique

Dans cette section, nous avons vu que certaines avancées réalisées dans les secteurs du web sémantique commencent à s'étendre au domaine du patrimoine à travers des solutions d'annotation 2D ou 3D laissant entrevoir de nouveaux scénarios d'analyse et de corrélation, à l'interface entre la vision par ordinateur et l'analyse géométrique.

Dans le cadre du relevé, face à la dispersion des données et des connaissances, l'enrichissement sémantique apparaît comme une étape essentielle pour permettre la convergence des regards issus d'études pluridisciplinaire autour d'un même objet d'étude. Ainsi, l'annotation constitue aujourd'hui le principal défi technique et méthodologique. Cependant, si elle fait l'objet de nombreuses recherches exploratoires, l'annotation sémantique ne bénéficie pas encore de solution fonctionnelle véritablement intégrée au processus de relevé.

Dans cette optique, nous avons étudié différentes approches en matière d'annotation sémantique, dépendant des données manipulées[19]. Nous avons pu constater que, dans le cas d'objets patrimoniaux,

[19] Pour davantage d'informations concernant la taxonomie des approches existantes d'annotation appliquées aux données 3D, nous invitons le lecteur à se référer à (Ponchio, Callieri, et al., 2019).

les approches automatiques sont exclues du fait des limites des algorithmes de segmentation et de la complexité – morphologique mais aussi informationnelle – des objets étudiés. L'étude des méthodes d'acquisition et de traitement des données nous a appris dans les sections précédentes que les ressources produites lors d'un relevé présentent toutes des apports spécifiques différents, c'est pourquoi en termes d'annotation, les approches basées sur l'utilisation exclusive de ressources 2D ou 3D sont insuffisantes.

Ainsi, les approches d'annotation 2D/3D semblent constituer une piste prometteuse, car elles ne privilégient pas une catégorie de support au détriment d'une autre. Comme nous l'avons montré, l'étude de l'art pariétal exige une certaine flexibilité dans les données manipulées, dans les modalités de visualisation, et particulièrement en ce qui concerne les modes de projection géométrique. À ce titre, les approches hybrides semblent les plus à même de répondre aux besoins de ce domaine d'étude.

Parmi elles, Aïoli présente des avantages importants. Cette plateforme *cloud* est adaptée au travail collaboratif, permet aux utilisateurs de réaliser des annotations et de les propager automatiquement sur l'ensemble des ressources du projet, est dotée d'une grande flexibilité en termes de description sémantique permettant aux utilisateurs de définir librement la structure des fiches descriptives associées aux régions. Cependant, elle présente également certaines limites, dont la principale est qu'elle n'autorise pas le travail *in situ*, pourtant indispensable dans le cas de l'art pariétal.

2.5. Bilan du relevé numérique : dispersion des données et rupture réel/virtuel

Nous pouvons constater à la suite de cet état de l'art que depuis les premières démarches de relevé d'art pariétal, l'aspect individuel du relevé s'est effacé au profit d'une approche multidisciplinaire. L'étude des œuvres pariétales s'est enrichie de données de natures variées, et de méthodes spécifiques permettant leur traitement et leur analyse.

> *« Un relevé moderne n'est pas un document unique. Il comporte des points de vue différents, fait appel à des techniques complémentaires et est constitué par une série de dessins, de cartes, de photographies et de prises d'empreintes. On aura mesuré tout l'intérêt méthodologique d'une lecture et d'un enregistrement technologiquement variés des tracés préhistoriques. »* (Paillet, 2005)

Nous avons vu dans cette partie que les données 3D sont aujourd'hui couramment exploitées dans le domaine du patrimoine. Dans le cas de l'art pariétal, leur capacité à permettre l'extraction de vues selon des modes de projection géométrique définis librement constitue un atout considérable. Outre le fait qu'elles aient largement prouvé leur intérêt dans le cadre de l'étude, de la documentation, ou de la conservation d'objets patrimoniaux, leur utilisation croissante est également liée aux améliorations matérielles et logicielles qui les ont rendues de plus en plus abordables, tant d'un point de vue économique que technique. Autrefois mises en opposition, les différentes techniques d'acquisition sont aujourd'hui considérées comme complémentaires, chacune présentant des avantages.

Dans ce contexte, le relevé implique la manipulation de données hétérogènes de natures, provenances, et niveaux de détails variés. Cette multiplication des supports d'étude ouvre de nombreuses possibilités aux communautés de chercheurs en termes d'extraction d'informations et d'analyse (section 2.3, Figure 54). Cependant, comme le soulignent Thivet et al. (2019), cette richesse constitue également un paradoxe, car les données demeurent souvent sous-exploitées à défaut d'outils efficaces pour permettre leur manipulation, leur mise en relation, et leur exploitation.

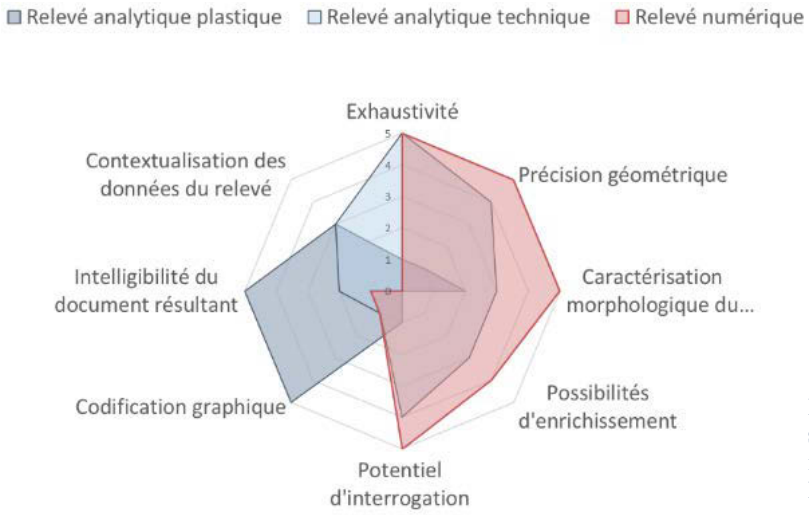


Figure 54 : Comparaison des spécificités du relevé numérique et des méthodes de relevé *classiques*. Illustration : auteur.

En termes d'extraction d'informations, la multiplication des outils logiciels permet de compenser dans une certaine mesure l'absence d'outil polyvalent dédié au relevé, mais n'est pas sans conséquence concernant l'interopérabilité des données et leur dispersion. L'essentiel des traitements d'extraction d'informations nécessite un retour en laboratoire et s'opère donc *ex situ*. Cet éloignement du site d'étude et la décontextualisation des données qu'il engendre ne semblent pas intentionnels, en témoignent les entretiens menés par Fuentes (2017) auprès des communautés de chercheurs, ou les recommandations de Le Quellec et al. (2013) et Fritz et al. (2016) concernant les *workflows* de relevé pariétal dans lesquels le phasage du travail alterne les périodes sur site et hors site. Que l'on parle de relevé analytique plastique ou technique, l'engagement de l'observateur face à la paroi étudiée est considéré comme fondamental, bien que certaines étapes du relevé puissent s'effectuer *ex situ* pour profiter des capacités computationnelles apportées par les outils informatiques.

Ce contexte cause donc d'importantes difficultés concernant la mise en relation des différentes ressources mobilisées lors d'une étude, et plus encore des différents regards des acteurs mobilisés. Nous avons pourtant pu voir dans la section 2.4 que des approches existent pour faire converger des observations pluridisciplinaires au sein des représentations numériques d'un objet d'étude. Cependant celles-ci sont non seulement encore très expérimentales – aucune n'est véritablement intégrée dans le processus de relevé – mais elles sont, elles aussi, reléguées hors site dans la mesure où les approches étudiées présentent l'enrichissement sémantique comme une étape ultime de post-traitement. **Nous pouvons donc constater une rupture franche dans la continuité du relevé numérique, entre les opérations d'acquisition de données menées *in situ* et les opérations ultérieures de traitement et d'interprétation.** Ce contexte cause une discontinuité majeure sur le plan de la connaissance qui se ressent très nettement dans le déroulement des différentes phases du relevé, qui passe d'un moment unique à un processus séquentiel linéaire.

Pour cette raison, il semble essentiel aujourd'hui d'examiner dans quelle mesure les opérations de production d'informations quantitatives (à partir de procédures d'acquisition et des traitements informatiques) et qualitatives (à partir de l'interprétation des données acquises par les acteurs du relevé) peuvent converger au sein d'un environnement numérique unique permettant aussi bien la collecte que l'analyse.

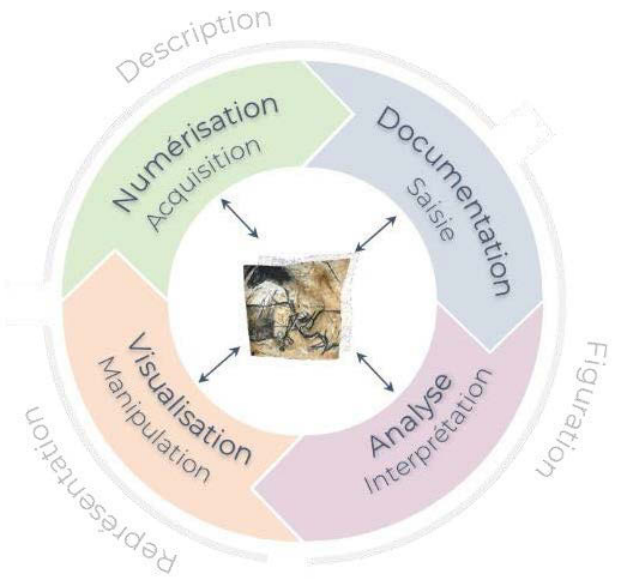


Figure 55 : Schéma de notre approche intégrée. Le relevé d'art pariétal doit être considéré comme un cycle dont les différentes phases sont centrées sur l'objet d'étude, éventuellement à travers sa copie numérique. Illustration : auteur.

Face à ce contexte, notre hypothèse principale pour la suite de ce travail est que le relevé d'art pariétal nécessite la définition d'une approche intégrée unique capable de **fédérer à la fois les données et les acteurs**. Globalement, notre approche revient à considérer le relevé d'art pariétal comme un cycle, un processus dynamique et itératif comprenant des phases de prise de connaissances sur site, d'acquisition de données, d'extraction d'informations, d'enrichissement sémantique, et de représentation (Figure 55). Cette ambition soulève deux verrous :

- Un premier verrou est celui de la formalisation d'un environnement numérique comprenant des outils d'annotation efficaces pour permettre aux chercheurs de comprendre, s'approprier, et enrichir les données numériques, ou autrement dit, de faire converger les aspects visuels, géométriques et sémantiques au sein d'un environnement de travail unique. Nous utiliserons pour cela comme levier des travaux existants dans les domaines des systèmes d'information patrimoniaux et de l'annotation sémantique 2D/3D.
- Un second verrou est celui de l'**hybridation** afin de **réconcilier les moments *in situ* et *ex situ***, c'est-à-dire de surmonter les ruptures existantes entre le réel et le virtuel. Pour cela, nous utiliserons comme levier la réalité mixte, dont la raison d'être est justement *« de permettre à une personne de réaliser des activités sensori-motrices et cognitives dans un nouvel espace en associant l'environnement réel et un environnement virtuel »* (Fuchs et al., 2010).

Ces deux verrous nous amènent donc à nous intéresser respectivement à deux domaines : **les systèmes d'information patrimoniaux**, et **la réalité mixte**. Nous étudierons dans un premier temps les approches existantes concernant les systèmes d'information patrimoniaux, afin de mieux comprendre les enjeux de la définition d'une approche intégrée dédiée au relevé d'art pariétal. En particulier, nous verrons que certains travaux peuvent servir de base de départ à notre système d'information.

Nous verrons dans un second temps que l'hybridation réel/virtuel est la raison d'être de la réalité mixte et étudierons les technologies existantes, ainsi que les méthodes de recalage permettant aux utilisateurs d'interagir avec des données virtuelles contextualisées dans leur environnement réel. À l'issue de cet état de l'art, nous définirons précisément le cahier des charges de notre système d'information et opèrerons des choix concernant notre protocole expérimental.

Chapitre 3. Les systèmes d'information et l'hybridation réel/virtuel dans le domaine du patrimoine

3.1. Systèmes d'information patrimoniaux

Différents travaux s'intéressent à l'élaboration de systèmes d'information patrimoniaux compatibles avec l'exploitation de données numériques tridimensionnelles. Peu d'entre eux s'intéressent spécifiquement à l'art pariétal comme ArTaPOC ou certaines approches basées sur l'utilisation d'un système d'information géographique (SIG). D'autres, comme CHER-OB, AGATA, NUBES ou Aïoli, se veulent plus génériques et peuvent théoriquement être exploités pour tout type d'objet patrimonial.

3.1.1. ArTaPOC

L'objectif du projet ArTaPOC (**Ar**chéologie et **Ta**phonomie de la **P**aroi **O**rnée à **C**ussac), consiste au développement d'une application dédiée à l'étude intégrée pluridisciplinaire de l'art pariétal. L'ambition est de proposer un outil de relevé 3D collaboratif, permettant centraliser les observations des pariétalistes et des géoarchéologues. Selon Feruglio et al. (2013), cet outil doit permettre au minimum de disposer des mêmes fonctionnalités qu'un relevé 2D, c'est-à-dire la sélection de zones sur les parois, la possibilité de choisir la couleur des tracés et de leur associer un texte explicatif, l'individualisation des sélections afin de permettre leur hiérarchisation, et la possibilité de désigner un point précis pour lui attacher des informations. Le choix de procéder au relevé sur un modèle 3D repose sur la nécessité de faciliter la lecture par le biais des éclairages ou des outils d'accentuation de gravures, mais aussi de préserver les propriétés métriques de l'objet, que l'on perd en photographie.

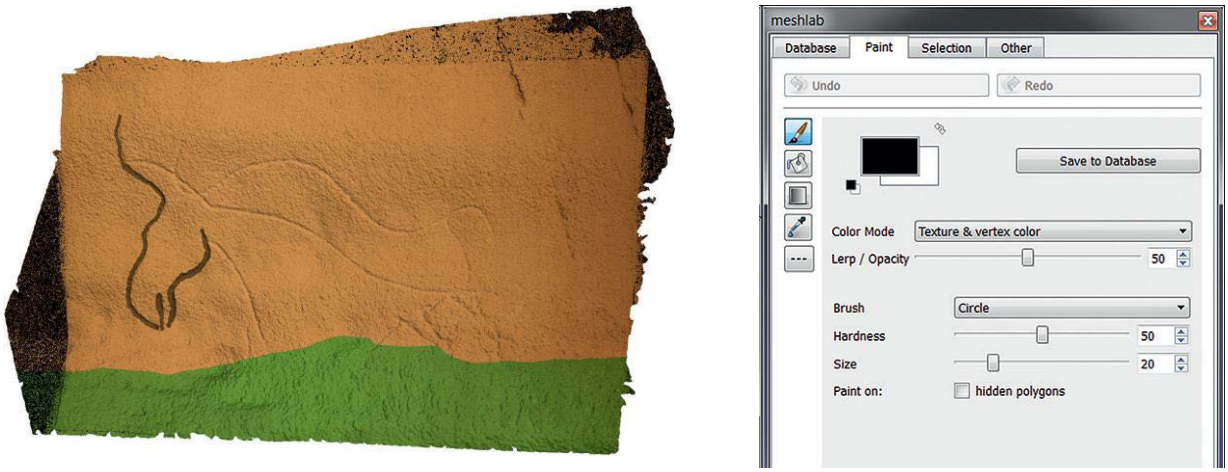


Figure 56 : Relevé de tracés gravés (grotte de Cussac) à l'aide du *plugin* Meshlab. Illustration : (Feruglio et al., 2013).

Pour cela, l'équipe a opté pour la programmation d'un *plugin* pour le logiciel MeshLab afin de profiter de ses fonctionnalités préexistantes. Il est associé à une base de données en ligne permettant de stocker les éléments produits par les utilisateurs. Le *plugin* dispose de différents outils pour permettre la sélection de régions (par polygone, pinceau, ou point 3D individuel), ensuite nommées et décrites par l'utilisateur (Figure 56). Un site web permet aux différents acteurs du projet de télécharger le maillage et la dernière version de la base de données pour l'utiliser dans MeshLab et reprendre le travail de relevé, mais aussi de téléverser ces mêmes éléments actualisés à l'issue d'une session. Ce fonctionnement permet aux utilisateurs de disposer toujours de la dernière version des relevés effectués sur une paroi.

Pour les auteurs, la preuve de concept développée lors de ce projet est probante, et fonctionnelle malgré des éléments d'ergonomie à reconsidérer. Malheureusement, aucune publication ne fait état des suites de ce projet, nous ne savons pas à ce jour si la preuve de concept a évolué ni si elle est couramment utilisée. À l'image d'autres initiatives (Robert et al., 2014 ; Fritz et al., 2016 ; Fuentes et al., 2019), ce projet démontre bien l'envie des acteurs du domaine de l'art pariétal de s'orienter vers des approches intégrées spécifiquement dédiées au relevé, et capables de bénéficier de l'apport des ressources 3D.

Cependant, l'architecture applicative choisie pour ce projet présente des limites notables. D'une part, elle impose l'installation d'un logiciel et d'un *plugin* pour tous les acteurs amenés à travailler sur les données, d'autre part l'organisation choisie pour garantir l'actualisation de la base de données en ligne n'autorise pas de simultanéité dans le travail pluridisciplinaire. En outre, les outils d'annotation nécessitent l'utilisation d'un modèle maillé préétabli ce qui limite fortement les perspectives, par exemple en termes de suivi temporel. Enfin, les données, bien qu'à l'échelle, sont visualisées de manière isolée, et ne sont liées à leur contexte et au reste de la collection que par le système de métadonnées.

3.1.2. Approches basées sur l'utilisation de SIG

D'après Bordin (2002), les SIG, ou Systèmes d'Informations Géographiques sont *« des cas particuliers de bases de données contenant, outre des renseignements quantitatifs et qualitatifs, des informations sus la localisation spatiale de chaque observation »*. Plus précisément, il s'agit de systèmes de gestion de bases de données permettant le stockage, l'extraction, l'interrogation, l'analyse et l'affichage de données géolocalisées. Ils comprennent deux éléments principaux (CCIG, 2012) : des données permettant de connaître la forme et la localisation dans l'espace des objets étudiés (couches géographiques), et des données alphanumériques permettant de décrire les objets géographiques, d'en connaître des caractéristiques qualitatives et/ou quantitatives (données attributaires).

Outre la possibilité de représenter visuellement de l'information quantitative, les SIG sont particulièrement intéressants par leur capacité à permettre l'étude des relations spatiales pouvant exister entre des informations voisines, afin de déduire de nouvelles connaissances. De fait, conceptuellement les SIG présentent des intérêts proches des problématiques soulevées par le relevé patrimonial.

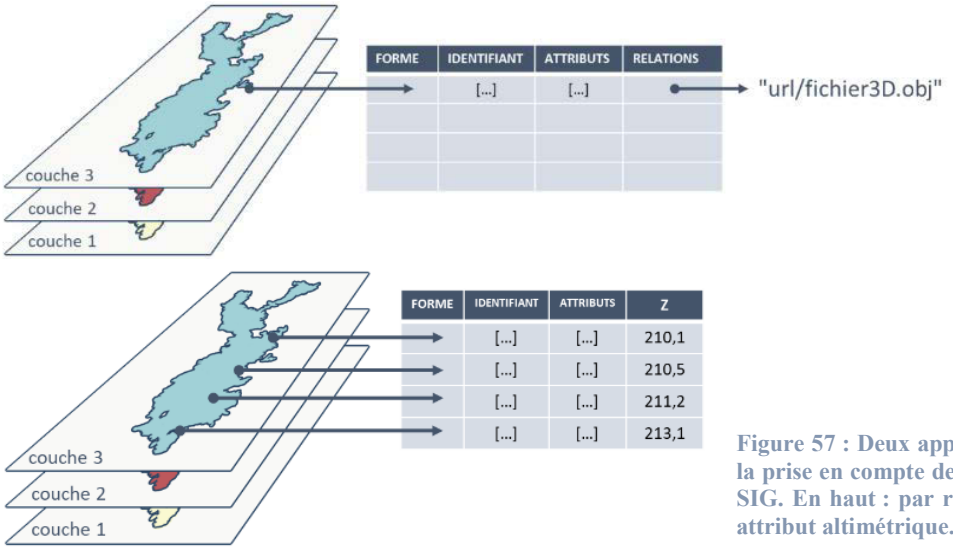


Figure 57 : Deux approches possibles pour la prise en compte de données 3D dans des SIG. En haut : par relations. En bas : par attribut altimétrique. Illustration : auteur.

Cependant, s'ils sont efficaces sur des données de types cartographiques 2D ou 2,5D, les SIG sont encore globalement peu adaptés aux données 3D. Celles-ci peuvent être traitées comme des relations liées à des zones spatiales spécifiques, dans ce cas elles ne bénéficient pas de l'ensemble des outils d'analyse disponibles dans le sens où elles n'agissent pas comme *support* d'informations, ou bien elles peuvent être traitées comme des données quantitatives, c'est-à-dire des valeurs d'élévation *Z* pour tout point de la zone concernée, auquel cas elles ne permettent pas de s'adapter à la complexité morphologique des sites tels que les grottes ornées (Figure 57). Étant donné les atouts non négligeables des outils de corrélation spatiale d'informations hétérogènes, certains travaux s'intéressent à leurs potentielles applications aux domaines du patrimoine et de l'archéologie, et particulièrement à leur adaptation aux données 3D. On parle alors de SIG 3D.

Dans le cadre de l'étude de la grotte Chauvet-Pont d'Arc, Ployon et al. (2010) se sont intéressés à la mise en place d'un SIG 3D pour centraliser les contributions des différents acteurs. La grotte comprend des représentations pariétales riches, des vestiges et traces de différentes natures, mais aussi une importante diversité géomorphologique. Face à la quantité de données produites par chacune des disciplines impliquées, l'équipe a été confrontée à la difficulté de proposer un environnement unique efficace pour les centraliser et permettre le recoupement des observations. Ce contexte les a amenés à proposer un système d'information. Parmi les pistes possibles, le SIG a été choisi pour permettre de traiter des problématiques spécifiques à chaque discipline et les confronter par l'entrée spatiale.

Pour aborder le problème de la gestion des données 3D, les auteurs ont testé deux approches. L'une consiste à nettoyer et découper le nuage de points de la grotte de façon à le diviser en plusieurs modèles 2,5D. Ces données sont ensuite importées dans le SIG en tant que MNT[20], et sont alors prises en charge naturellement au même titre que n'importe quelle couche géographique. Cependant, cette approche impose un prétraitement conséquent et s'adapte difficilement aux éléments complexes, par exemple les stalactites. La seconde approche a été de considérer la donnée 3D comme une ressource externe, liée au SIG par un lien hypertexte. Ce fonctionnement permet d'accéder à davantage d'outils de manipulation 3D mais a tendance à désolidariser les traitements de l'objet 3D du reste du système d'information. Finalement, les auteurs soulignent le caractère fédérateur du SIG et l'apport du croisement des observations selon des critères spatiaux, mais l'intégration des données 3D complexes demeure problématique et les approches proposées ne constituent que des réponses partielles.

Plus éloignés de l'art pariétal mais toujours dans le domaine du patrimoine, les travaux de Campanaro et al. (2016) s'intéressent à la mise en place d'un système d'information 3D pour la gestion des objets patrimoniaux, exploitant simultanément les atouts de la visualisation 3D et ceux des SIG. Initiée lors du projet suédois *Pompeii*, cette expérimentation vise à fournir une chaîne de traitement efficace aux spécialistes de la conservation et de la restauration. Pour cela, les auteurs proposent de décomposer l'objet d'étude en parties approximativement planes – dans leur cas, les façades. À partir de chacun de ces plans de référence sont fabriqués des clones horizontaux, utilisés comme couches thématiques pour le SIG. Ces couches servent de bases aux analyses statistiques, notamment pour l'évaluation des risques de dommages. Les données produites à partir des clones sont ensuite retransférées sur les éléments de façades de l'objet 3D (Figure 58). La méthode proposée offre la possibilité de mettre à jour les attributs de ces éléments en temps réel, car les faces des maillages utilisées (façades et clones) conservent leurs

[20] Modèle Numérique de Terrain

indices. Cette approche semble intéressante, particulièrement pour profiter des outils d'analyse à partir des attributs de couches. Cependant, le passage imposé à des éléments plans la rend inadaptée aux géométries complexes telles que les grottes.

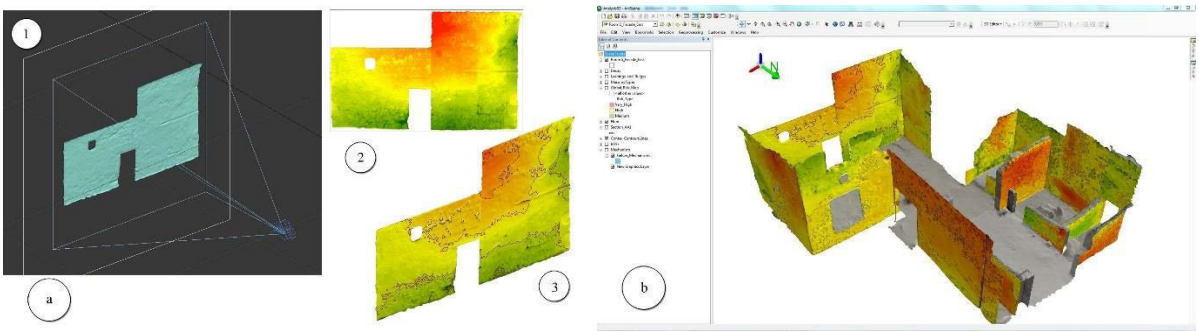


Figure 58 : Principe d'affectation de la carte d'élévation comme texture du modèle 3D de la façade et exploitation dans un SIG. Illustration : (Campanaro et al., 2016).

Nous pouvons constater que malgré l'intérêt notable des SIG en termes de mise en relation de ressources spatialisées et de corrélation, l'insuffisante intégration des données 3D nécessite le passage par une multitude de procédés empiriques, ce qui constitue un obstacle majeur pour le relevé.

3.1.3. CHER-Ob

CHER-Ob (Shi et al., 2016 ; Wang et al., 2018) est un logiciel né du constat que la réussite d'un projet de recherche sur le patrimoine culturel dépend fortement des interactions entre les experts des différents domaines. Les approches actuelles sont confrontées à trois verrous majeurs : la difficulté à centraliser les données issues de recherches différentes concernant le même objet d'étude, le manque de méthodes d'enrichissement sémantique et d'analyse, et le manque d'outils permettant de partager les résultats d'analyse de manière formelle.

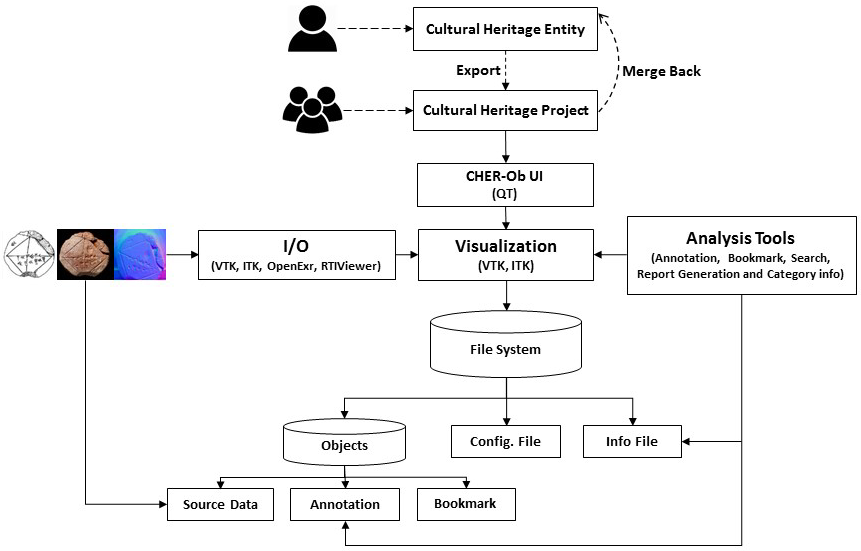


Figure 59 : Architecture logicielle de CHER-Ob. Illustration : (Shi et al., 2016).

Pour surmonter ces problèmes, les auteurs proposent CHER-Ob, un logiciel destiné à la recherche sur le patrimoine culturel et disposant d'outils d'analyse et d'évaluation. Il introduit la notion de *Cultural Heritage Entity* ou *Objet Patrimonial* (CHE), qui constitue un conteneur d'informations concernant une

entité patrimoniale. Il s'agit d'une méthode de stockage hiérarchique qui permet d'associer les projets se rapportant au même objet (Figure 59). Dans le cas de l'art pariétal par exemple, un CHE pourrait être la grotte Chauvet-Pont d'Arc, et comprendrait ainsi tous les sous-projets se rapportant aux différentes zones (galerie du Mégacéros, salle des Brunel-Deschamps, etc.) ou aux différents panneaux (panneau de la Panthère, panneau des Chevaux, etc.).

Le logiciel prend en charge plusieurs types de données (2D, 3D et RTI) et propose différents outils d'annotation. Pour les données 2D, les outils de sélection permettent des annotations ponctuelles ou sous forme de régions. Pour les données 3D, les outils de sélection permettent de désigner des sommets, des surfaces, ou des volumes. Les descriptions sont réalisées sous forme de textes libres.

Le lien hiérarchique établi entre les projets et les objets patrimoniaux permet de recouper les données et éventuellement de mettre en relation les équipes différentes s'intéressant au même objet d'étude. La prise en charge de plusieurs formats de données est également un aspect important, qui permet de ne pas renoncer aux apports des différentes ressources. Cependant, le système proposé pour l'annotation ne permet pas de gérer finement la description sémantique des observations réalisées. Pourtant, quelle que soit l'approche choisie pour le relevé, nous avons pu voir précédemment que les processus d'interprétation et de documentation nécessitent de structurer, hiérarchiser les observations. En outre, leur description sous forme de texte libre ne permet pas de bénéficier d'outils de recherche ou de corrélation efficaces. Enfin, le système proposé ne permet pas aux utilisateurs d'annoter simultanément le même CHE, ce qui constitue une limite en termes de collaboration.

3.1.4. Agata

Le logiciel Agata (Soler et al., 2017) est un système d'informations patrimoniales fournissant aux spécialistes un environnement pour l'annotation sémantique de modèles 3D haute-résolution. Ce travail part du même constat initial que pour CHER-Ob. Le système permet de charger des modèles maillés texturés, et de les annoter en réalisant des sélections directement sur la surface des objets 3D (Figure 60). Les annotations ainsi créées sont stockées dans des calques, chacun étant associé à un schéma XML spécifiant la structure des descriptions sémantiques. Les descriptions sont donc réalisées sous forme d'*attributs* qui peuvent être utilisés pour effectuer des requêtes, par exemple *« Quelles zones ont été restaurées le 24 mars 2017 ? »*. Enfin, le logiciel propose un outil pour estimer automatiquement les zones exposées aux dépôts de poussières à partir des orientations des faces du maillage et des occlusions.

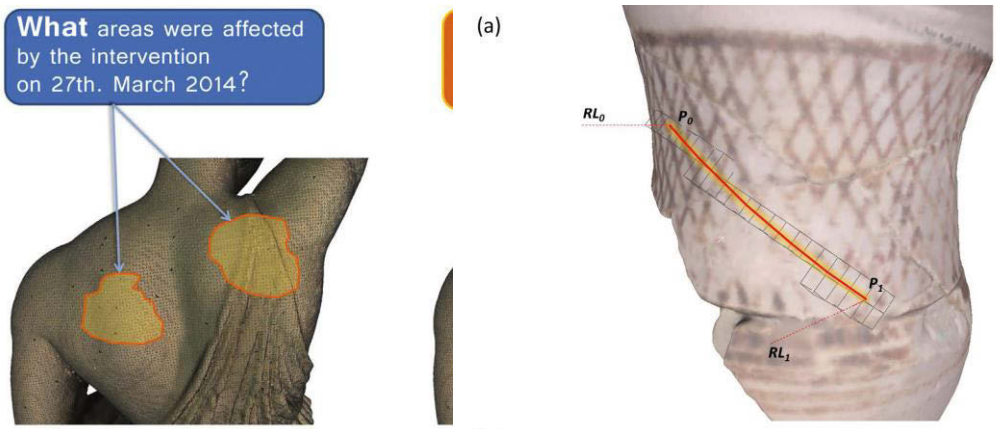


Figure 60 : À gauche : exemple de requête qu'Agata peut traiter. À droite : principe d'annotation proposé. Illustration : (Soler et al., 2017).

Cette méthode ne prend pas en charge des données autres que des maillages 3D, ce qui ne permet pas de centraliser les ressources relatives à un projet, ni de transférer les annotations réalisées à d'autres représentations du même objet. Par ailleurs elle impose un traitement préalable des données 3D qui doivent être maillées et texturées en amont, alors que les démarches d'analyse ou de documentation sont souvent évolutives et s'inscrivent dans la durée. Ainsi, dans le cadre d'un suivi temporel, chaque lot de données doit faire l'objet d'un projet spécifique. Par ailleurs, l'aspect collaboratif n'est pas évoqué.

3.1.5. Netptune IS

Dans le cadre de la restauration de la Fontaine de Neptune à Bologne (Italie), Apollonio et al. (2018) se sont intéressés à la définition d'un système d'information centré sur une numérisation 3D de la fontaine pour de stocker les données de restauration et y accéder efficacement. Contrairement aux approches évoquées précédemment, ce système se présente comme une application web. Cela présente plusieurs avantages, en particulier sur le plan collaboratif puisque cela permet un accès simultané à plusieurs utilisateurs, évite l'installation d'applications locales et assure la centralisation des données issues des sessions d'utilisation. Neptune IS a été réalisée sur la base de 3DHop, un *web-viewer* 3D développé par la même équipe (Potenziani et al., 2015) basé sur l'API WebGL (Khronos Group, 2019).

Le système proposé comprend un espace de visualisation 3D, des fonctionnalités d'annotation pour l'ajout d'informations sur le modèle 3D, ainsi que des outils de recherche et de navigation. Pour cela, les données 3D ont été préalablement traitées, maillées, et présegmentées. La Figure 61 présente la structure hiérarchique définie. Cette structure sémantique, fondée sur la façon dont les restaurateurs organisent leur travail et leur documentation, est utilisée pour la navigation dans la scène mais aussi pour structurer les annotations réalisées par les utilisateurs.

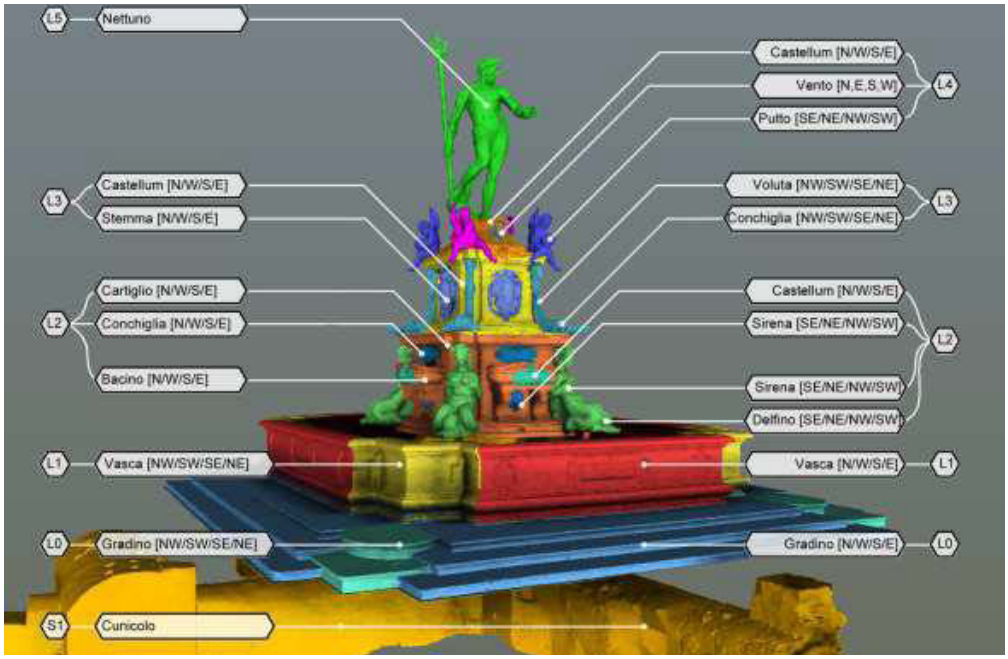


Figure 61 : Structure hiérarchique définie pour la Fontaine de Neptune et segmentation subséquente. Illustration : (Apollonio et al., 2018).

Les annotations peuvent être réalisées sous forme de points, polylignes ou surfaces. Leur description est réalisée à l'aide de fiches d'opérations, comprenant l'auteur, la date, l'élément concerné parmi la

segmentation préétablie, le type d'activité, ainsi qu'un texte libre, des *tags*, et d'éventuels médias additionnels (images ou documents PDF). Des attributs géométriques (aire et périmètre) sont automatiquement calculés et ajoutés à ces descriptions.

Cependant, l'application présente plusieurs limites. Elle ne prend en charge que les données 3D maillées, et ne permet donc pas de bénéficier de l'apport des données présenté en section 2.2. En outre, le système proposé impose une segmentation préalable. Par ailleurs, comme nous l'avons vu dans la section 2.4 les descriptions sémantiques des annotations sous forme de *tags* ou de textes libres ne décrivent qu'implicitement les propriétés des éléments sélectionnés et sont donc sujettes à l'ambiguïté dans l'interprétation. Elles semblent insuffisantes au regard de la complexité des processus d'analyse et de documentation d'un objet patrimonial. Enfin, la plateforme présentée est un prototype spécifiquement conçu pour le cas d'étude de la Fontaine de Neptune. Les auteurs soulignent leur intention de la faire évoluer vers une plateforme générique, elle pourra peut-être s'enrichir à cette occasion d'outils ou modifications nécessaires pour l'adapter aux besoins du relevé.

3.1.6. NUBES

NUBES (Busarayat, 2010 ; Stefani, 2010) est un projet débuté au laboratoire MAP il y a environ vingt ans. L'ambition du projet était d'examiner comment les informations quantitatives (extraites d'opérations d'acquisition et de mesures) et qualitatives (produites par l'interprétation des données acquises) pouvaient converger pour être analysées et affichées au sein d'une plateforme intégrée unique. NUBES se présente comme une plateforme web pour l'analyse pluridisciplinaire de bâtiments patrimoniaux. Le système repose sur l'utilisation d'une maquette numérique préétablie comme support d'informations. Cette approche impose donc l'élaboration d'un modèle 3D, segmenté en éléments caractéristiques qualifiés par des attributs sémantiques provenant notamment de thésaurus.

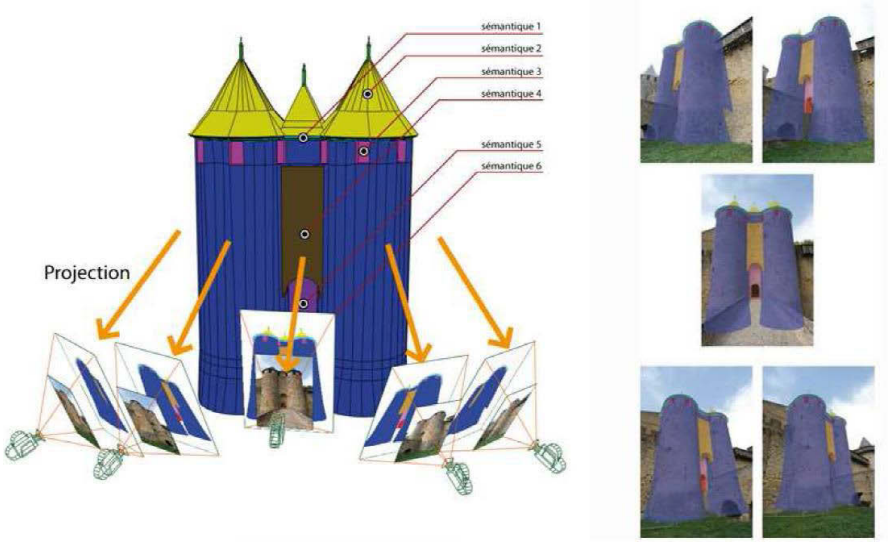


Figure 62 : Le référencement spatial d'images autour d'une maquette numérique segmentée leur permet de recevoir des annotations par projection 3D → 2D. Illustration : (Busarayat, 2010).

Les relations entre attributs sémantiques et géométriques sont exploitées pour réaliser des recherches avancées et filtrer les éléments affichés. Les ressources iconographiques 2D (dessins, peintures, cartes postales, photographies, etc.) peuvent être référencées spatialement autour de la maquette 3D à l'aide

de différents outils. Les relations projectives entre les images orientées et le modèle 3D sont utilisées pour reprojeter sur chaque source 2D des calques comprenant les informations sémantiques provenant de la maquette (Figure 62). De Luca et al. (2010) ont complété ce système par la prise en charge d'informations temporelles, pour permettre le suivi de l'évolution de l'objet d'étude. Le principe repose sur l'utilisation de modèles 3D représentant l'objet d'étude à différents états temporels.

Ce système d'information montre comment les données 3D peuvent permettre de centraliser les ressources concernant un édifice. Ce principe permet de mettre en relation des attributs géométriques, sémantiques, et temporels, afin de mieux comprendre l'objet d'étude. Cependant, l'utilisation d'un modèle 3D maillé et présegmenté est contraignante. Par ailleurs, les données 2D orientées autour du modèle 3D ne servent qu'à recevoir les informations sémantiques provenant du modèle, mais ne permettent pas de réaliser des annotations.

3.1.7. Aïoli

Aïoli (De Luca et al., 2017) est un système d'information patrimonial développé au laboratoire MAP, faisant suite à NUBES. Cette approche propose de considérer l'objet réel – à travers sa copie numérique – comme le seul dénominateur commun capable de fédérer les ressources et les connaissances des communautés qui l'étudient. Concrètement, Aïoli se présente comme une plateforme *cloud* dédiée à l'annotation sémantique 2D et 3D pour la documentation collaborative d'objets patrimoniaux, réalisée sur la base des travaux de Manuel et al. (2014) concernant la propagation automatique d'annotations sémantiques. Contrairement aux travaux précédemment évoqués, Aïoli intègre un outil de reconstruction photogrammétrique et n'impose pas de manipulation préalable des données.

Accessible à partir d'un navigateur depuis un PC, une tablette ou un *smartphone*, elle permet aux utilisateurs de réaliser à partir de simples photographies, des annotations spatialisées en 3D autour d'un objet d'étude (qu'il s'agisse d'un édifice, d'une sculpture, d'une peinture, d'un objet d'art ou de fragments archéologiques), et d'en faire bénéficier leurs pairs. La plateforme comprend :

- **TACO, un procédé de spatialisation 3D** incrémentielle pour gérer la fusion géométrique de plusieurs images provenant d'acteurs différents à des états temporels différents ;
- **Un *viewer* 2D et un *viewer* 3D** permettant aux utilisateurs de naviguer dans leurs jeux de photographies et de visualiser ces mêmes images orientées autour d'une représentation 3D dynamique ;
- **Un cadre d'annotation 2D** permettant aux utilisateurs de tracer, visualiser et enregistrer les régions pertinentes (points, polylignes, surfaces) en manipulant de simples images 2D ;
- **Un modèle de structuration de données par calques** pour décrire les objets dans toute leur complexité géométrique et sémantique, selon les observations réalisées.

La plateforme repose sur un moteur de corrélation multidimensionnelle, basé sur l'automatisation des processus de spatialisation d'images par photogrammétrie et sur la possibilité de réunir, traiter et diffuser de grandes quantités de données via le *cloud computing*. Les photographies envoyées par chaque acteur lors de la création d'un projet sont exploitées pour générer un nuage de points dense de l'objet étudié.

En complétant ce processus par une phase d'indexation, Aïoli établit une correspondance entre chaque point du nuage de points et chaque pixel des différentes images (Figure 63).

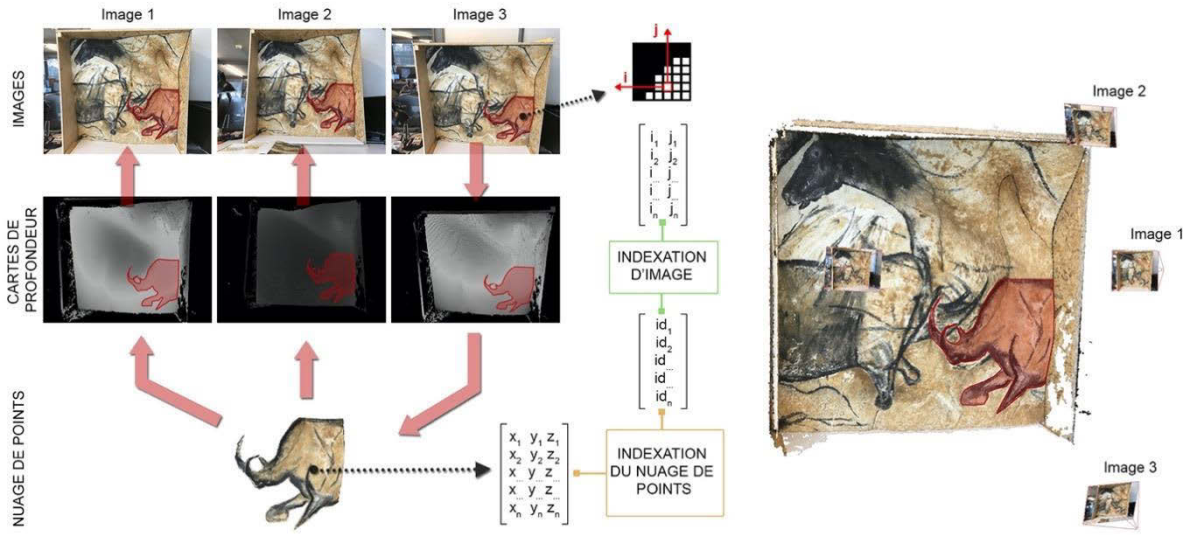


Figure 63 : Principe de l'indexation exploitée par Aïoli pour propager des annotations par la relation 2D → 3D → 2D. Illustration : auteur, d'après (Manuel, 2016).

La relation ainsi établie permet de connaître, pour chaque point 3D, la liste des images sur lesquelles le point apparaît. Inversement, partant d'un ou plusieurs points sur une image, Aïoli peut obtenir la liste des points 3D concernés. Grâce à ce processus, les annotations réalisées sur n'importe quelle image de l'objet sont automatiquement projetées sur toutes les autres images (passées, présentes et futures), mais aussi continuellement corrélées sur le plan géométrique, visuel et sémantique, avec d'autres annotations ayant une localisation spatiale proche.

Les annotations sont structurées sous forme de régions appartenant à des calques, associées à des descripteurs géométriques (calculés automatiquement) et sémantiques (définis par les différents utilisateurs). Enfin, en plus des champs de description par *attributs*, les utilisateurs peuvent exploiter des champs de description par *relations* pour associer des ressources multimédia complémentaires (Figure 64). Des preuves de concept (Messaoudi et al., 2018) montrent que l'approche pourrait permettre d'intégrer également des vocabulaires contrôlés et des ontologies pour la description des annotations. La plateforme offre aux utilisateurs la possibilité de partager les projets de façon à rendre le processus d'analyse collaboratif, et offre pour cela des outils permettant une gestion fine des droits d'accès aux calques (lecture/écriture)[21].

Ce système d'information constitue une avancée importante concernant la création d'un cadre pour l'intégration des connaissances humaines au sein de représentations numériques 3D d'objets patrimoniaux. Le nuage de points, placé au centre de l'écosystème numérique, est utilisé pour fédérer les connaissances autour de l'objet, en limitant ainsi l'effet de dispersion habituellement constaté dans le domaine du patrimoine. L'approche intégrée proposée permet d'envisager le relevé comme un processus cyclique dynamique, dans lequel les données peuvent être complétées de manière itérative au gré des observations, et continuellement enrichies sémantiquement. La méthode de propagation

[21] Cette fonctionnalité est en cours d'implémentation.

proposée permet de diffuser les annotations sur l'ensemble des ressources concernées, et épargne ainsi aux observateurs un travail fastidieux d'annotation manuelle.

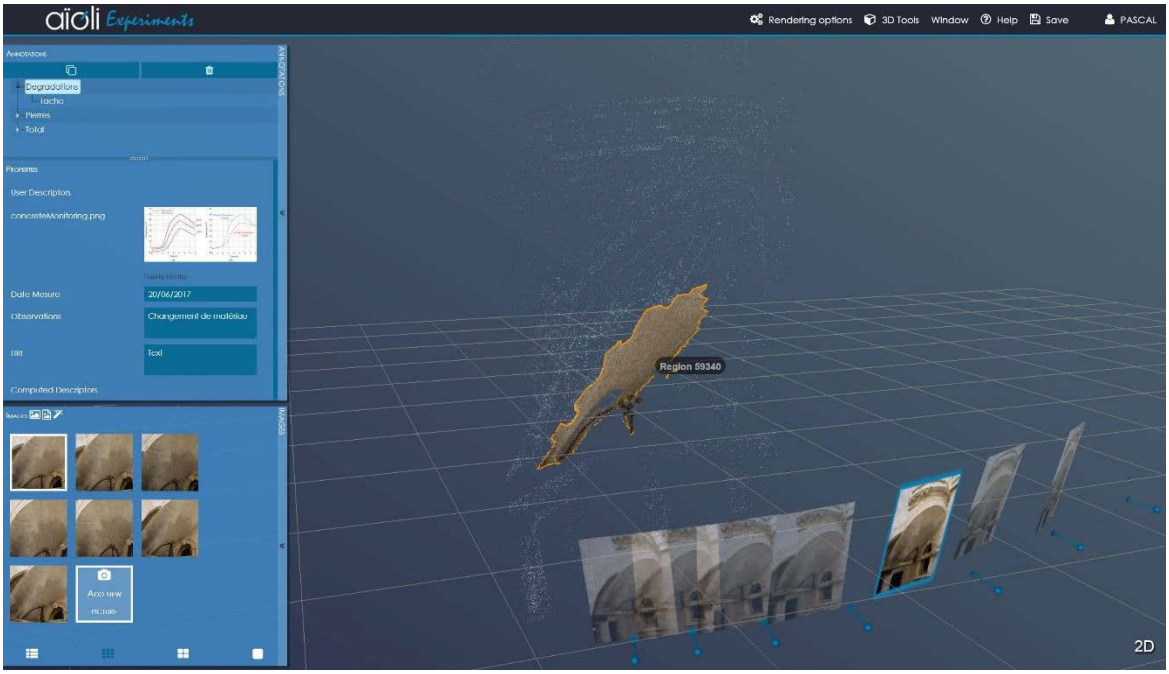


Figure 64 : Interface d'Aïoli, le *viewer* 3D permet de visualiser les images orientées autour d'un nuage de points annoté.

Aïoli présente cependant certaines limites. Chaque annotation ne peut être réalisée qu'à partir d'une seule photographie, ce qui peut se révéler insuffisant en cas de formes complexes. Les nuages de points ne sont pas exploités comme support d'annotation, alors qu'ils pourraient s'avérer plus efficaces notamment dans le cas d'objets tortueux sujets aux phénomènes d'occlusion. Par ailleurs, la décomposition de l'interface en deux *viewers* distincts, l'un consacré aux données 2D, l'autre aux données 3D, prive les utilisateurs d'une certaine simultanéité dans les procédés d'annotation, et impose des temps de calcul conséquents pour la propagation. Nous avons vu précédemment que ces ressources sont complémentaires et que leur utilisation combinée facilite l'observation de certains stigmates sur les parois. Enfin, la fenêtre de visualisation ne permet de représenter les données que de manière isolée, c'est-à-dire que la scène comprend les entités dont le positionnement relatif est cohérent, mais l'ensemble est décontextualisé.

3.2. Bilan sur les systèmes d'information patrimoniaux

Nous avons vu précédemment que les supports 2D et 3D présentent des intérêts complémentaires au regard du relevé d'art pariétal. Ainsi, l'utilisation combinée de ces deux catégories de ressources constitue une attente importante. Pourtant, suite à cet état de l'art sur les systèmes d'information patrimoniaux, nous pouvons constater que peu d'entre eux permettent de les mettre en relation et de profiter de leurs apports respectifs.

Parmi eux, Aïoli présente un atout considérable en permettant la propagation automatique des annotations à l'ensemble des ressources d'un projet. En outre, à l'exception d'Aïoli, toutes les approches proposées imposent un traitement préalable des données 3D, voire même une segmentation irrévocable. Pourtant, l'intégration des processus d'acquisition et de numérisation permet de proposer une approche

véritablement intégrée, considérant davantage le relevé comme un processus cyclique itératif que comme une étape informatisée d'un traitement linéaire. Ce fonctionnement permet non seulement d'assurer une certaine traçabilité dans les données, mais aussi de les faire évoluer en les complétant au cours du temps pour assurer un suivi continu de l'objet d'étude.

Au regard des enjeux du relevé, Aïoli semble donc l'approche la plus aboutie, c'est pourquoi nous proposons de nous inscrire dans sa continuité pour la suite de ce travail. Nous l'exploiterons comme base de départ et chercherons à étendre ou améliorer ses fonctionnalités afin de répondre aux problématiques soulevées par le relevé d'art pariétal.

Nous pouvons constater qu'en termes de visualisation, toutes les approches proposées représentent les données de manière décontextualisée. Dans le meilleur des cas, les outils sont mobilisables sur tablette sur le terrain, mais pas réellement pensés pour une utilisation *in situ*, pourtant indispensable dans le cas de l'art pariétal. Dans le cas d'Aïoli, il n'existe pas de possibilités d'enrichissement sémantique à partir d'une vue globale de la scène. Sur le terrain, l'observateur doit choisir manuellement la vue la plus proche de l'entité qu'il cherche à observer et s'y conformer, ou bien prendre une nouvelle photographie et l'ajouter au projet, pour l'utiliser comme nouveau support d'annotation. En termes d'adaptabilité aux besoins du terrain, cette approche est donc en avance par rapport aux autres systèmes d'information, mais ce fonctionnement peut s'avérer très contraignant dans le cas de projets de grandes ampleurs ou morphologiquement complexes.

Comme nous le détaillerons en (Chapitre 4), notre posture est de considérer qu'un système d'information dédié au relevé d'art pariétal ne peut s'affranchir du contact nécessaire avec l'objet réel. À ce titre, il est nécessaire de **résorber la fracture entre l'objet réel et l'objet virtuel** pour garantir la continuité du travail réalisé *in situ* et *ex situ*. Autrement dit, il s'agit de rétablir le lien qui lie les données numériques à l'objet réel. Cependant, l'hybridation représente un défi complexe qui soulève des questions à la fois technologiques et méthodologiques. Or la vocation de la **réalité mixte** est de permettre, par définition, l'hybridation d'environnements réels et virtuels, pour cette raison, ce domaine constitue une piste sérieuse en réponse au problème de **contextualisation**. Nous allons donc maintenant nous intéresser plus spécifiquement à ce domaine afin d'en mieux cerner les enjeux et d'identifier des leviers permettant de doter Aïoli de modalités d'interactions en réalité mixte.

3.3. Hybridation réel / virtuel

3.3.1. L'expérience de Brunelleschi

Au XV^{e} siècle, Filippo Brunelleschi réalise à Florence une expérience devant le baptistère San Giovanni, s'apparentant étonnamment à ce que l'on appelle aujourd'hui communément *Réalité Mixte*. En effet, dans la continuité directe de ses travaux sur la perspective (section 1.1.1.2), et afin de prouver que ce nouveau mode de représentation pouvait permettre aux observateurs de se projeter dans la réalité peinte par l'artiste, Brunelleschi réalise, vers 1415, la démonstration d'un procédé permettant de superposer l'une de ses peintures à la réalité, mais également de projeter certains éléments réels dans son tableau. Pour cela, un spectateur se positionne précisément sur la place, au point exact où l'image et l'objet réel se superposent. Ainsi placé, il observe le baptistère d'un seul œil, à travers un trou creusé dans une *tavoletta*[22]. Entre la *tavoletta* et le monument, un miroir (lui aussi percé) réfléchit la peinture de Brunelleschi. De cette manière, peinture et réalité coïncident parfaitement. Plus surprenant encore, l'arrière-plan du tableau n'est pas peint, mais remplacé par une couche d'argent poli. Ainsi, c'est bien le ciel réel qui pénètre dans le tableau, offrant alors au baptistère un ciel animé de nuages (Figure 65).

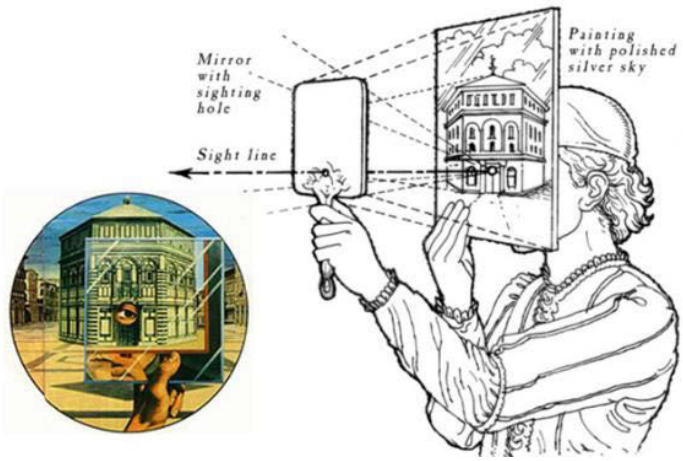


Figure 65 : Schéma de l'expérience de Brunelleschi. Illustration : (Argan, Wittkower, 1990).

S'il ne s'agit là que d'une expérience visant à démontrer l'intérêt d'une construction géométrique rigoureuse en peinture, elle souligne tout de même des préoccupations modernes inhérentes au domaine de la réalité mixte. En tout premier lieu, le positionnement précis de l'observateur par rapport aux scènes réelles et virtuelles est encore aujourd'hui considéré comme un important verrou scientifique ou technologique, compris dans la problématique plus générale de l'alignement ou recalage réel-virtuel.

L'expérience soulève également des interrogations quant aux frontières entre le réel et le virtuel (ici, la peinture). Puisque le ciel est projeté dans la peinture, il semble légitime de considérer que la réalité fait partie de la représentation. Mais cette représentation est contextualisée, superposée à l'objet réel. Est-elle alors un élément virtuel intégré dans une scène réelle ? Quel est l'environnement dominant ? Cette nécessité d'établir les frontières du monde *réel* et de classifier les applications de réalités virtuelle et mixte a mené à différentes théories schématisant ce que l'on nomme aujourd'hui le ***continuum réel-virtuel***. Enfin, cette démonstration soulève également la question du support de visualisation. Ici, le

[22] Littéralement « planchette » en italien. Ici il s'agit du tableau.

spectateur n'observe que d'un œil, nous faisons donc face à un dispositif monoculaire. En outre, l'environnement du baptistère est observé en vision directe. Dans les classifications modernes, le procédé proposé par Brunelleschi s'apparenterait ainsi à un dispositif de réalité mixte de type *optical see-through* Monoculaire (section 3.3.4).

Cette expérience nous montre que, tout comme la perspective, la réalité mixte est profondément liée à la question de la mesure de l'existant. Conceptuellement, nous pouvons même considérer que la réalité mixte est l'aboutissement logique d'un processus qui démarre de la lecture du réel (dans notre cas l'acquisition), passant par l'analyse et l'interprétation de sa complexité (forme, surface, composition, etc.), et qui retourne vers le réel pour l'enrichir de toutes les informations produites.

3.3.2. Le continuum Réel –Virtuel

Bien que les termes *réel* et *virtuel* soient relativement bien définis dans la littérature, les frontières qui les séparent sont bien plus ténues lorsque l'on fait appel à ces notions pour qualifier des environnements, qu'ils soient numériques ou non. De nombreux chercheurs se sont ainsi attachés à proposer des taxonomies des environnements réels et virtuels dans un premier temps, puis des applications de réalité virtuelle et de réalité mixte dans un second. Un premier pas en ce sens a été proposé par Milgram et Kishino (1994) qui s'appuient sur la considération suivante pour construire le continuum réel – virtuel, encore largement utilisé de nos jours comme point de départ de nombreuses taxonomies : un environnement qui n'est ni complètement réel, ni complètement virtuel, est nécessairement une hybridation des deux. En effet, plutôt que de considérer ces deux concepts antithétiques, les auteurs estiment plus correct de les considérer comme les extrémités d'un continuum entre lesquelles s'étend ce qu'ils nomment *réalité mixte* (Figure 66). Cette notion désigne indifféremment les environnements dans lesquels éléments physiques et virtuels coexistent (Milgram et al., 1995).

Figure 66 : Continuum Réalité – Virtualité. Illustration : auteur, d'après (Milgram et al., 1995).

Ce continuum repose donc sur l'idée que les notions de réel et de virtuel sont trop exclusives pour permettre de décrire l'ensemble environnements possibles. Cependant, s'il semble faire consensus, il demeure très générique. La distinction entre la réalité augmentée et la virtualité augmentée, toutes deux incluses dans la notion de réalité mixte, n'est établie que par la nature de l'environnement dominant, difficile à quantifier. En outre, le postulat d'une frontière entre le réel et le virtuel est aujourd'hui remis en question, au profit du concept de réalité hybride (Vial, 2016).

3.3.3. Définitions

En nous en tenant à ce continuum, la réalité mixte (MR) est un terme générique qui se définit comme l'ensemble des environnements impliquant la fusion d'environnements réels et virtuels. Il comprend

deux sous-ensembles aux frontières incertaines, la réalité augmentée (AR) et la virtualité augmentée (AV) [23].

L'expression *Réalité Augmentée* a été introduite dans les années 1990. Sa première utilisation est attribuée à Caudell et Mizell (1992), qui désignent par ce biais une technologie destinée à *« augmenter le champ visuel de l'utilisateur avec les informations nécessaires à la réalisation de la tâche en cours »*. Néanmoins ce concept a été exploré bien avant qu'un nom ne lui soit attribué. On considère habituellement les travaux d'Ivan Sutherland dans les années 1960 comme un tournant dans le domaine, bien que le brevet déposé par Howard Grubb[24] évoque dès 1901 la volonté de projeter une image virtuelle sur un objet réel (Aukstakalnis, 2016). Notons en outre l'utilisation d'affichages tête haute (ou *Head-Up Displays*) dans le domaine militaire à partir de la fin des années 1950 (Nijboer, Patterson, 2016).

En 1994, Paul Milgram, cherchant à établir une taxonomie des dispositifs de réalité virtuelle et réalité mixte, définit la réalité augmentée comme l'ensemble des cas dans lesquels l'affichage d'un environnement autrement réel est augmenté au moyen d'objets virtuels (graphiques informatiques). En 1997, Ronald Azuma précise cette définition en établissant trois caractéristiques fondamentales permettant de souligner la coexistence d'éléments réels et virtuels au sein d'un environnement unique (Azuma, 1997). La réalité augmentée est ainsi définie comme un système qui :

- Combine le réel et le virtuel ;
- Permet d'interagir de manière interactive en temps réel ;
- Est enregistré en trois dimensions et respecte l'alignement réel-virtuel ;

Cette définition permet notamment d'exclure les films à effets spéciaux, et introduit un important lien sémantique entre le réel et le virtuel reposant sur la notion de spatialité. Selon Bottecchia (2010), elle demeure néanmoins discutable sur la dernière règle. En effet, la présence d'objets virtuels en trois dimensions n'est pas justifiée dans les cas où la présence d'éléments signalétiques est suffisante (flèches, indications textuelles, icônes, etc.). Dans la même optique, le recalage réel-virtuel n'est pas toujours indispensable, puisque certaines informations n'imposent pas une localisation précise (par exemple des informations génériques comme une vitesse, des consignes générales, ou encore des instructions d'intervention à destination d'un opérateur, etc.).

Bien que la définition d'Azuma demeure aujourd'hui une référence, celle de Fuchs et al. (2010), plus générique, s'adapte plus facilement à la multiplicité des applications existantes. Ils définissent ainsi la réalité augmentée par sa finalité qui est *« de permettre à une personne de réaliser des activités sensori-motrices et cognitives dans un nouvel espace en associant l'environnement réel et un environnement*

[23] Le concept de *virtualité augmentée* est beaucoup moins exploré dans la littérature. Nous pouvons simplement souligner que dans la taxonomie proposée par Milgram, les environnements de virtualité augmentée sont plus proches des environnements virtuels que réels. Par analogie avec la réalité augmentée, nous pouvons déduire que ce terme renvoie aux environnements virtuels augmentés d'éléments du monde réel. Autrement dit, il s'agirait d'introduire dans des environnements virtuels des représentations temps-réel de l'état actuel d'éléments existants. Bien que ce concept ne soit pas dénué d'intérêt dans le cas de l'art pariétal, particulièrement pour des applications de *monitoring*, nous l'écarterons pour la suite de ce travail qui s'intéresse spécifiquement au relevé.

[24] Brevet intitulé « *A New Collimating-Telescope Gun-Sight for Large and Small Ordnance* », soit « Une nouvelle lunette de visée pour grandes et petites munitions » (Grubb, 1901)

virtuel ». Pour la suite de ce travail, nous adopterons cette définition, bien que l'approche que nous proposons en (Chapitre 5) respecte également celle d'Azuma.

3.3.4. Technologies

En règle générale, l'architecture matérielle des dispositifs de réalité augmentée comprend quatre éléments principaux (Ghouaiel, 2014) :

- Un **ordinateur**, pour réaliser les calculs (estimation des poses, rendu graphique, etc.) ;
- Un **dispositif d'affichage**, pour présenter à l'utilisateur les images calculées et lui permettre ainsi de visualiser son environnement augmenté ;
- Une ou plusieurs **caméras**, pour filmer l'environnement réel ;
- Des **capteurs**, principalement utilisés pour la localisation et le suivi (GPS, centrale inertielle, etc.) mais aussi pour les fonctions courantes (Wifi, Bluetooth, etc.).

Dans certains dispositifs ces composants matériels sont regroupés en une seule entité, c'est par exemple le cas des tablettes ou des casques autonomes. Il est possible de diviser les technologies d'affichage de réalité augmentée selon leur support physique. Elles comprennent alors trois grandes catégories[25] (Figure 67) : les dispositifs *fixés à la tête* de l'utilisateur, plus connus sous l'appellation anglophone *Head-Mounted Displays (HMD)*, les dispositifs *mobiles*, et les dispositifs dits *spatiaux* c'est-à-dire disposés dans l'environnement (Bimber, Raskar, 2006 ; Arnaldi et al., 2018).

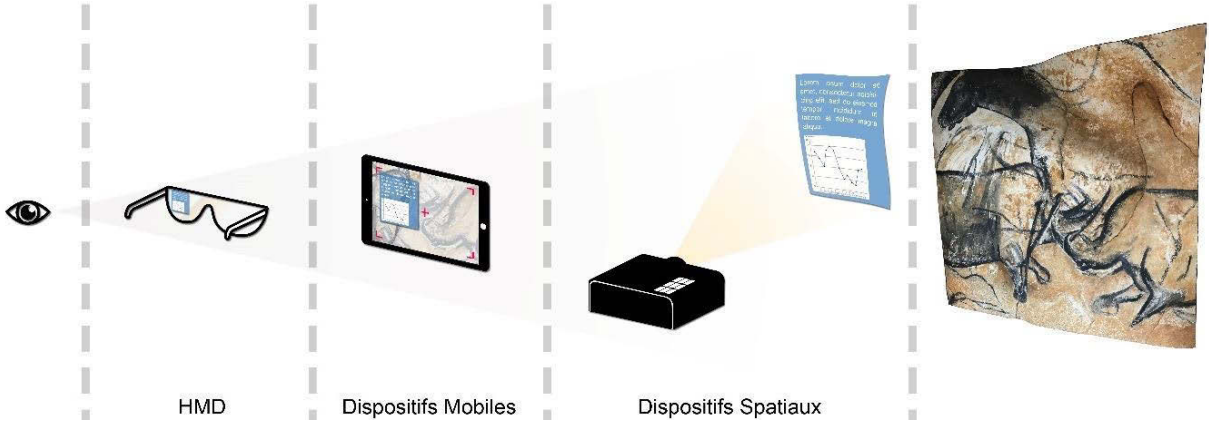


Figure 67 : Schéma des différentes catégories de dispositifs d'affichage de réalité augmentée. Illustration : auteur.

3.3.4.1. Dispositifs fixés à la tête

Cette première catégorie comprend les dispositifs de type casques ou lunettes. Ces derniers sont catégorisés selon deux critères : le principe exploité pour l'affichage : *Optical see-through* ou *video see-through* (Figure 68), et leur mode de vision : monoculaire, bi-oculaire, ou binoculaire (Figure 69).

[25] D'autres systèmes basés sur des lentilles de contact ont récemment été mis au point, mais leur développement technologique n'est pas encore suffisamment avancé pour envisager des utilisations concrètes. Ces systèmes ne permettent aujourd'hui que d'afficher des symboles ou images très simples. Pour l'instant, les défis principaux de ces technologies concernent principalement l'amélioration matérielle (résolution, consommation d'énergie, etc.) et la prise en compte des facteurs humains (Arnaldi et al., 2018). Dans les sections suivantes, nous éluderons donc ces dispositifs.

Affichages *optical see-through* : Les dispositifs dits o*ptical see-through* ou *dispositifs en vision directe*, fonctionnent à l'aide d'éléments optiques qui sont mi-transparents et mi-réfléchissants pour combiner le monde réel et les éléments virtuels. Ces éléments laissent passer une quantité suffisante de lumière du monde réel, ce qui permet à l'utilisateur de voir directement son environnement. Des images générées par ordinateur sont simultanément projetées sur ces dispositifs, ce qui permet à l'utilisateur de voir également les éléments virtuels (Aukstakalnis, 2016). Ces systèmes sont majoritairement basés sur l'utilisation de guides d'ondes, des structures permettant de contraindre la propagation de la lumière à travers un système optique par réflexion interne. Ces dispositifs sont confortables dans la mesure où la vision directe de l'environnement est conservée, ce qui semble naturel pour l'utilisateur. En revanche, ils présentent des limites notables, par exemple concernant le champ de vision relativement faible, le contraste et la luminosité des images projetées, ou encore les aberrations chromatiques (Schmalstieg, Höllerer, 2016).

Affichages *video see-through* : Le fonctionnement des dispositifs dits *video see-through* repose sur l'utilisation d'une ou plusieurs caméras capturant des vidéos de l'environnement réel de l'utilisateur. Ces vidéos sont traitées en temps réel par le processeur graphique pour être combinées avec des images virtuelles, puis affichées à l'écran. Avec cette méthode, l'utilisateur ne perçoit donc pas directement son environnement, mais une image numérique de celui-ci. Cette approche est plus flexible que la précédente dans la mesure où la balance entre les contenus réels et virtuels peut être définie librement par le créateur de contenus. Elle présente des limites, principalement la résolution des écrans utilisés, la latence entre l'acquisition de la vidéo et la restitution sur l'écran, et le champ de vision limité. L'immersion et la proximité de l'écran avec les yeux peuvent également procurer un certain inconfort lors de longues sessions d'utilisation (Lucas et al., 2013). Dans le cas des grottes ornées, ces dispositifs sont délicats à envisager étant donné les risques potentiels de collisions liés à l'absence de contact direct avec l'environnement, nous les écarterons donc pour la suite de ce travail.

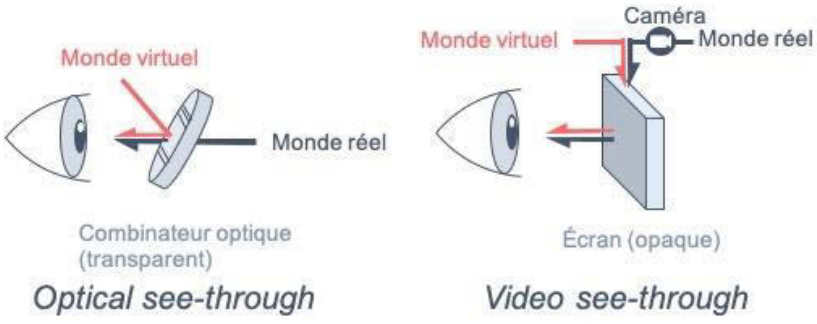


Figure 68 : Schéma des principes d'affichage AR. Illustration : auteur, d'après (Yadav, 2018).

La technologie d'affichage utilisée n'est pas suffisante pour classifier les différents HMD. Ces derniers se distinguent également par le mode de vision exploité (Figure 69).

Monoculaire : Une monture monoculaire dispose d'un seul dispositif d'affichage placé devant l'un des deux yeux de l'utilisateur. Ce dernier peut donc percevoir son environnement réel à travers son autre œil. Ces dispositifs ne permettent pas la perception de la profondeur, et sont donc principalement utilisés comme écrans d'informations.

Bi-oculaire : Une monture bi-oculaire fournit une image identique aux deux yeux de l'utilisateur, ce qui est principalement exploité lorsque le dispositif ne dispose que d'une seule caméra. Cependant, le fait

que les images perçues par chaque œil soient identiques prive l'utilisateur de la perception de la profondeur.

Binoculaire : Une monture binoculaire dispose de deux dispositifs d'affichage distincts, fournissant à chaque œil une image différente. Ce principe permet vision stéréoscopique, nécessaire à la perception de la profondeur. D'un point de vue logiciel, ce principe est plus coûteux en calculs car il exige de produire simultanément deux vues différentes de l'environnement virtuel.

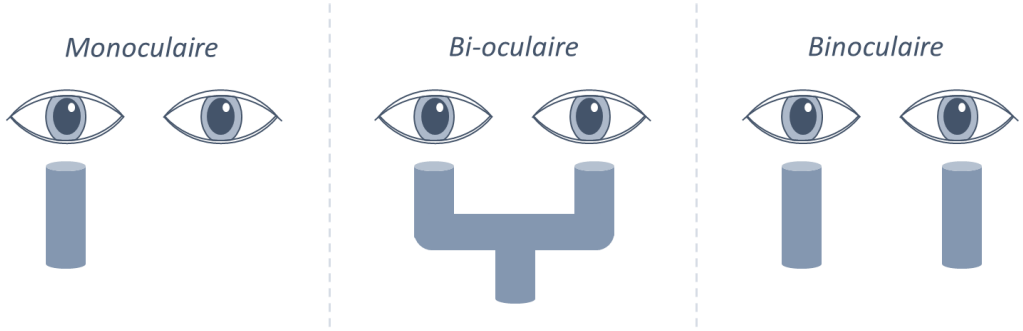


Figure 69 : Schéma des différents modes de visions. Illustration : auteur, d'après (Aukstakalnis, 2016).

3.3.4.2. Dispositifs Mobiles

Le terme de *réalité augmentée mobile* se réfère à l'utilisation de *smartphones* ou de tablettes. Ces dispositifs autonomes intègrent directement l'ensemble des composants matériels nécessaires à l'élaboration d'une application de réalité augmentée. Pour cette raison, leur distribution massive sur le marché[26] constitue un vecteur efficace pour démocratiser les usages de réalité augmentée, y compris auprès du grand public. En pratique, les dispositifs mobiles utilisent l'affichage *video see-through*. La caméra du mobile capture l'environnement réel, les images du flux vidéo sont traitées en temps réel par le processeur graphique pour être combinées avec des images virtuelles. Les images résultantes sont affichées sur l'écran (tactile) de l'utilisateur, lequel sert également d'Interface Homme-Machine (IHM) pour permettre l'interaction avec les contenus.

Si du point de vue de la perception la scène affichée est moins convaincante que sur HMD *video see-through* en raison de l'absence de la stéréoscopie, le contact visuel permanent avec l'environnement rend l'utilisation AR Mobile plus confortable. Par ailleurs, étant donnée la généralisation des dispositifs mobiles dans la population, nous pouvons supposer que les interactions basées sur l'utilisation de l'écran tactile demandent un temps d'apprentissage relativement faible.

3.3.4.3. Réalité augmentée spatiale

Les dispositifs de réalité augmentée spatiale sont des systèmes fixes dans l'espace, basés sur l'utilisation de vidéoprojecteurs qui projettent des images directement sur des objets de l'environnement réel, provoquant ainsi une superposition naturelle. Ils ne nécessitent pas d'équipement de visualisation spécifique pour l'observateur, ce qui rend ces dispositifs multiutilisateurs. Ils imposent cependant l'utilisation de capteurs pour permettre les interactions avec les éléments réels augmentés. De plus, le fait que le vidéoprojecteur soit fixe constitue une contrainte importante, qui implique notamment une

[26] D'après une étude menée par (Hootsuite Media Inc, 2019) à partir des données de *Google Consumer Barometer*, en France, en janvier 2018, 71% de la population adulte déclare utiliser un *smartphone*. Pour les tablettes, le taux est de 41%.

scénarisation préalable du cadre d'utilisation. De fait, ce principe n'offre pas une flexibilité suffisante pour permettre de servir de base à une approche générique pour le relevé en réalité augmentée.

3.3.4.4. Enjeux technologiques des systèmes d'affichage de réalité augmentée

Les technologies utilisées pour l'affichage dans des systèmes de réalité augmentée sont encore sujettes à différentes limites qui restent à surmonter. Le mécanisme de la vision humaine, particulièrement, soulève de nombreuses questions d'ordre matérielles.

Le champ de vision humain moyen couvre environ 180° horizontalement. La perception de l'environnement y est inégale : elle n'est optimale que sur un champ restreint qui ne couvre qu'environ 3 à 5° du champ visuel total, et diminue en s'éloignant de cette zone. Ainsi, les différentes activités visuelles mobilisent des zones variables du champ de vision (Figure 70). Ainsi, selon les contextes d'application, les besoins en termes de champ de vision (FOV) varient. Pourtant, les champs de vision proposés par les technologies HMD actuelles demeurent très faibles en comparaison avec les capacités humaines (entre 20° et 40° selon les modèles). Cette limitation est liée à l'utilisation de guides d'ondes. Des solutions alternatives imposent des compromis en termes de résolution (Maimone et al., 2014). Ce choix n'est pas anodin en tenant compte d'une autre particularité du mécanisme de la vision humaine : l'acuité visuelle. Il s'agit de la capacité de l'œil à séparer visuellement deux détails distincts. Une acuité normale correspond environ à 1/60 de degré. Un dispositif d'affichage souhaitant respecter cette valeur devrait respecter une résolution de 60 pixels par degrés[27]. Dans le cas de la réalité augmentée mobile, ces enjeux sont moindres dans la mesure où la surface d'affichage se limite à la taille de l'écran, c'est-à-dire globalement entre 5,5 et 13 pouces selon l'appareil utilisé.

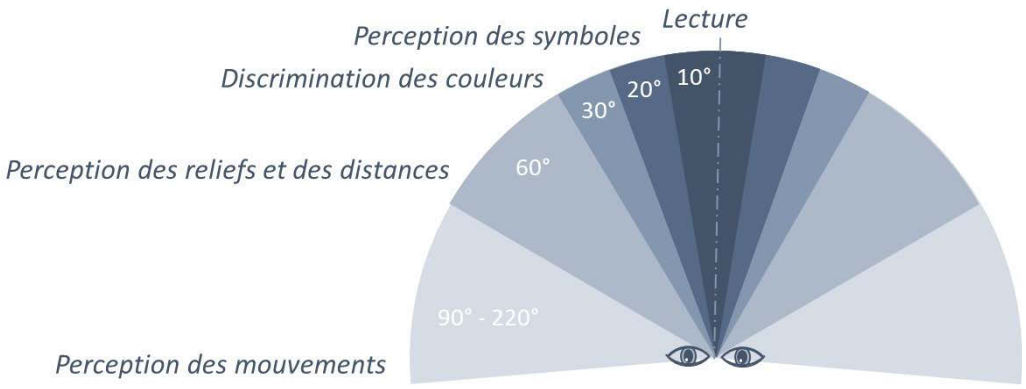


Figure 70 : Les différentes activités visuelles mobilisent des zones variables du champ de vision. Illustration : auteur.

Un autre défi technologique soulevé par le mécanisme de la vision est lié à l'accommodation. En conditions naturelles, les yeux convergent et accommodent à la même distance, c'est-à-dire sur l'objet observé. Dans le cas des HMD, les yeux accommodent sur le dispositif d'affichage, alors qu'ils convergent devant ou derrière ce plan. Ce phénomène, connu sous le nom de conflit accommodation-convergence cause pour l'utilisateur de la fatigue visuelle, qui, au mieux se limite à un inconfort, au pire peut entraîner des douleurs, des maux de tête, et des vertiges (Aurat et al., 2013)[28]. Enfin, d'autres

[27] Sur un champ de vision de 40°, cela impliquerait déjà une définition horizontale de 2400 pixels par œil !
[28] En réponse à ce problème et pour améliorer le réalisme des scènes augmentées, des recherches se concentrent sur le développement d'écrans plénoptiques, plus connus sous le nom *light field displays*. Le principe consiste à afficher simultanément les images de la scène virtuelles sur différents plans focaux (Tewes, 2017 ; Maimone et al., 2017).

facteurs liés au matériel utilisé influent sur la qualité d'une expérience de réalité augmentée. À ce sujet, le lecteur pourra se référer à Kuntz et al. (2018).

Nous pouvons constater que les dispositifs autonomes de réalité augmentée, et particulièrement les HMD, font encore l'objet de développements. S'ils ne constituent pas encore, à ce jour, une technologie mature, ils demeurent toutefois très prometteurs pour l'avenir. **En ce qui nous concerne, nous concentrerons nos efforts dans un premier temps sur les supports mobiles, sans nécessairement occulter ces dispositifs, et tâcherons dans la mesure du possible de proposer une approche polyvalente.**

3.3.5. Applications au patrimoine

Dans le domaine du patrimoine culturel, les applications de réalité augmentée concernent principalement la médiation et la diffusion. Plus spécifiquement, la majorité cible le tourisme ou des cadres muséaux. Elles nécessitent un traitement préalable des données 3D, prennent lieu dans des environnements connus ou contrôlés, et sont scénarisées en amont afin de définir à l'avance les contenus associés aux données, et chaque modalité d'interaction (Bekele et al., 2018).

Ridel et al. (2014) proposent par exemple « The revealing flashlight », un système de réalité augmentée spatiale permettant aux utilisateurs de visualiser un objet patrimonial en rendu expressif. La direction de l'index de l'utilisateur guide l'orientation d'une lampe torche virtuelle, qui, au lieu d'éclairer l'objet, le texture à l'aide de techniques de rendu expressif basées sur les concavités et convexités calculées à partir d'une numérisation 3D (Figure 71). Ce travail montre l'intérêt de la réalité mixte pour l'aide à la visualisation des reliefs, mais implique une importante préparation des données en amont et un contrôle total de l'environnement. En outre, les dimensions des objets observés doivent être contenues.

L'application proposée par Pierdicca, et al. (2015) a pour but de permettre la visualisation en réalité augmentée mobile de sites archéologiques disparus. Leur proposition repose sur la définition de métadonnées décrivant chaque site, comprenant notamment des informations concernant sa géolocalisation. Le recalage des modèles 3D des sites archéologiques sur la réalité est alors basé sur leur géoréférencement et sur les coordonnées GPS du mobile de l'utilisateur (Figure 72). La visualisation peut être accompagnée d'une description sommaire fournie par le reste des métadonnées. Ce travail montre qu'il est possible de proposer une approche plus générique permettant la visualisation AR de sites sans connaissance préalable ni maîtrise de l'environnement dans lequel évoluera l'utilisateur. Cependant, elle n'offre aucune forme d'interaction avec les données à l'exception de leur visualisation. Par ailleurs, comme nous l'avons vu en (section 2.4), la description sous forme de texte libre est insuffisante pour profiter de l'apport des informations sémantiques.

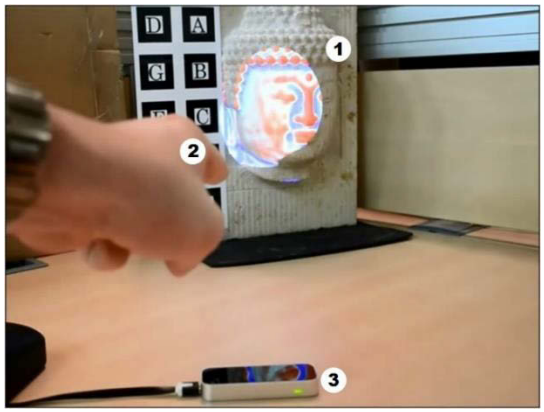


Figure 71 : Application de réalité augmentée spatiale *The revealing flashlight*. Illustration : (Ridel et al., 2014).

Figure 72 : Visualisation AR du modèle 3D du site archéologique de Chan Chan (Pérou). Illustration : (Pierdicca et al., 2015).

Dans le cas de l'art pariétal, peu de travaux s'intéressent à la réalité augmentée. Parmi eux, MARCH (Choudary et al., 2009) est une application mobile de réalité augmentée qui permet l'augmentation des images du flux vidéo par les dessins de relevé des experts (Figure 73). Les auteurs partent du constat que les traces anthropiques dans les grottes sont très souvent difficiles, voire impossibles, à discerner à l'œil nu. Leur application permet de les mettre en évidence à partir de tracés fiables établis par les experts. Le suivi est basé sur l'utilisation d'images de références. Cette application, qui date déjà d'une décennie, constitue un premier pas intéressant vers un système dédié au relevé. Bien sûr, en l'état il ne s'agit pas tant d'une application dédiée au relevé que d'un outil de visualisation contextualisée de travaux déjà effectués. En outre, elle ne permet pas de profiter des attributs géométriques de l'objet observé ni des descriptions établies par les experts et associées au relevé.

Blanco-Pons et al. (2019) partent d'un constat similaire, et proposent une application de réalité augmentée mobile destinée à améliorer l'expérience des utilisateurs lors de visites guidées du site *Cova dels Cavalls* à Tirig (Espagne). Il s'agit donc principalement de médiation : les utilisateurs pointent leur mobile vers les parois, et des informations supplémentaires sont affichées en superposition. L'application permet en particulier de faciliter l'identification des éléments peints en accentuant les peintures (Figure 74) par la superposition de tracés colorés préparés en amont.

Figure 73 : Principe de l'application MARCH. Illustration : (Choudary et al., 2009).

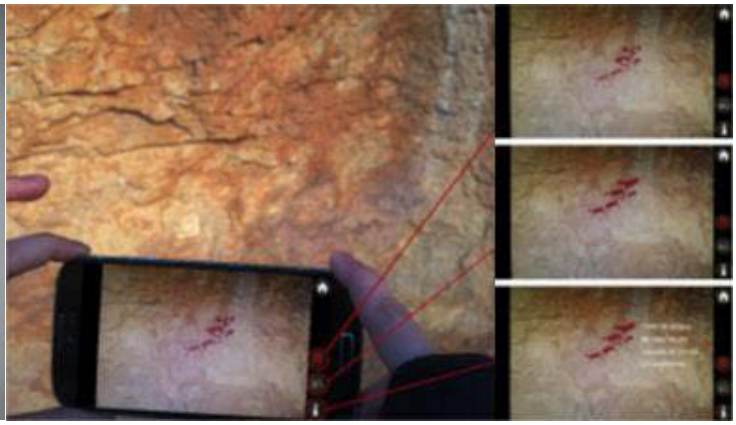


Figure 74 : Application de réalité augmentée pour le site Cova dels Cavals. Illustration : (Blanco-Pons et al., 2019).

Citons également la technique du relevé par vidéoprojection, déjà expérimentée pour l'étude des parois de la Chaire-à-Calvin en 2005 (Pinçon et al., 2018), qui s'apparente fortement à de la réalité augmentée

spatiale. Elle consiste à mettre en place un espace de travail collaboratif pour le relevé des panneaux, exploitant une charte graphique partagée, tout en préservant l'intégrité des parois étudiées. Pour cela, les releveurs utilisent une tablette graphique et un logiciel de dessin vectoriel, utilisant une vue orthographique du panneau pour réaliser les observations. Cet espace de travail est projeté sur la paroi à l'aide d'un vidéoprojecteur, ce qui permet de visualiser dans l'environnement réel et en temps réel les tracés réalisés par les observateurs (Figure 75). Cette expérience est très convaincante, mais elle ne permet pas de profiter véritablement de l'apport de la 3D dans la mesure où seule une projection orthographique est utilisée comme support d'annotation. En outre, elle impose un point de vue fixe.

Figure 75 : Relevé par vidéoprojection. Illustration : (Pinçon et al., 2018).

Enfin, comme évoqué précédemment, le travail de Barbier et al. (2017) se concentre sur l'art mégalithique du Néolithique, et s'intéresse aux potentiels de la réalité augmentée pour l'annotation sémantique de modèles 3D. L'application proposée, nommée MAAP, cible la monture Hololens (Figure 76). Elle permet de charger un modèle 3D de l'objet d'étude, et de l'annoter au moyen de différents outils d'accentuation et de sélection (voir section 2.4.2). La description des annotations est réalisée sous forme de texte libre à l'aide de reconnaissance vocale. Cependant, contrairement aux travaux précédents, la réalité augmentée n'est pas exploitée ici à des fins de contextualisation. En effet, l'objet 3D n'est pas visualisé dans son environnement réel mais uniquement *ex situ*.

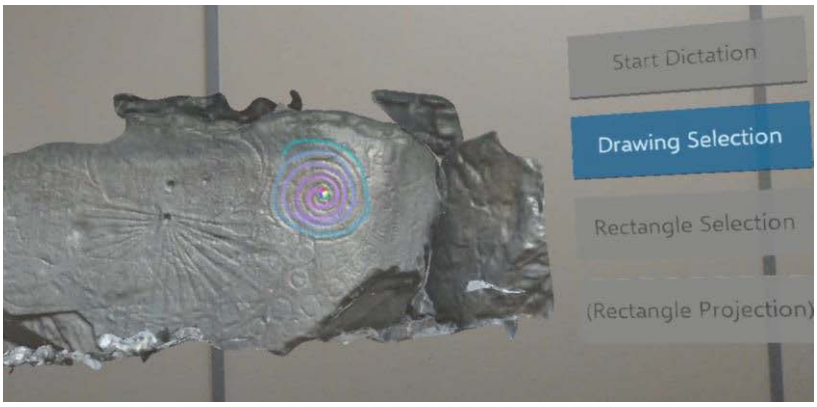


Figure 76 : Vue de l'application MAAP pour Hololens. Illustration : (Barbier et al., 2017).

Finalement, nous pouvons constater, comme le soulignent Bekele et al. (2018) que les applications de réalité augmentée dans le domaine du patrimoine concernent principalement la *médiation* et l'*éducation*, la catégorie *exploration* étant largement minoritaire[29]. Parmi les travaux de cette catégorie, une partie

[29] Pour établir une classification des applications AR et VR pour le patrimoine, les auteurs proposent cinq catégories : **l'éducation** qui vise l'apprentissage des aspects historiques, **la médiation** qui vise à améliorer l'expérience des visiteurs dans

s'intéresse à l'aide à la visualisation en profitant des attributs géométriques des données numériques, mais sans explorer l'apport des descriptions sémantiques ou, plus globalement, des systèmes d'information. Inversement, les travaux s'intéressant à la définition d'approches génériques font généralement l'impasse sur l'apport des attributs géométriques et offrent peu d'outils d'interaction et d'enrichissement. En outre, elles reposent principalement sur des méthodes de suivi basées sur le géoréférencement des données et l'utilisation de capteurs GPS, ce qui est incompatible avec des environnements intérieurs. Enfin, nous devons également souligner que la grande majorité des applications étudiées implique un traitement préalable des données 3D ainsi qu'une connaissance *a priori* de l'environnement. Or, dans notre cas, les projets réalisés sur Aïoli doivent être disponibles en réalité augmentée indifféremment de leur contenu et de leur nature, ce qui signifie que l'environnement de l'objet représenté ne peut être considéré comme connu.

3.3.6. Le suivi et le recalage

Nous avons vu précédemment que la réalité augmentée implique la superposition d'images virtuelles sur des vues de l'environnement réel. Cette exigence est d'ailleurs explicite dans la définition de Ronald Azuma, et concerne la capacité du système à ancrer le contenu virtuel 3D dans le monde réel afin qu'il semble faire partie de l'environnement physique. Ainsi, le défi principal d'un système de réalité augmentée est de déterminer la relation entre les différents systèmes de coordonnées des éléments de la scène augmentée (Figure 77). Lorsque toutes ces relations sont connues, les images des objets virtuels et la vue de l'environnement peuvent être correctement alignées, pour fusionner en une seule image cohérente. On parle d'alignement réel/virtuel ou de recalage, ou encore de *registration* en anglais.

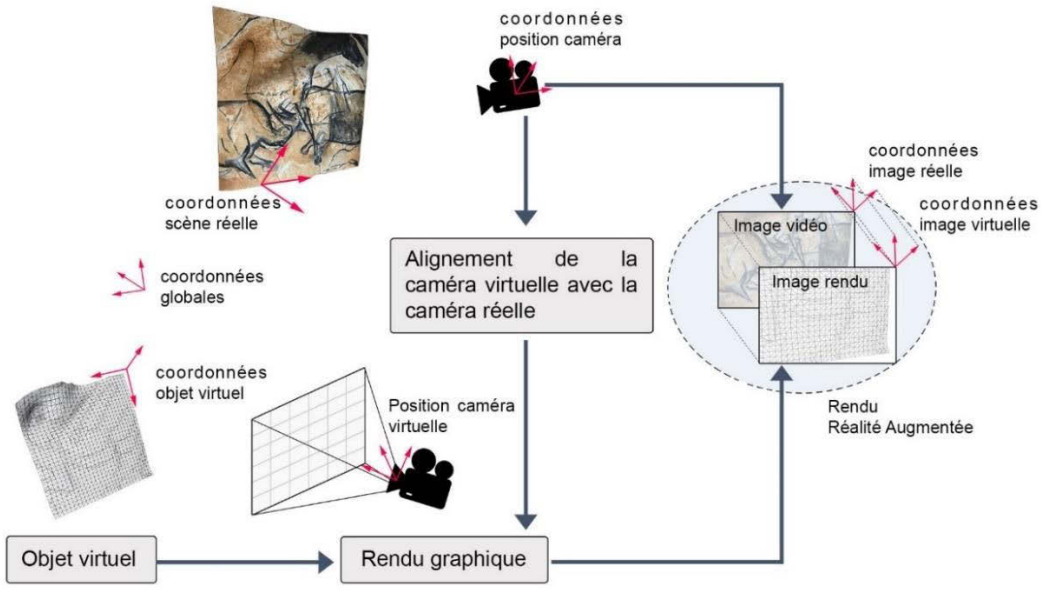


Figure 77 : Schéma du principe du recalage. Il s'agit d'établir une relation entre les poses successives de la caméra du dispositif et celles de la caméra virtuelle agissant sur le rendu de la scène 3D, afin de rendre les deux images superposables. Illustration : auteur.

Ce processus repose plus particulièrement sur la capacité du système à déterminer en continu la pose (position et orientation) du dispositif d'affichage par rapport à un ou plusieurs point(s) connu(s) de

des cadres muséaux, **l'exploration** qui aide les utilisateurs à visualiser et explorer, **la reconstruction** qui vise à visualiser des vues historiques reconstituées, et **les musées virtuels** qui présentent au public des collections numériques.

l'espace physique. On parle de suivi ou de *tracking*. Les points utilisés comme références dans l'espace physique sont nommés *ancres*. Il s'agit d'un problème complexe pour plusieurs raisons :

- **Précision :** Le calcul des poses doit être très précis. En effet, une erreur de 1 degré sur la pose représente environ un écart de 17,5 centimètres à une distance de 10 mètres, soit 1,75 mètres à 100 mètres. Cet écart n'est pas acceptable dans de nombreuses applications, y compris la nôtre : nous avons pu voir dans le chapitre précédent à quel point les stigmates observés sur les parois pouvaient être ténus !
- **Vitesse :** Le calcul doit être réalisé rapidement pour éviter la latence, qui nuit fortement à la fluidité de l'expérience et au confort de l'observateur. Or les dispositifs autonomes disposent d'une capacité de calcul relativement limitée ;
- **Généricité :** Les différentes approches ne s'adaptent pas nécessairement à tous les environnements. Par exemple, les méthodes nécessitant un GPS ne sont pas fonctionnelles en intérieur, celles basées sur la vision par ordinateur peuvent présenter des limites dans certaines conditions lumineuses, etc.

Pour aborder ce problème, il existe plusieurs approches. Certaines se basent sur l'utilisation de capteurs externes ou *traqueurs* spécifiquement dévolus au suivi de l'évolution de l'observateur dans l'environnement réel. Ces approches peuvent être intéressantes dans des cadres applicatifs dans des environnements maîtrisés, mais sont inappropriées pour l'art pariétal, dont la fragilité impose de minimiser autant que possible l'aménagement de l'environnement et l'installation de matériel à demeure. D'autres approches, que nous étudierons dans la section qui suit, se basent sur la vision par ordinateur et incidemment sur l'utilisation d'une caméra embarquée sur le dispositif d'affichage.

3.3.6.1. Suivi inertiel

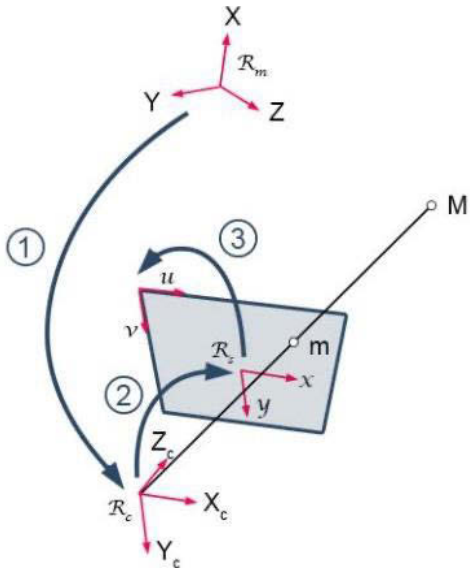
Le suivi inertiel est une méthode basée sur l'utilisation d'une centrale inertielle (ou IMU pour *Inertial Measurement Unit*) embarquée sur l'objet suivi pour mesurer sa pose. Il s'agit d'un dispositif généralement composé de trois éléments : un accéléromètre trois axes, un gyroscope trois axes et un magnétomètre trois axes, mesurant respectivement l'accélération linéaire, la vitesse angulaire, et la direction du champ magnétique terrestre. La position et l'orientation de la centrale inertielle dans un référentiel donné peuvent être calculées à partir de la double intégration des mesures d'accélération linéaire, et de la vitesse angulaire. Cette méthode est peu coûteuse et peut fournir des taux de mise à jour élevés avec une faible latence. Cependant, ces dispositifs sont imprécis et l'accumulation d'erreurs d'intégration au cours du temps introduit une dérive. Pour compenser, une solution est de recourir à la fusion de données. Un exemple d'algorithme utilisé pour cela est le filtre de Kalman, qui permet d'estimer les biais de mesure en utilisant des données issues de capteurs tiers, par exemple GPS (Kalman, 1960). Cette méthode peut être intéressante dans le cas de la réalité augmentée mobile étant donné que les *smartphones* actuels sont généralement dotés de centrales inertielles. Du fait de son imprécision, elle est surtout mobilisée dans les méthodes de suivi hybrides.

3.3.6.2. Basé sur la vision par ordinateur

Le suivi basé vision, aussi appelé suivi visuel est une technique reposant sur l'utilisation d'une caméra embarquée sur le dispositif d'affichage. Le suivi du dispositif d'affichage, nécessaire pour le recalage des images virtuelles sur l'environnement réel revient alors au problème d'estimation de pose, bien connu en vision par ordinateur.

3.3.6.2.1. L'estimation de pose

L'estimation de pose, aussi appelée calibrage extrinsèque a pour objectif de calculer la position et l'orientation de la caméra par rapport à la scène. Pour ce problème, on suppose que les paramètres intrinsèques de la caméra sont connus (distance focale et coordonnées du point principal), ce qui est le cas dans la plupart des méthodes de localisation par vision (Tamaazousti, 2013). La caméra est modélisée par le modèle sténopé. Il s'agit de considérer mathématiquement la formation d'une image comme une projection perspective. Ce modèle transforme un point 3D *M* de l'espace en un point *m* sur le plan image, et peut se décomposer en trois transformations géométriques successives (Figure 78).



La composition de ces trois transformations peut se schématiser de la façon suivante :

$$(X_m\ Y_m\ Z_m) \xrightarrow{T} (X_c\ Y_c\ Z_c) \xrightarrow{P} (x\ y) \xrightarrow{A} (u\ v)$$

Figure 78 : Transformations successives du modèle sténopé. Illustration : auteur, d'après (Orteu, 2008).

Avec :

① La transformation entre le repère du monde et celui de la caméra. Il s'agit de la combinaison d'une rotation et d'une translation, transformant un point 3D de coordonnées (X_m, Y_m, Z_m) en un point 3D de coordonnées (X_c, Y_c, Z_c) ;

② La transformation entre le repère caméra et le repère capteur. Il s'agit d'une projection perspective transformant un point 3D de coordonnées (X_c, Y_c, Z_c) en un point image de coordonnées (x, y) en unités métriques ;

③ La transformation entre le repère capteur et le repère image, qui permet de convertir les coordonnées images métriques (x, y) en coordonnées images discrètes (u, v), en pixels.

- **La transformation ① entre le repère monde et le repère caméra :**

Les paramètres de la transformation entre le repère monde R_m (placé de manière arbitraire) et le repère caméra R_c (dont l'origine correspond au centre optique de la caméra) **sont appelés paramètres extrinsèques** de la caméra. Cette transformation rigide peut se décomposer en une rotation [R] et une translation [t] telles que :

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} = [R] \begin{bmatrix} X_m \\ Y_m \\ Z_m \\ 1 \end{bmatrix} + [t]$$

Ou, de façon équivalente :

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} = \begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} X_m \\ Y_m \\ Z_m \\ 1 \end{bmatrix} \Leftrightarrow \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} = [T] \begin{bmatrix} X_m \\ Y_m \\ Z_m \\ 1 \end{bmatrix} \text{ où } [T] = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

- **La transformation ② entre le repère caméra et le repère capteur :**

La transformation entre le repère caméra R_c et le repère capteur R_s est une projection perspective, notée [P], transformant un point 3D de coordonnées (X_c, Y_c, Z_c) en un point 2D de coordonnées (x, y) (unités métriques). Elle peut s'écrire de la façon suivante :

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \sim \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix}$$

Où f désigne la focale de la caméra, en unités métriques et $\sim$ désigne une égalité à un facteur d'échelle multiplicatif près (les coordonnées homogènes (x, y) désignant l'ensemble des vecteurs proportionnels à (x, y)). Plus précisément, en notant s le facteur d'échelle, on obtient :

$$s \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix}$$

On en déduit $s = Zc$ et $x = \frac{f\,X_c}{Z_c}$, $y = \frac{f\,X_c}{Z_c}$

- **La transformation ③ entre le repère capteur et le repère image :**

La transformation entre le repère capteur R_s et le repère image R_i permet de convertir les coordonnées métriques (x, y) d'un point 2D de l'image en coordonnées discrètes (u, v), en pixels. En considérant le cas simplifié où le *skew* est négligeable, c'est-à-dire le cas où les lignes et les colonnes de l'image sont orthogonales, cette projection, notée [A] peut s'exprimer de la façon suivante :

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} k_x & 0 & c_x \\ 0 & k_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

Où :

k_x et k_y désignent le nombre de pixels par unité de longueur dans les directions x et y du capteur. Dans le cas de pixels carrés, ce qui est généralement le cas des photographies numériques, $k_x = k_y$;

(c_x, c_y) sont les coordonnées de l'intersection de l'axe optique avec le plan image.

Ces trois transformations nous permettent d'écrire l'équation du modèle de sténopé :

$$m \sim [A][P][T]\, M$$

Ou encore, en posant [C] = [A][P], $f_x = k_x f$, et $f_y = k_y f$:

$$\begin{array}{ccccc} \text{m} & \sim & \text{[C]} & \text{[T]} & \text{M} \\ \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} & \sim & \begin{bmatrix} f_x & 0 & c_x & 0 \\ 0 & f_y & c_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} & \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \\ 0 & 0 & 0 & 1 \end{bmatrix} & \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \end{array}$$

Où :

$[u \quad v \quad 1]^T$ sont les coordonnées homogènes du point 2D m dans le repère image R_i. ;

[C] est la matrice correspondant à la transformation entre le repère caméra et le repère image. Elle combine les transformations (2) et (3). Ses paramètres sont appelés **paramètres intrinsèques** de la caméra. On l'appelle **matrice de calibrage** ;

(f_x, f_y) correspondent à la distance focale, en pixels par unités ;

(c_x, c_y) sont les coordonnées du point principal de l'image, en pixels. Le point principal correspond à l'intersection de l'axe optique avec le plan image. Généralement, il se situe approximativement au centre de l'image ;

[T] est la matrice correspondant à la transformation (1) entre le repère monde et le repère caméra. Elle comprend la rotation et la translation de la caméra dans le repère monde. Les paramètres de cette transformation sont appelés **paramètres extrinsèques** de la caméra, il s'agit de sa pose ;

$[X \quad Y \quad Z \quad 1]^T$ sont les coordonnées homogènes du point 3D M dans le repère monde.

Autrement dit, en utilisant le modèle sténopé, onze paramètres de la caméra permettent de déterminer les coordonnées homogènes $(u \quad v \quad 1)$ d'un point de l'image en fonction des coordonnées homogènes $(X \quad Y \quad Z \quad 1)$ du point de la scène : **cinq paramètres intrinsèques** connus par calibration de la caméra (quatre si l'on néglige le *skew*), et **six paramètres extrinsèques** (en exprimant la rotation avec les angles d'Euler).

Notons que le modèle sténopé présenté ci-dessus ne modélise qu'une caméra *parfaite*, sans prendre en compte les éventuelles distorsions géométriques induites par le système optique utilisé (voir section 2.2.1.1). Pour limiter les erreurs, il convient de prendre en compte ces distorsions. En pratique, en partant du modèle sténopé, ces distorsions peuvent être modélisées en introduisant une quatrième transformation [D], liant les coordonnées capteur *idéales* et les coordonnées capteur *réelles*[30].

Ainsi, lorsque suffisamment de correspondances entre des points 2D et des points 3D sont établies, il devient possible de calculer les paramètres de la matrice [T], et donc la pose de la caméra[31]. **Le suivi basé vision, et donc l'estimation temps réel des poses de la caméra, suppose donc une connaissance *a priori* de plusieurs points de l'environnement réel dans lequel évolue l'utilisateur**. Dans ce contexte, plusieurs stratégies existent : le suivi avec marqueurs, le suivi avec une image de référence, le suivi basé sur un modèle 3D de l'objet à suivre, ou le suivi basé sur des méthodes de localisation et cartographie simultanées (SLAM).

3.3.6.2.2. Suivi avec marqueurs

Cette méthode repose sur l'utilisation de marqueurs insérés dans l'environnement réel. La complexité des marqueurs varie selon les applications, cependant la méthode la plus répandue consiste à utiliser des quadrilatères de dimensions connues et dont les coins servent d'ancres. Des informations d'identification peuvent être codées à l'intérieur du quadrilatère. Cette méthode a été largement popularisée par la bibliothèque logicielle ARtoolkit (Kato, 2017), mais de nombreuses alternatives ont été ensuite proposées, par exemple la bibliothèque ArUco (Garrido-Jurado et al., 2016), implémentée dans OpenCV (Bradski, 2019). Le suivi avec marqueurs se fait en 5 étapes (Havard, 2018) :

1. L'image provenant de la caméra est passée en niveaux de gris, puis binarisée par seuillage.
2. Les contours sont extraits de l'image binarisée. Parmi eux, les formes carrées sont détectées grâce à une transformée de Hough (Hough, 1962).
3. Les coins des carrés sont détectés en utilisant un algorithme de détection de coins, par exemple Harris (Harris, Stephens, 1988)
4. Les positions des coins détectés sont utilisées pour calculer la pose de la caméra.
5. Les informations d'identification sont extraites des marqueurs (dans la partie centrale dans le cas des marqueurs ARToolKit), et comparées à une base de marqueurs pour déterminer quelle augmentation doit être rendue sur le dispositif d'affichage.

Cette méthode présente l'avantage d'être rapide, simple à mettre en œuvre, et relativement stable. Cependant, elle présente des inconvénients. La détection des marqueurs peut être entravée en cas d'occlusions partielles, et la méthode est sensible aux conditions lumineuses ambiantes. De plus, l'environnement doit être préparé préalablement, ce qui peut s'avérer contraignant, particulièrement dans le cas d'applications au domaine du patrimoine (Figure 79). Dans les cas s'inscrivant dans la durée, par exemple le suivi d'un site d'art pariétal, les marqueurs doivent être entretenus.

[30] Le lecteur intéressé pourra se référer à (Orteu, 2008 ; Tamaazousti, 2013).
[31] Le lecteur intéressé par la résolution de ce système d'équations pourra se référer à (Sturm, 2015).

Figure 79 : Marqueurs installés dans le cabinet de travail de Charles V (château de Vincennes) pour une application de réalité augmentée, 2009.Installation : © CMN/CNRS/AGP/Axyz. Photographie : © Mr~Poussnik.

3.3.6.2.3. Suivi avec image de référence

Avec l'amélioration de la puissance de calculs des appareils utilisés pour faire de la réalité augmentée, il est devenu possible d'utiliser des images plus complexes pour calculer les poses de la caméra. Pour estimer la pose de la caméra, des points d'intérêt sont recherchés dans l'image puis caractérisés à l'aide de descripteurs, par exemple SIFT, ou SURF. Il s'agit ensuite d'évaluer la correspondance entre les points d'intérêt de l'image provenant de la caméra et ceux de l'image de référence. S'ils sont considérés comme équivalents, c'est-à-dire si l'image de référence est détectée dans l'image de la caméra, la pose peut être calculée à l'aide d'algorithmes similaires à ceux utilisés dans les techniques avec marqueurs.

Cette méthode s'apparente à celle du suivi avec marqueurs tout en étant visuellement plus agréable. Elle présente l'avantage d'être plus robuste (notamment aux occlusions) et permet d'utiliser des images de l'environnement, mais exige plus de calculs. Par ailleurs, les images de référence doivent provenir d'objets globalement plans ou de forme peu complexe. Dans le cas contraire, une image référence ne sera reconnue que lorsque le point de vue de l'utilisateur coïncidera parfaitement avec celui de la photographie. Enfin, pour des applications s'inscrivant dans la durée, les images de référence doivent être stables c'est-à-dire non seulement fixes dans l'espace mais aussi dans le temps. Dans le cas d'objets patrimoniaux susceptibles d'évoluer voire de se dégrader dans le temps, cela peut être problématique.

3.3.6.2.4. Suivi basé sur un modèle 3D

Cette méthode repose sur l'utilisation d'un modèle 3D représentant l'objet que l'on cherche à suivre dans l'environnement réel. Il existe plusieurs approches de suivi basées sur l'utilisation de modèles. On distingue ainsi les méthodes par détection, qui permettent de calculer la pose de la caméra sans connaissance préalable de la position de l'objet à détecter, et les méthodes récursives qui exploitent la pose de la caméra précédente pour estimer la pose actuelle. Ces méthodes sont nécessairement initiées par une détection, qui peut être également utilisée pour réinitialiser la boucle en cours de suivi en cas d'erreur (Lima et al., 2010). Globalement, il s'agit de considérer l'objet 3D comme une collection de vues 2D de référence, à partir desquelles sont extraites des caractéristiques (points d'intérêts, arrêtes, etc.). Pour chaque image du flux vidéo de la caméra, les mêmes caractéristiques sont calculées, puis utilisées dans une phase de mise en correspondance (dans la totalité de l'image dans le cas des méthodes de détection, dans la zone prédite dans le cas des méthodes récursives). Une fois l'objet détecté, la pose

de la caméra peut être calculée de manière similaire aux méthodes précédentes, et la vue peut être augmentée (Havard, 2018).

Comme les méthodes basées sur l'utilisation d'images de références, ces méthodes présentent l'intérêt d'être visuellement agréables dans la mesure où elles ne nécessitent pas de modification de l'environnement, tout en étant plus polyvalentes. Selon les caractéristiques utilisées pour la mise en correspondance, elles ne présentent pas les mêmes avantages et limites. Les méthodes basées sur la texture peuvent poser des problèmes dans le cas d'éléments monochromes ou trop répétitifs. Les approches basées uniquement sur la forme de l'objet ou sur des attributs géométriques (Drost et al., 2010 ; Hinterstoisser et al., 2013) ne sont pas soumises à ces difficultés. Notons enfin que certaines de ces méthodes nécessitent l'utilisation de caméras RGBD (Krull et al., 2015).

3.3.6.2.5. Suivi basé sur le SLAM

La localisation et cartographie simultanées ou *Simultaneous Localization And Mapping* (SLAM) est une méthode qui consiste à construire de façon simultanée une carte de l'environnement et à s'y localiser (Figure 80). Elle ne nécessite pas de connaissance préalable de l'environnement. Initialement dédiée au domaine de la robotique, la méthode présente de nombreuses applications, notamment dans les domaines du *Structure From Motion* (SFM) ou, dans le cas qui nous intéresse, la réalité augmentée sans marqueurs. Le SLAM vise à traiter deux problèmes majeurs dans le domaine de la robotique – particulièrement pour le déplacement autonome :

- La compréhension des caractéristiques de l'environnement dans lequel le robot évolue (à quoi ressemble mon environnement ?), c'est-à-dire sa cartographie 2D ou 3D ;
- L'interprétation des données collectées par ses différents capteurs et la localisation du robot (où suis-je ?)

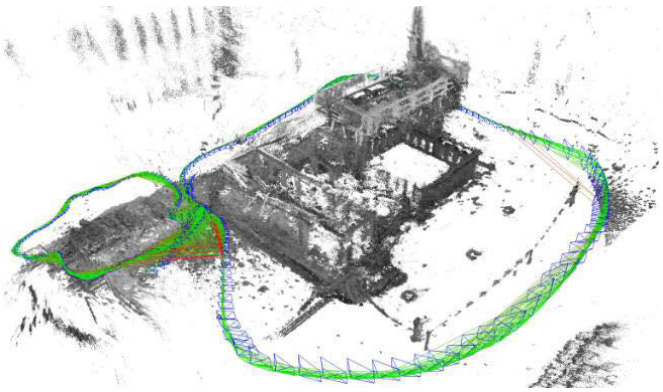


Figure 80 : Données produites par SLAM. En vert : poses successives des caméras. En gris : cartographie 3D de l'environnement. Illustration : (Engel et al., 2014).

Les résolutions de ces deux problèmes sont intrinsèquement liées, dans la mesure où le robot a besoin d'une cartographie pour se situer, et inversement, de se localiser pour obtenir une cartographie. Le SLAM se définit donc comme une méthode liant ces deux processus dans une boucle visant l'amélioration des résultats des deux problèmes par retours itératifs. Les stratégies adoptées pour cela peuvent impliquer des capteurs spécifiques comme des scanners laser pour établir des nuages de points denses de l'environnement, ou encore des caméras RGBD (dans ce cas on parle de RGBD-SLAM), comme dans le cas de l'algorithme proposé par Microsoft pour le Hololens (Newcombe et al., 2014).

L'approche qui concerne particulièrement le suivi pour la réalité augmentée est le *Visual SLAM*, le plus souvent monoculaire. Elle n'implique que des caméras et repose sur des algorithmes de vision par ordinateur. Il en existe plusieurs approches. Celles basées sur les filtres, par exemple un filtre de Kalman, considèrent la localisation et la cartographie de manière entrelacée. Les poses de chaque vue et la cartographie de l'environnement doivent être actualisées à chaque nouvelle image traitée. Des méthodes plus performantes ont vu le jour avec le *Keyframe-based SLAM* (Strasdat et al., 2012).

Dans ces systèmes, la localisation et la cartographie sont séparées en deux étapes. La localisation s'effectue pour chaque image, tandis que la reconstruction 3D et l'optimisation ne reposent que sur des images clés. Toutes les images du flux vidéo n'ont donc pas le même rôle. Certaines images sont utilisées uniquement pour localiser la caméra dans son environnement à partir des données accumulées sur les poses précédentes et des points caractéristiques dans l'espace. Il ne serait pas efficace de toutes les exploiter pour la cartographie 3D de l'environnement, car le déplacement entre deux images successives d'un flux vidéo est trop faible pour permettre d'appliquer les algorithmes de reconstruction avec précision (Tamaazousti, 2013).

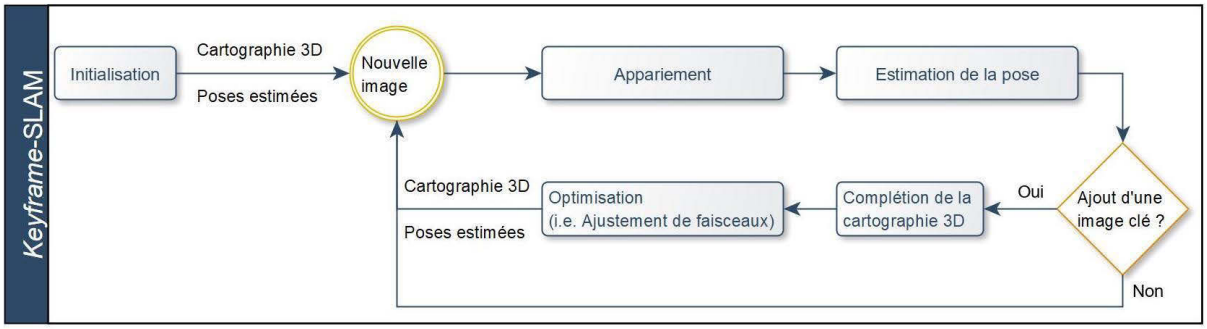


Figure 81 : Diagramme résumant le fonctionnement d'une méthode Keyframe-based Visual Monocular SLAM. Illustration : auteur, d'après (Younes et al., 2017).

Le principe des méthodes *Keyframe SLAM* est schématisé dans l'illustration ci-dessus (Figure 81), et comprend plusieurs étapes.

- **Initialisation :** Pour accumuler suffisamment de données pour engager la suite du processus, il s'agit dans un premier temps de détecter des éléments fixes qui pourront être suivis dans les images suivantes, et calculer leur position en 3D. Les points d'intérêt détectés dans les trois premières images clés de la caméra (C_1, C_2, et C_3) sont appariés. Pour calculer les poses de ces images, l'algorithme détermine d'abord un **repère arbitraire**, par exemple en fixant la pose de C_1 à l'origine. La position et l'orientation relative de C_2 par rapport à C_1 sont calculées grâce aux points homologues. Ces derniers sont ensuite triangulés pour générer un nuage de points 3D épars. Enfin, la pose de C_3 est calculée par mise en correspondance 3D-2D. À l'issue de cette étape, trois poses de caméras sont donc connues ainsi qu'une liste de points 3D. Pour finir, ces données sont raffinées par ajustement de faisceaux, puis la suite du processus est déclenchée.

- **Nouvelle image :** Les poses de chaque image du flux vidéo peuvent être calculées à partir des données issues des images clés précédentes. Pour cela, les points d'intérêt détectés sont mis en relation avec ceux des images précédemment sélectionnées. Lorsqu'au minimum trois correspondances entre des points 2D et des points 3D sont établies, il devient possible de calculer les paramètres extrinsèques de la vue actuelle.

- **Image clé**[32] **:** Si l'image est sélectionnée comme une image clé, les positions 3D des points détectés sont calculées par triangulation pour compléter la cartographie de l'environnement. Cette étape est suivie d'une phase d'optimisation, généralement par ajustement de faisceaux. L'objectif est de réestimer simultanément les coordonnées des points 3D et les poses des caméras afin de trouver la solution qui minimise les erreurs de reprojection (Triggs et al., 2000). Les stratégies d'optimisation peuvent également reposer sur la détection de fermetures de boucles dans la trajectoire pour limiter les dérives accumulées au cours du temps.

Dans le cas du suivi AR, ces méthodes[33] présentent l'avantage de fournir les poses de chaque *frame* en temps réel et de nombreuses ancres (potentiellement chaque point de la cartographie 3D). Le fait qu'elles n'exigent pas de connaissance préalable de l'environnement permet de limiter les contraintes en termes de préparation. Dans des applications patrimoniales, l'absence de marqueurs et de traqueurs est un atout non négligeable, ainsi que le fait qu'elles n'exigent pas de matériel particulier à l'exception d'une caméra. Cependant, le caractère incrémental du processus cause des problèmes de dérive dus à l'accumulation d'erreurs au cours du temps qui nuit au recalage des données virtuelles. Par ailleurs. La méconnaissance du facteur d'échelle de la trajectoire de la caméra et de la cartographie générée limite très fortement l'utilisabilité pratique.

Une façon d'aborder ce problème est de recourir à des méthodes hybrides, basées sur le principe du *Visual SLAM* mais exploitant les données issues de capteurs supplémentaires pour améliorer la robustesse du système. En particulier, les données provenant de la centrale inertielle du dispositif et d'un récepteur GPS permettent de déterminer le facteur d'échelle et de limiter la dérive de la trajectoire. Ces approches sont exploitées dans les SDK ARCore (Nerurkar et al., 2017) et ARKit (Maerz, Stoppa, 2018).

3.3.7. Bilan des méthodes de suivi

Nous avons étudié dans cette section les méthodes de suivi utilisées pour le recalage en réalité augmentée. Nous avons vu que la localisation 3D précise du dispositif d'affichage est essentielle pour que l'information virtuelle ajoutée à la scène soit correctement positionnée par rapport à l'environnement réel. Or, si des cadres applicatifs dans des environnements maîtrisés permettent d'obtenir des données précises à l'aide de dispositifs externes, ce n'est pas le cas de notre cas d'étude – l'art pariétal – dont la fragilité n'autorise pas l'aménagement de l'environnement ni l'installation de matériel à demeure. Ce contexte exclut donc également d'office les méthodes basées sur l'utilisation de marqueurs. Les méthodes basées sur l'utilisation d'images de référence s'adaptent mal aux formes complexes. Les méthodes basées sur l'utilisation d'un modèle 3D sont intéressantes par leur robustesse, mais imposent une connaissance préalable de l'environnement dans lequel évoluera l'utilisateur, ce qui constitue une contrainte dans la mesure où le relevé pariétal peut concerner aussi bien des parties très localisées du site étudié que sa globalité. Enfin, les méthodes basées sur le SLAM peuvent être relativement précises, peu coûteuses, adaptées à l'utilisation de matériel mobile – en particulier les

[32] Selon les approches, différents critères permettent de définir si une nouvelle image est une image clé ou non, par exemple en se basant sur le nombre d'appariements 2D avec l'image clé précédente, ou encore sur la proximité des poses successives.
[33] De nombreuses implémentations ont été proposées dans la littérature, par exemple PTAM (Klein, Murray, 2007), DT-SLAM (Herrera C et al., 2014), LSD-SLAM (Engel et al., 2014), RKSLAM (Liu et al., 2016), ou DSO (Engel et al., 2017). Des études de ces méthodes ont été proposées par (Fuentes-Pacheco et al., 2015 ; Younes et al., 2017).

tablettes, déjà sollicitées pour le relevé pariétal *classique* – et s'adaptent à des contextes variés (intérieur ou extérieur). Elles semblent les plus adaptées aux besoins de notre étude.

3.3.8. Les composantes applicatives d'un système de réalité augmentée

Dans les sections précédentes, nous nous sommes concentrés sur les technologies d'affichage pour la réalité augmentée, ainsi que sur les approches permettant le recalage des données virtuelles. Il est évident que ces deux éléments ne suffisent pas seuls à constituer une application de réalité augmentée. En partant des « *building blocks* » de Bimber et Raskar (2005), nous proposons de schématiser de façon suivante la constitution générale d'un système de réalité augmentée (Figure 82).

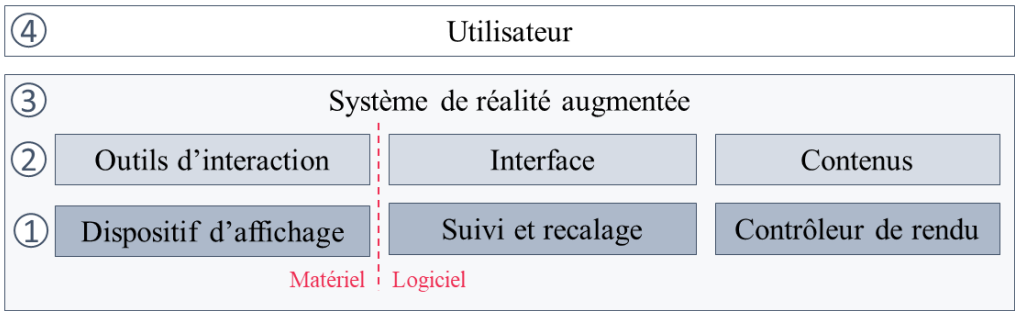


Figure 82 : Composants d'un système de réalité augmentée. Illustration : auteur.

- **1er niveau** : Il s'agit de la couche le plus bas niveau. Elle correspond aux éléments fondamentaux du système de réalité augmentée, nécessaires à son fonctionnement. Elle comprend trois éléments principaux, tant matériels que logiciels : le dispositif d'affichage, le suivi et le recalage, ainsi que le contrôleur de rendu graphique.
- **2nd niveau** : Cette couche intermédiaire est principalement implémentée sous forme logicielle. Elle comprend principalement les contenus, la façon de les présenter et gère la façon dont l'utilisateur interagit avec.
- **3e niveau** : Cette couche correspond au système de réalité augmentée en lui-même, en tant qu'entité complètement tournée vers un usage spécifique – en ce qui nous concernera pour la suite de ce travail, le relevé d'art pariétal. Ce niveau fait donc intervenir le contexte d'utilisation.
- **4e niveau** : Dans la mesure où le système est destiné à un utilisateur final, son acceptabilité soulève un verrou *humain* qui dépend de ses habitudes de travail, de son expérience, de ses compétences techniques, etc. Ce niveau constitue donc la couche le plus haut niveau, et fait intervenir différentes disciplines, notamment les sciences cognitives et l'ergonomie.

Dans le cadre de notre travail, nous intervenons principalement sur les couches supérieures, le développement de dispositifs d'affichage et d'algorithmes de suivi n'étant pas de notre ressort. En particulier, nous nous concentrerons en priorité sur les couches 2 et 3 pour proposer une preuve de concept, la 4e couche n'étant possiblement sollicitée qu'à terme pour l'évaluation de nos propositions. Nous devons donc, à la lumière de l'état de l'art, faire des choix concernant les éléments de la couche inférieure. Concernant le suivi et le recalage, le développement de solutions de réalité augmentée est considérablement facilité par la disponibilité de bibliothèques logicielles, SDK, et solutions logicielles intégrées.

De nombreuses méthodes de *Visual SLAM* sont disponibles sous forme de bibliothèques *open-source*, écrites en C/C++. Leur utilisation est intéressante dans les cas nécessitant un contrôle total sur tous les composants applicatifs. Cependant, elles se concentrent uniquement sur le suivi, et ne proposent pas de composants d'interfaces graphiques ou d'accès aux données. À l'opposée, les moteurs temps-réel comme Unity ou Unreal Engine, se présentent comme des solutions logicielles intégrées comprenant l'ensemble des éléments nécessaires à la création et au déploiement d'une application (système, graphisme, son, etc.). Ils sont dotés d'éditeurs 3D et permettent de compiler le même contenu sur différentes plateformes, ce qui simplifie le travail de développement. Des *frameworks* dédiés à la réalité augmentée (par exemple Vuforia ou Wikitude) élargissent leurs fonctionnalités en offrant des outils *clé en main* spécifiques, notamment pour le suivi ou le recalage.

Les SDK de réalité augmentée, par exemple Apple ARKit, Google ARCore, ou Microsoft MRTK, offrent différents outils de suivi et de recalage, de compréhension de l'environnement (détection de plans ou estimation de la lumière), et de rendu. Ces solutions permettent de profiter de briques technologiques tout en conservant un certain contrôle sur l'ensemble de l'application. En outre, elles ciblent des terminaux spécifiques, ce qui leur permet d'optimiser les méthodes de suivi proposées en fonction des capteurs disponibles. En termes de développement, les langages et environnements de développement proposés dépendent des SDK et des terminaux auxquels ils se destinent.

Si ces trois solutions sont intéressantes dans les cas de développement d'applications natives ou entièrement conçues depuis le même cadre logiciel, leur utilisation est délicate dans notre contexte, d'une part parce qu'elles sont globalement inadaptées à des utilisations web, d'autre part parce nous cherchons à étendre les fonctionnalités d'une plateforme préexistante.

Le WebVR est une API proposée en 2016 par les équipes de Mozilla-VR et Google Chrome pour prendre en charge les fonctionnalités des dispositifs de réalité virtuelle dans les navigateurs web (Casey, 2016). Elle permet ainsi d'étendre les capacités du web et de proposer des contenus VR lorsque l'utilisateur dispose d'un système d'affichage compatible. Dans la lignée de ce travail, en 2017, Mozilla propose le WebXR, une expérimentation visant à explorer les possibilités d'extension du WebVR pour y inclure les capacités des dispositifs de réalité mixte[34]. Repris par le *Immersive Web Community Group* du W3C, ce travail a mené à une ébauche d'API publiée en février 2019 (W3C®, 2019), remplaçant le WebVR devenu obsolète. À partir des SDK existant pour chaque plateforme, elle formalise les différentes façons dont les technologies XR exposent les données issues de leurs capteurs et synthétise les concepts communs, tels que les ancres et les poses de caméra. Enrichis de cette façon unifiée de communiquer avec les dispositifs d'affichage et de l'API WebGL pour le rendu 3D, les navigateurs web peuvent donc proposer des outils permettant le développement d'applications web polyvalentes et multiplateformes.

Ce fonctionnement présente différents avantages. D'une part Aïoli est une plateforme web, ces outils nous permettent donc de nous inscrire dans sa continuité sans réécrire les éléments existants ni modifier son architecture logicielle. D'autre part, le fait qu'il unifie les capacités des dispositifs permet de profiter des méthodes de suivi et de recalage propres aux différentes technologies, et de bénéficier de leurs améliorations progressives au cours des mises à jour des SDK sous-jacents, puisque les navigateurs agissent comme intermédiaire entre la 1ère couche et la 2nde (Figure 82). Cependant, il présente aussi des limites notables. Il ne s'agit pas encore d'une API stable, et elle évolue très rapidement. Les navigateurs

[34] Le terme XR renvoie alors à l'ensemble des environnements du continuum de Milgram, incluant VR et MR.

courants ne prennent pas encore en charge ces fonctionnalités, mais des navigateurs expérimentaux sont développés[35] par Mozilla et Google pour iOS et Android (Mozilla, 2017 ; Google-AR, 2017b ; 2017a).

3.4. Bilan des méthodes d'hybridation

Dans l'optique de réconcilier les moments d'observation *in situ* et *ex situ* du relevé d'art pariétal, nous avons étudié dans cette section les aspects matériels et logiciels des systèmes de réalité mixte. La volonté de tenir Aïoli comme base de départ et de proposer diverses modalités d'interactions adaptées aux différents contextes d'analyse (hors site ou sur site) implique l'élaboration d'une approche générique permettant d'étendre un système d'information doté de modalités *desktop* classiques à des modalités AR lorsque l'utilisateur dispose de la possibilité d'accéder à son site d'étude.

Or, nous avons pu voir qu'en ce qui concerne la réalité augmentée, le contexte des grottes ornées nécessite l'élaboration d'une méthode générique et flexible, ne nécessitant ni connaissance ni modification ou scénarisation préalable de l'environnement réel. Ce contexte exclut donc d'office les dispositifs de réalité augmentée spatiale. En termes de matériel, nous choisissons de nous concentrer dans un premier temps sur les dispositifs mobiles, d'une part parce que les tablettes sont déjà exploitées lors de relevés analytiques, d'autre part parce que les HMD, bien que prometteurs, ne constituent pas encore une technologie mature.

Nous avons étudié dans ce chapitre les méthodes de suivi utilisées pour le recalage en réalité augmentée. La localisation 3D précise du dispositif d'affichage est essentielle pour que l'information virtuelle ajoutée à la scène soit correctement positionnée par rapport à l'environnement réel. Toutefois, si des cadres applicatifs dans des environnements maîtrisés permettent d'obtenir des données précises à l'aide de traqueurs externes ou de marqueurs, ce n'est pas le cas de notre cas d'étude dont la fragilité n'autorise pas l'aménagement de l'environnement. Par ailleurs, la nature des sites d'art pariétal impose des contraintes supplémentaires, d'une part parce que les contextes souterrains sont inadaptés à l'utilisation de GPS, et d'autre part parce que les traces anthropiques à analyser – et les annotations correspondantes – peuvent être très fines, ce qui impose une grande précision en termes d'alignement réel/virtuel.

La question du recalage constitue donc un verrou majeur, étroitement lié aux méthodes exploitées pour le suivi AR. À l'issue de cet état de l'art, les approches de suivi sans marqueurs basées sur le *keyframe-based* SLAM apparaissent les plus adaptées aux besoins de notre étude. Or, l'estimation de pose constitue également le cœur du fonctionnement d'Aïoli dans la mesure où il s'agit d'une étape incontournable pour la spatialisation des ressources issues du relevé autour d'un nuage de points fédérateur. Nous pouvons donc supposer que la convergence de ces deux processus pourrait permettre de lever le verrou du recalage.

Enfin, nous avons étudié les différents aspects impliqués par le développement d'applications de réalité augmentée sans marqueurs. À partir du schéma des couches applicatives des systèmes de réalité augmentée (section 3.3.8), nous avons pu préciser notre champ d'intervention, qui cible en priorité les niveaux 2 et 3 pour l'élaboration d'une preuve de concept, c'est-à-dire respectivement l'implémentation logicielle s'intéressant aux interfaces, aux contenus et aux outils d'interactions, et le système en lui-

[35] Il n'existe pas encore de tel navigateur pour le Microsoft Hololens, mais Mozilla a récemment annoncé travailler sur le sujet (Bergstrom, 2019), nous pouvons donc supposer que les applications développées à l'aide de l'API WebXR deviendront naturellement compatibles (ou à moindres frais) lorsque le navigateur sera disponible.

même en tant qu'entité spécifiquement dévolue au relevé d'art pariétal. Ce positionnement nécessite donc de faire des choix en ce qui concerne les outils à exploiter.

À partir de l'étude des différents outils logiciels permettant le développement d'applications, nous avons conclu que l'ébauche d'API WebXR présentait des avantages très intéressants. En permettant le développement de contenus AR pour le web, elle nous permet en effet de nous inscrire dans la continuité de la plateforme Aïoli sans remettre en question son architecture logicielle. D'autre part, le fait que cette API unifie les différentes capacités des dispositifs AR permet de bénéficier, par le biais des navigateurs, des méthodes de suivi propres aux différentes technologies, et ainsi, de proposer une approche naturellement multisupport.

Chapitre 4. Discussion et approche proposée pour la définition d'un environnement numérique multimodal dédié au relevé d'art pariétal

4.1. Un moment privilégié de connaissance

Nous avons vu dans le Chapitre 1 que le contexte difficile des grottes (environnements souterrains, morphologie des parois, fragilité des sites, obscurité, manque de recul, etc.) et le caractère complexe des figures rendent le travail du relevé particulièrement délicat. À la fois œuvres d'art et témoins d'une activité humaine préhistorique, les figures pariétales comportent des dimensions multiples qui imposent à l'observateur un respect total des figurations tout en tenant compte du plus grand nombre possible de leurs caractères. Le contexte et les besoins de l'étude ainsi que la complexité des éléments de parois observés influencent donc logiquement la démarche de l'observateur, et à ce titre, l'absence de consensus sur un protocole global adapté à la diversité des situations semble tout à fait naturelle. Comme nous l'avons vu précédemment, les approches proposées pour le relevé d'art pariétal se divisent en deux grandes familles, qui se distinguent notamment par le degré de subjectivité accordé à l'opération, et par la diversité des éléments relevés.

Le relevé contemporain, à travers la méthode du *relevé analytique technique,* aspire à une certaine exhaustivité, considérant que la collecte du plus grand nombre possible d'informations aide à comprendre l'entité dans toutes ses dimensions et à la préserver. Dans cette même optique, cette méthode de relevé s'attache à minimiser les erreurs personnelles en recherchant une plus grande objectivité dans la démarche. La multiplication des lectures (dans le cadre d'études pluridisciplinaires par exemple) ou des procédés (diversification des outils d'acquisition pour lever les doutes ou compléter des observations) permet dans une certaine mesure d'affirmer les choix d'assignation et d'interprétation et d'affiner la compréhension d'un élément dans ses dimensions multiples. Néanmoins, il est évident que malgré ces efforts, le relevé ne peut prétendre à une parfaite exhaustivité.

Comme l'explique Aujoulat (1993a) : *« Une représentation pariétale se comporte, en quelque sorte, comme un objet fractal, dont une des propriétés est d'apporter des informations de plus en plus précises à mesure que l'on se rapproche de lui »*. Cette particularité est reprise par Lorblanchet (1993a) qui explique que, selon la distance qui sépare l'observateur de son objet d'étude, un tracé peut apparaître plus ou moins continu. À un mètre de distance, le tracé peut sembler constitué d'une seule ligne continue. Si l'observateur s'approche, ce même tracé peut se révéler formé de taches disjointes. En réduisant encore la distance, chaque tache peut apparaître composée d'une nappe de grains agglomérés. Ces mêmes grains, observés à l'aide d'un instrument optique, peuvent se révéler constitués de sous-éléments, et ainsi de suite.

Le degré de précision du relevé tel qu'il est tracé par l'observateur dépend donc en partie des objectifs qu'il s'est fixé, mais aussi des outils sur lesquels il appuie son tracé. Ainsi, comme dans le cas du relevé patrimonial, la technique employée pour rendre compte des observations dépend notamment de l'échelle de l'élément observé, de sa nature, ou de son contexte. De ce fait, qu'il s'agisse du processus ou de son résultat, le relevé d'art pariétal est inévitablement empreint de subjectivité, même dans ses applications les plus rigoureuses.

L'auteur du relevé opère donc nécessairement *in situ* une sélection d'éléments signifiants, qu'elle soit désirée ou assumée (par exemple lorsqu'elle dépend de ses objectifs ou de son environnement), ou subie (par exemple lorsqu'elle dépend des limites des outils utilisés ou des limites de ses capacités perceptives). La sélection opérée peut être également influencée par l'expérience de l'observateur, ses connaissances personnelles, mais aussi par l'héritage scientifique collectif sur lequel il s'appuie. C'est d'autant plus flagrant dans le cas du *relevé analytique plastique*, dont l'approche ne cherche pas particulièrement l'exhaustivité ou l'objectivité, et dans lesquels nous pouvons retrouver parfois la *patte* de l'auteur. Dans les deux cas, la démarche du relevé implique une lecture de l'existant pour aboutir à une compréhension scientifique communicable de l'objet d'étude (Figure 83).

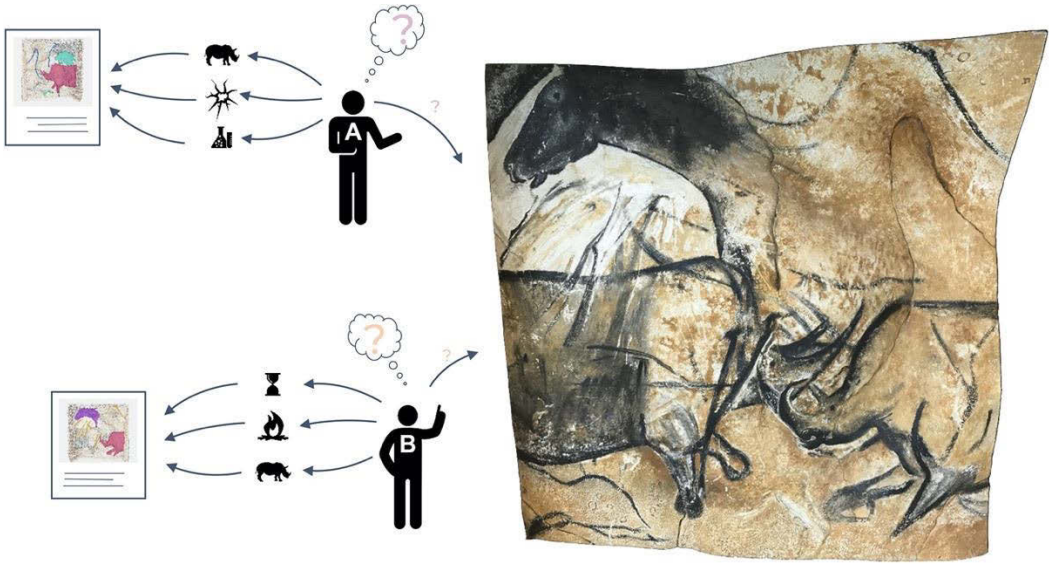


Figure 83 : Schéma du relevé *classique* : un moment privilégié d'observation du réel, qui aboutit à une représentation graphique destinée à illustrer un propos scientifique. Les données résultantes présentent un faible potentiel d'interrogation (section 1.3). Illustration : auteur.

4.2. Une modification profonde du processus de relevé : le rôle de l'observateur

Nous avons vu dans le Chapitre 2 que la démocratisation des outils et méthodes d'acquisition de données tridimensionnelles a sensiblement modifié le processus du relevé. Les progrès technologiques se sont principalement intéressés à la multiplication des points de mesure, à l'amélioration de leur précision, et à la facilité de mise en œuvre. Ainsi, l'automatisation du moment de *description*, tel que nommé par Saint-Aubin, a déporté le caractère subjectif du relevé – la lecture de l'existant, à travers la sélection *in situ* d'éléments signifiants – auprès d'une instance numérique, reléguant l'engagement cognitif et analytique de l'observateur derrière l'instrument de mesure, voire même derrière le dispositif d'affichage des données résultantes (Figure 84).

La dissociation et le séquençage des activités dans des moments distincts d'acquisition puis d'interprétation et de représentation constituent un changement de paradigme important : le relevé ne relève plus d'un moment privilégié d'observation du réel, mais plutôt d'une capacité à trier, écarter, et mettre en relation des données numériques.

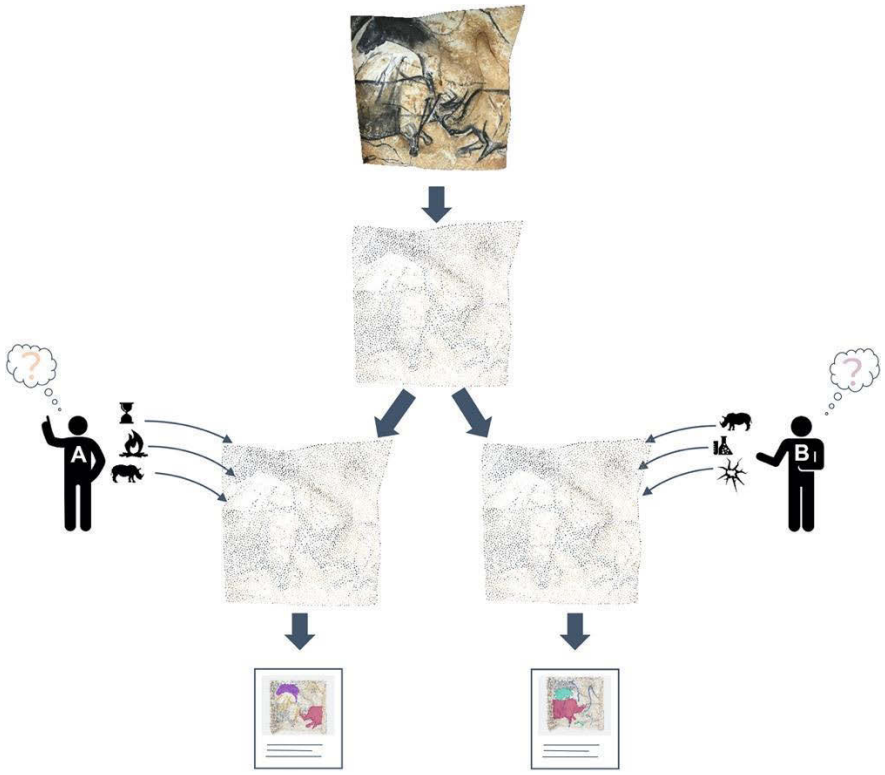


Figure 84 : Schéma du relevé numérique, centré sur des copies virtuelles de l'objet d'étude. Pour les acteurs, l'enjeu est de réussir à intégrer des connaissances aux représentations numériques. Les pratiques actuelles mènent à un effet de dispersion des données et connaissances (section 2.5). Illustration : auteur.

Par ailleurs, si, comme nous l'avons vu en section 2.3, les données 3D présentent un fort potentiel d'interrogation de plus en plus exploité dans le cadre des grottes ornées, globalement leur cadre d'utilisation ne vise qu'à fournir des documents (des représentations graphiques, principalement des captures d'écran) réinjectables dans les processus *traditionnels* de relevé. Les techniques relatives aux activités d'analyse et d'interprétation n'ont pas encore été suffisamment remises en question pour s'adapter à l'exploitation de ces nouvelles ressources. De fait, bien que l'usage de la 3D se soit progressivement imposé en raison de ses nombreux atouts, les acteurs des différents champs disciplinaires impliqués dans l'étude et la conservation de l'art pariétal ne disposent pas encore des outils ou méthodes nécessaires pour intégrer efficacement leurs connaissances à ces nouvelles données.

Il existe pourtant plusieurs expériences notables qui démontrent la volonté de faire converger les pratiques. Comme nous l'avons vu précédemment, des travaux (Robert et al., 2014 ; Fritz et al., 2016) proposent bien d'exploiter la relation projective entre les données 2D et 3D pour reprojeter au moyen de différentes techniques les résultats de relevés sur les modèles 3D des parois étudiées. Mais d'une part les chaînes opératoires proposées impliquent le relevé en tant que donnée figée, et d'autre part les opérations de projection impliquent des étapes manuelles.

Ainsi, l'absence de simultanéité entre le geste de l'observateur et la visualisation – sans doute liée à la diversité des outils impliqués dans les chaînes opératoires testées – fait de ces tentatives des outils de visualisation de travaux déjà effectués plus que des utilisations concrètes de données 3D au cours de l'étude. Fuentes et al. (2019) proposent bien une méthode permettant l'annotation dynamique d'un modèle 3D de la grotte de Comarques qui, associée à une charte graphique préalablement définie par l'équipe, permet de lier de l'information sémantique à un contenu 3D. Néanmoins, cette expérience prometteuse est, pour l'heure, un cas isolé dont la mise en œuvre est relativement complexe. En outre,

le contenu 3D utilisé est une représentation statique de l'objet d'étude à la segmentation préétablie, ce qui limite très fortement les possibilités en matière d'analyse ou de structuration des données.

4.3. Vers une approche intégrée

Nous avons vu en sections 2.4 et 3.1 que de nombreuses recherches s'intéressent à la création et à l'enrichissement de représentations numériques 3D, ou à l'amélioration de la gestion des données interconnectées. Mais la numérisation 3D et les systèmes d'information sont généralement considérés comme des sujets distincts, et peu de travaux les considèrent comme les parties complémentaires d'un cadre unique. Ainsi, nous pouvons distinguer, d'une part, les systèmes d'information centrés sur les connaissances, destinés à la gestion des activités de documentation, mais sans exploitation explicite ou directe des propriétés géométriques et visuelles des données 3D. D'autre part, les approches dotées d'outils d'annotations de données 3D se concentrent principalement sur la description des données numérisées, en manquant toute référence aux relations complexes propres au processus d'observation, d'interprétation ou de documentation.

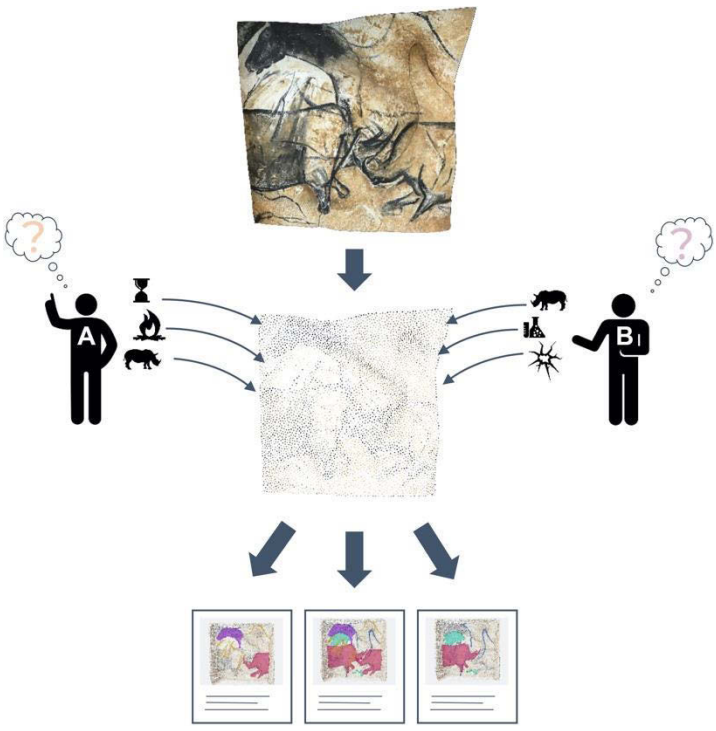


Figure 85 : Schéma de l'approche intégrée proposée par Aïoli : la donnée 3D est au cœur du système d'information et permet de fédérer les ressources liées à l'objet d'étude. Illustration : auteur.

L'approche proposée par De Luca et al. (2017) à travers l'outil Aïoli part du constat que le passage de pratiques basées sur les données à des pratiques basées sur la sémantique constitue aujourd'hui un enjeu essentiel dans le domaine du patrimoine culturel, où la connaissance est toujours le résultat d'une combinaison de compétences et de profils disciplinaires complémentaires. À ce titre, elle propose de considérer l'objet réel – à travers son clone numérique – comme le seul dénominateur commun capable de fédérer les ressources et les connaissances des communautés qui l'étudient (Figure 85).

Cette approche, qui passe par la fusion des technologies d'acquisition et de visualisation avec les méthodologies d'analyse et d'interprétation des données, permet d'envisager la production de masses de contenus sémantiques plutôt que de masses de données brutes non interprétées. Néanmoins, au regard

des enjeux concernant le relevé d'art pariétal, Aïoli présente encore certaines limites notables (section 3.1.7).

Il semble légitime de penser que la reconquête de l'engagement analytique de l'observateur ne passe pas que par la spontanéité de son geste, mais aussi par la relation qui le lie à son objet d'étude et qui lui permet, dans le cas d'un relevé traditionnel, de mettre en relation ses connaissances avec ses perceptions.

Pourtant, bien qu'Aïoli offre une certaine continuité entre la capture du réel et l'enrichissement sémantique, les données relatives à un objet d'étude demeurent **un ensemble décontextualisé** : la relation projective entre les ressources 2D et 3D est bien établie, les données 3D sont bien basées sur la réalité, mais **le lien entre les données spatialisées et l'objet réel demeure absent**. De fait, même si le nuage de points exploité constitue bien une représentation fidèle de l'objet d'étude – à un niveau de détails donné –, le moment de *description* demeure un processus déporté *ex situ* auprès d'une copie numérique qui ne peut rendre compte de toute sa complexité.

4.4. Replacer l'objet réel au cœur du processus d'interprétation

Pour ce travail, nous avons posé précédemment l'hypothèse que l'engagement cognitif et analytique de l'observateur passe par la fusion des trois moments du relevé au sein d'un moment unique face à l'objet réel, comme c'est le cas pour le relevé *traditionnel* (section 2.5).

Un enjeu majeur est alors de **faire converger les pratiques du relevé *classique* et du relevé numérique**, afin d'autoriser la fusion des aspects géométriques, visuels et sémantiques au sein d'une approche intégrée. Comme nous l'avons vu, cela implique non seulement de faire converger les approches des systèmes d'information centrés sur les connaissances et celles centrées sur la 3D, comme le propose déjà efficacement Aïoli, mais cela sous-entend également d'intégrer des problématiques liées à l'hybridation réel – virtuel pour autoriser la contextualisation des données numériques sur la réalité et faciliter le processus d'analyse.

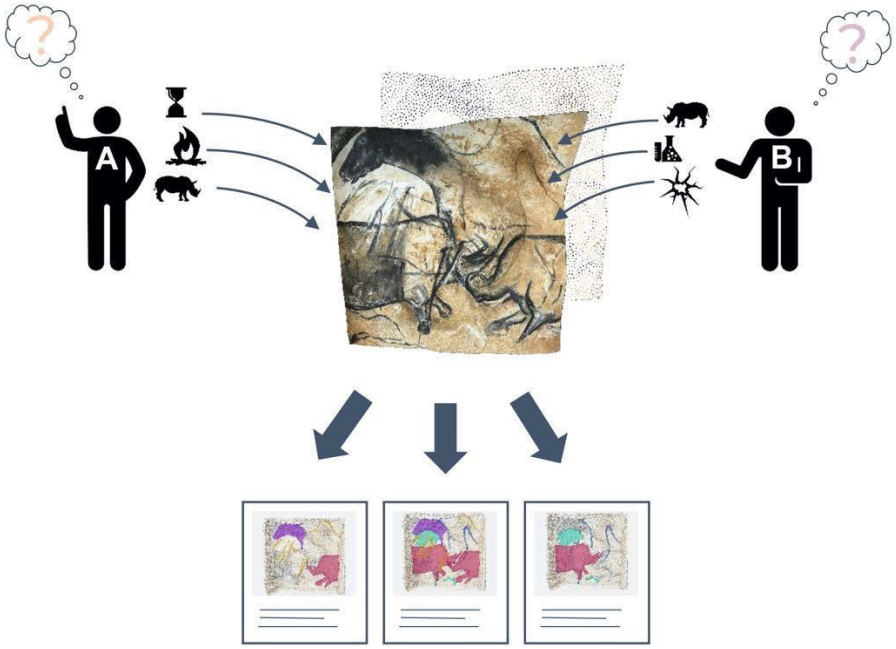


Figure 86 : Schéma de notre approche : un système d'information dédié au relevé d'art pariétal centré sur l'objet réel, permettant de spatialiser des observations. Le nuage de points 3D permet de fédérer les ressources liées à l'objet d'étude, et joue le rôle d'intermédiaire entre ces ressources et l'objet réel. Illustration : auteur.

Pour ce faire, nous nous inscrivons dans la continuité du travail mené par le laboratoire MAP dans le cadre du développement d'Aïoli. Notre ambition est de proposer un système d'information adapté au relevé d'art pariétal. Notre objectif est d'étendre à l'objet réel la relation projective établie entre les différentes ressources 2D et 3D relatives à un objet d'étude (Figure 86).

Concrètement, il s'agit de faire converger ce qui relève de la saisie de données spatiales avec ce qui relève de l'interprétation de l'objet dans un environnement unique. Comme nous l'avons vu précédemment (section 3.1.7), le nuage de points 3D peut être exploité pour fédérer les différentes ressources liées à un objet d'étude. Considérant la capacité des technologies actuelles à faire coexister des contenus réels et virtuels, notamment à travers la notion de *persistance*, notre approche propose d'utiliser ce même nuage de points comme légitime intermédiaire entre l'objet réel et les données associées, c'est-à-dire aussi bien les données issues de procédés d'acquisition 2D ou 3D, que les annotations sémantiques issues de leur interprétation (Figure 87).

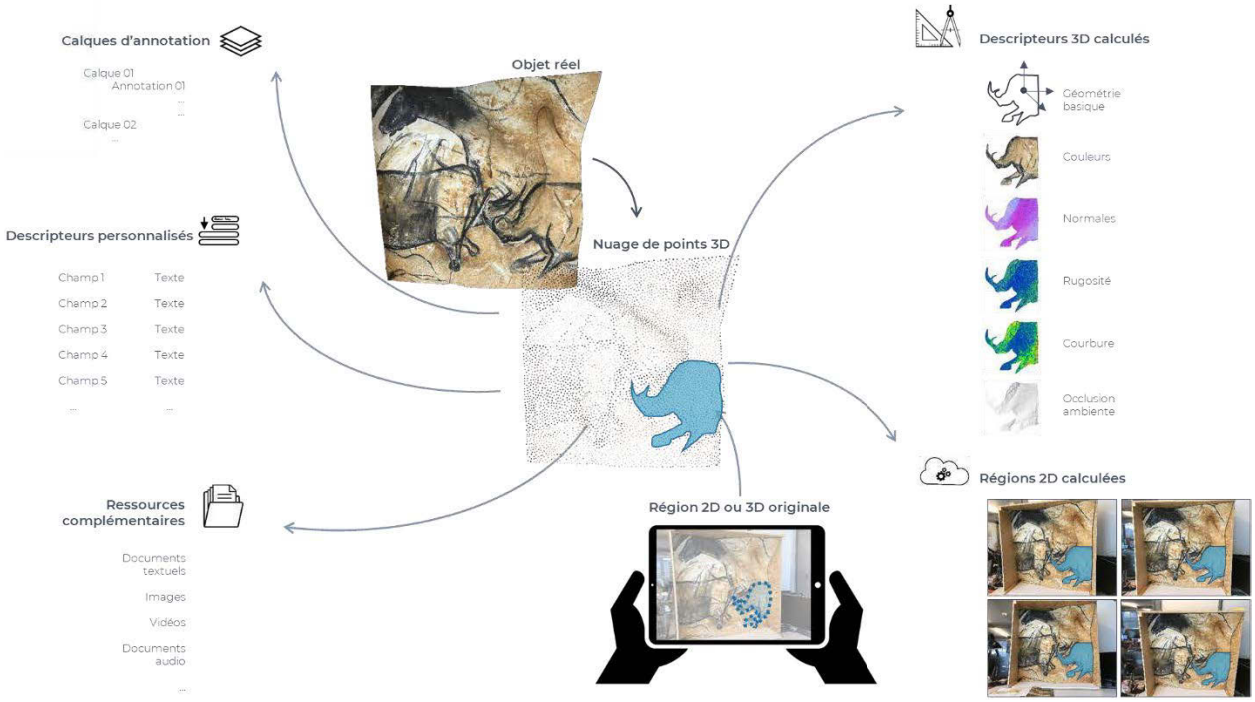


Figure 87 : Schéma de notre approche. Le nuage de points agit comme intermédiaire entre l'objet réel et les données spatialisées associées, elles-mêmes liées à des descripteurs géométriques et sémantiques. Illustration : auteur.

Nous avons vu dans le Chapitre 3 que la réalité mixte, et en particulier la réalité augmentée, constituait pour cela un levier majeur permettant de surmonter les verrous soulevés par Aïoli en termes d'interactions et d'hybridation. Ce chapitre nous a permis de mieux cerner les enjeux actuels de la réalité mixte et de comprendre les différentes pistes qui s'offrent à nous en termes de technologies et de méthodes.

Concernant les dispositifs d'affichage, nous avons pu voir qu'il était souhaitable de privilégier dans un premier temps les dispositifs autonomes, mobiles ou HMD *optical see-through*, avec une préférence pour les tablettes, qui sont déjà actuellement utilisées sur sites lors de relevés analytiques. Nous avons exclu les approches de suivi et recalage nécessitant le contrôle ou la modification de l'environnement, au profit d'approches basées vision, en particulier les approches hybrides basées sur le SLAM. En termes de développement applicatif, nous avons pu voir ici encore qu'il existait plusieurs pistes, parmi

lesquelles l'ébauche de standard WebXR qui semble présenter des intérêts majeurs vis-à-vis de notre contexte de travail. Associée à l'API WebGL, elle permet de développer des applications web multiplateformes dotées de contenus 3D visualisables en réalité augmentée.

À la lumière des différents états de l'art, nous proposons donc de répondre aux enjeux actuels du relevé d'art pariétal en définissant un environnement numérique dédié, se présentant sous la forme d'une application web multimodale. Cette application tient Aïoli comme base de départ (architecture logicielle, structuration des données, création de projets, propagation d'annotation, fonctionnalités collaboratives, etc.), étendu de trois principes fondamentaux interreliés.

- **Fédérer :** L'application ne doit pas seulement centraliser l'ensemble des données, informations, et connaissances relatives à l'objet d'étude, elle doit aussi les mettre en relation spatialement au sein d'un espace de visualisation unique temps réel autorisant leur manipulation. Compte tenu de la complexité des sites étudiés, les annotations doivent être réalisables quelle que soit la donnée d'entrée, 2D ou 3D. L'hypothèse sous-jacente est que l'approche proposée par Aïoli est généralisable à l'ensemble des données impliquées dans l'étude de l'art pariétal. ;

- **Hybrider :** Les données virtuelles doivent être accessibles *in situ* en réalité augmentée afin de permettre aux utilisateurs de les visualiser de manière contextualisée et d'interagir avec un environnement hybride comprenant l'objet réel et les informations numériques associées ;

- **Augmenter :** L'application doit permettre d'exploiter les attributs géométriques, visuels, et sémantiques à disposition pour faciliter le processus d'analyse et de suivi et enrichir la compréhension de l'objet. Cela passe par la définition de modalités de rendu facilitant le processus d'annotation et de description sémantique. Nous avons vu en (section 2.3) que certains procédés permettaient de faciliter l'observation de stigmates. Il s'agit donc de s'appuyer sur l'hybridation réel/virtuel pour augmenter l'environnement des releveurs par les données, informations et connaissances existantes ou calculées.

Si la spatialisation et l'augmentation constituent davantage des verrous méthodologiques que technologiques, l'hybridation représente un défi complexe qui soulève des questions à la fois technologiques et méthodologiques.

Pour atteindre ces objectifs, les problèmes à traiter sont les suivants :

1. Développement d'un *viewer* hybride orienté web, permettant de charger les données hétérogènes 2D et 3D impliquées dans un relevé numérique (sections 2.2 et 2.3) au sein d'un espace de représentation unique, dynamique et cohérent (section 5.2).

Ce point est fondamental car, comme nous l'avons vu précédemment, le relevé d'art pariétal nécessite le croisement de différentes ressources 2D et 3D dont la mise en relation selon des critères spatiaux constitue un préalable indispensable pour permettre la mise en relation selon des critères sémantiques. Il s'agit également d'un préalable indispensable pour permettre l'alignement réel/virtuel des ressources liées à l'objet d'étude dans le cas d'une utilisation en réalité augmentée. ***Pour ce faire, nous proposons de nous baser sur l'API WebGL pour l'affichage de données 3D dans le navigateur ainsi que sur les informations issues des calculs de spatialisation pour leur chargement et leur positionnement.***

2. Développement d'une méthode de recalage basé vision d'une scène 3D sur la réalité, indépendante du contenu manipulé, et avec pour seules données de départ les informations communiquées par le WebXR, c'est-à-dire les poses des caméras du mobile exprimées dans un système de coordonnées arbitraire différent de celui de la scène (section 5.3).

Il s'agit d'un élément clé pour permettre l'articulation de modalités de visualisation et d'interaction *desktop* destinées aux moments d'études hors site avec des modalités en réalité augmentée lors des moments sur site. Or, nous avons pu constater à partir de l'état de l'art sur les méthodes d'hybridation que le recalage constitue un verrou majeur. Nous avons également identifié les méthodes de suivi basées sur l'estimation de poses d'images clés comme un levier susceptible de converger avec les approches basées sur la photogrammétrie actuellement exploitées par Aïoli pour la spatialisation de ressources autour du nuage de points fédérateur. ***Notre approche consiste à proposer une méthode de recalage basée sur un calcul de résection spatiale appliqué à une image clé issue du suivi.***

3. Proposition d'outils d'assistance à la visualisation permettant de faciliter l'observation des traces anthropiques difficilement perceptibles (gravures et peintures détériorées, reliefs, etc.), mais aussi d'outils adaptés à l'annotation temps-réel des ressources 2D et 3D permettant de profiter de leurs apports respectifs (Chapitre 6).

Comme nous avons pu le constater dans les chapitres précédents, les grottes ornées sont des sites complexes. Les vestiges et traces anthropiques sont profondément liés aux caractéristiques géométriques et visuelles des parois, et souvent difficiles à apprécier à l'œil nu. Le Chapitre 2 montre bien la nécessité de surmonter les fossés sensoriels et sémantiques des données numériques par le biais de traitements permettant de faciliter le processus d'analyse. À l'inverse, les potentiels de ces mêmes données en termes de caractérisation géométrico-visuelle peuvent bénéficier aux activités d'interprétation *in situ* en améliorant les capacités perceptives des observateurs, par l'adjonction d'informations issues de calculs. ***Notre approche consiste donc à explorer différents traitements favorisant l'interprétation tout en proposant des outils d'interaction – et en particulier d'annotation – permettant aux utilisateurs de formaliser sémantiquement le résultat de leurs observations.***

Dans la suite de ce travail, nous nous attachons à apporter des réponses à ces trois problèmes par le biais d'expérimentations ayant mené au développement d'une preuve de concept. Nous détaillons dans les chapitres qui suivent nos approches ainsi que les résultats obtenus.

Partie II

Chapitre 5. Spatialisation et recalage des données numériques de relevé

5.1. Point de départ de notre travail : état actuel de la plateforme Aïoli

Comme nous l'avons vu précédemment, Aïoli est une plateforme web dédiée à l'annotation sémantique d'objets patrimoniaux.

À la base se trouvent les objets patrimoniaux (Figure 88). Ils correspondent à des entités physiques relevant du domaine du patrimoine culturel, par exemple des monuments, des sites archéologiques, ou des éléments mobiliers. Concrètement, il s'agit des objets d'étude sur lesquels interviennent les acteurs. Chaque projet créé sur la plateforme est lié à un objet patrimonial dans la mesure où il en représente nécessairement tout ou partie. Le concept d'objet patrimonial permet donc dans une certaine mesure de fédérer les acteurs s'intéressant à un même objet d'étude, même lorsqu'ils n'interviennent pas sur les mêmes éléments de l'objet. Par exemple, dans notre cas, un objet patrimonial pourrait être une grotte ornée, et comprendrait différents projets, correspondant chacun à un panneau spécifique. Chaque projet comprend des calques, qui correspondent à des niveaux d'observation, comme pour le relevé analytique technique. Ces calques sont associés à des descripteurs sémantiques qui les qualifient dans leur globalité. Les calques comprennent des régions, issues des annotations réalisées par les utilisateurs. Chaque région hérite des descriptions du calque parent, et est également associée à des descripteurs sémantiques spécifiques.

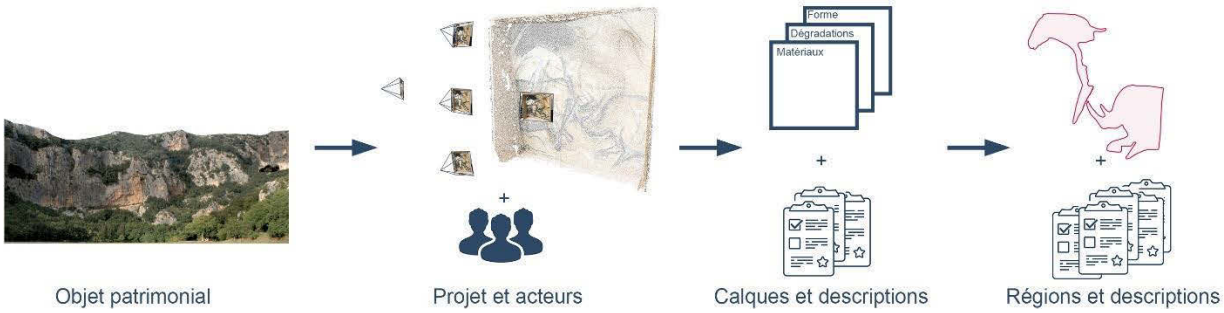


Figure 88 : Hiérarchie proposée par Aïoli. Illustration : auteur.

Pour créer un projet, l'utilisateur téléverse un jeu de photographies de son objet d'étude. Ces photographies sont traitées par le moteur de corrélation, et donnent lieu à un nuage de points dense autour duquel sont spatialisées les images. L'utilisateur peut alors annoter n'importe quelle image d'entrée, et visualiser, à l'issue du processus de propagation automatique, cette même annotation sur toutes les ressources 2D ou 3D sur lesquelles la région est visible.

Pour cela, la plateforme implique des processus variés, reposant notamment sur la photogrammétrie, le traitement d'images, ou encore l'indexation. Elle fait également intervenir des données hétérogènes. Pour cette raison, l'architecture de l'application a été pensée pour offrir une certaine flexibilité en isolant les différences processus, ce qui permet de les mobiliser indépendamment selon les besoins opérationnels.

5.1.1. Architecture applicative

L'architecture applicative comprend quatre éléments principaux (Figure 89) :

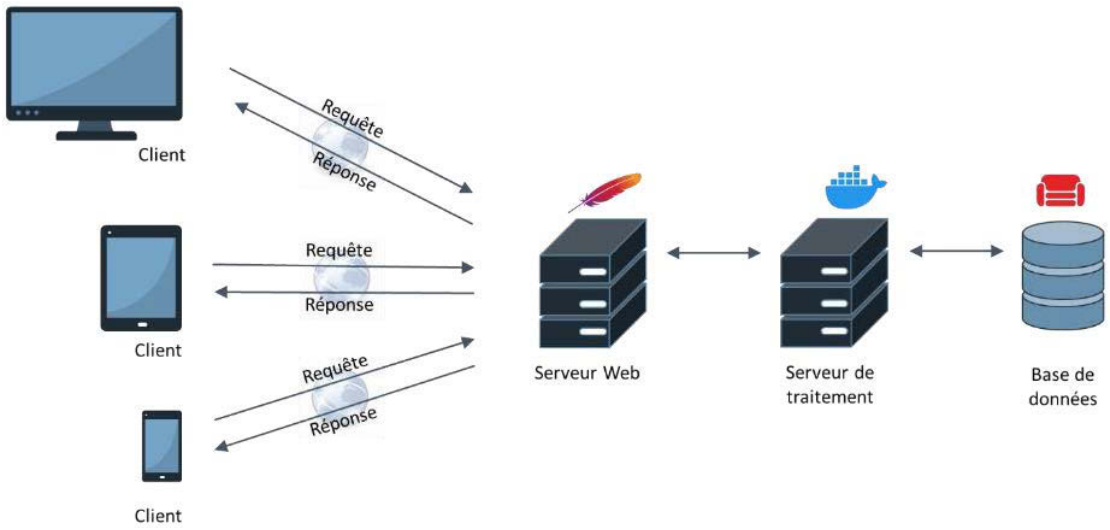


Figure 89 : Architecture applicative de la plateforme Aïoli. Illustration : auteur.

- Le **client léger**, en HTML et JavaScript, fonctionnant à partir d'un navigateur. Il communique avec le serveur web pour accéder aux services de l'application. Il comprend principalement les interfaces distinctes dédiées à la visualisation 2D et 3D, ainsi que les outils d'annotation sémantique.
- Le **serveur web**, qui constitue le nœud central de l'application et gère notamment les données associées aux projets. Tous les fichiers sont conservés sur ce serveur, par exemple les nuages de points produits par photogrammétrie, les jeux de photographies, ou encore les images produites à l'issue de l'indexation. Il communique avec le serveur de traitement qui contient les processus de calcul.
- Le **serveur de traitement**, qui contient tous les processus de calcul. Il s'appuie sur une technologie de virtualisation à base de conteneurs pour fournir des conteneurs distincts spécifiquement dédiés à chaque étape du processus de traitement (processus de spatialisation par photogrammétrie, conversion des données résultantes, indexation des données, propagation d'annotation, etc.).
- La **base de données orientée documents**, conçue pour stocker et interroger des données sous forme de documents au format JSON. Elle comprend toutes les données de l'application (utilisateurs, objets patrimoniaux, projets, calques, régions, etc.).

5.1.2. Données produites par la plateforme

Pour spatialiser dans un *viewer* unique les ressources impliquées dans un projet, il faut d'abord comprendre la nature des données générées par les différents modules de la plateforme. Lorsqu'un projet est créé à partir d'un jeu de photographies, le premier processus impliqué est celui du calcul photogrammétrique, basé sur le logiciel *open-source* MicMac (Rupnik et al., 2017). Il vise à orienter les images et créer un nuage de points dense de l'objet d'étude.

5.1.2.1. Fichiers d'orientation d'images et nuages de points globaux

Les paramètres intrinsèques (ou internes) expriment les propriétés projectives de l'appareil photographique utilisé pour les prises de vues. Comme nous l'avons vu précédemment, ces paramètres comprennent principalement la distance focale et les coordonnées du point principal, mais aussi les paramètres de distorsion pour la prise en compte des erreurs dues aux imperfections de la lentille (section 3.3.6.2.1). Les paramètres extrinsèques (ou externes) expriment la position et l'orientation de l'appareil photographique au moment de la prise de vue, dans le repère arbitraire défini au début du calcul et à un facteur d'échelle près. La position est définie par une translation dans l'espace selon les axes X, Y, et Z. L'orientation est définie par la rotation du plan image dans l'espace, exprimée à partir des angles θ, ψ, ω autour des trois axes. Ces paramètres, calculés par le module TAPAS (intégré au module APERO), sont stockés dans des fichiers XML :

- **Un fichier pour les données de calibration de chaque caméra** (un fichier unique dans le cas où toutes les images du projet partagent les mêmes paramètres internes), et
- **Un fichier par image pour les paramètres externes**.

Outre ces fichiers, le processus de calcul photogrammétrique produit également deux fichiers au format PLY représentant l'objet d'étude dans sa globalité :

- **Un nuage de points épars** (points SIFT)
- **Un nuage dense.**

5.1.2.2. Images maîtresses et images de caractérisation morphologique

Comme nous l'avons vu au Chapitre 2, les données 2D et 3D peuvent être exploitées pour extraire de nouvelles informations, principalement d'ordre géométrique ou visuel, et parfois physico-chimique selon les capteurs utilisés. Ces informations peuvent révéler efficacement un certain nombre d'aspects qui ne sont pas aisément perceptibles sur les données brutes. En particulier, le nuage de points contient, en plus des attributs colorimétriques de chaque point, les coordonnées 3D et les directions des normales. Ces données peuvent être utilisées pour caractériser la courbure de la surface de l'objet, sa rugosité, l'occlusion ambiante, mais aussi pour créer des cartes de profondeur selon des plans spécifiques, ou encore des masques caractérisant la présence ou non de points 3D (Annexe 2).

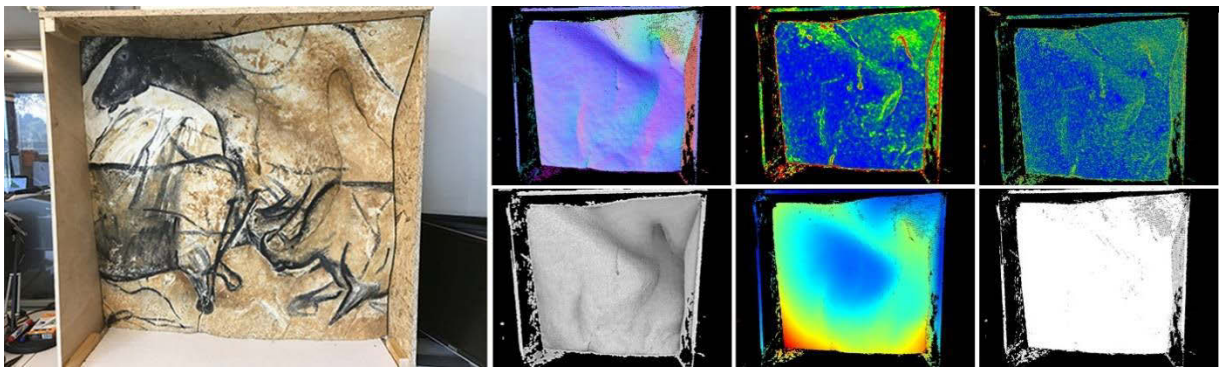


Figure 90 : Une image maîtresse et les images de caractérisation morphologique produites par reprojection 3D → 2D.

Pour aider le processus d'annotation sémantique en guidant l'utilisateur dans son analyse, la version actuelle de la plateforme Aïoli propose ainsi différents outils de caractérisation morphologique, automatiquement exécutés lors de l'indexation d'une image. La relation projective 2D → 3D établie par le calcul est utilisée pour exprimer toutes ces nouvelles informations comme les couches additionnelles des photographies maîtresses (Figure 90). À l'issue de la procédure d'indexation, nous disposons donc pour chaque photographie du jeu de données :

- D'**une photographie maîtresse**, celle utilisée pour la construction 3D du nuage de points dense par photogrammétrie ;
- D'**une série d'images complémentaires** caractérisant la morphologie du nuage de points issues de projections 3D → 2D (courbures, occlusion ambiante, rugosité, normales, cartes de profondeur, position des indices).

5.1.2.3. Fichiers d'annotation 2D et 3D

Lorsqu'une annotation est réalisée sur une image d'entrée, la sélection est envoyée au serveur de traitement sous la forme d'un tableau ordonné de coordonnées UV qui correspondent aux contours de la sélection dans le cas d'une région, ou au tracé d'une polyligne sinon (Figure 91 ①). La zone ainsi définie sur l'image est utilisée pour rechercher les coordonnées 3D correspondantes sur le nuage de points grâce à la relation 2D → 3D établie lors de l'indexation (Figure 91 ②). Les coordonnées des points 3D concernés sont stockées dans un fichier PLY, le nuage de points dense de l'annotation. Pour achever la propagation, la sélection 3D est ensuite projetée sur les autres images du projet à l'aide de la relation 3D → 2D (Figure 91 ③).

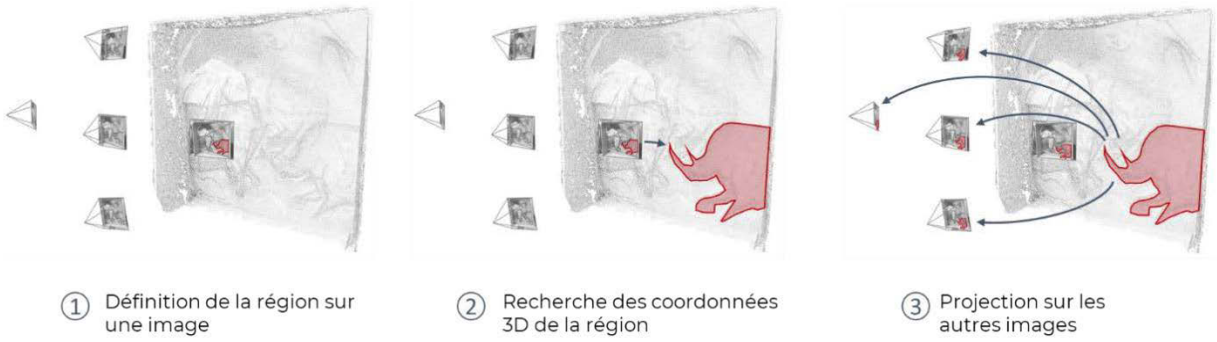


Figure 91 : Principe du processus de propagation d'annotations. Illustration : auteur, d'après (Manuel, 2016).

Un algorithme d'extraction de contours est appliqué sur les régions 2D extraites, et fournit un tableau ordonné de coordonnées pixel des points de contours. En l'état actuel, ces régions sont stockées sous la forme de fichiers vectoriels au format SVG, cependant cette étape impose un temps de traitement supplémentaire. À l'issue de la propagation d'une annotation, nous disposons donc de deux catégories d'éléments :

- **Le nuage de points 3D** correspondant à la région annotée ;
- **Une série de fichiers vectoriels** contenant, pour chaque image sur laquelle région est visible, la représentation géométrique 2D de l'annotation.

5.2. Méthode proposée pour la mise en relation de données spatialisées : un *viewer* hybride 2D/3D

Comme nous l'avons énoncé en Chapitre 4, notre approche se concentre dans un premier temps sur la partie client de l'architecture de la plateforme. Il s'agit d'envisager l'ensemble des ressources manipulées lors du relevé pariétal comme les *N* éléments d'un ensemble unique comprenant **un seul système de coordonnées**, et au sein duquel toutes les données peuvent être représentées de manière cohérente relativement au nuage de points. De cette manière, la compréhension de l'objet d'étude passe par la mise en relation spatiale des différentes ressources produites par les acteurs, et ce fonctionnement réduit la problématique du recalage AR des *N* ressources indépendantes à celui d'une unique entité 3D cohérente.

Cet objectif repose donc avant tout sur notre capacité à unifier les contenus 2D et 3D au sein d'un même espace de représentation, que ce soit pour la visualisation ou l'interaction. Pour ce faire, la première étape de notre travail consiste à traduire littéralement les interrelations entre les ressources 2D et 3D dans un *viewer* dynamique en spatialisant les différentes ressources iconographiques relatives à un nuage de points dans une scène unique. Comme nous avons pu le constater dans les sections précédentes, les données produites et manipulées au sein d'un projet Aïoli relèvent de trois catégories principales :

- **Les nuages de points**, comprenant les nuages de points principaux du projet (épars et dense) et les nuages de points des régions 3D, stockées sur le serveur web au format PLY ;
- **Les images 2D** comprenant les photographies maîtresses du projet et les images auxiliaires calculées lors de l'indexation. Stockées sur le serveur web dans des formats d'image communs (JPG, PNG, etc.), elles sont accompagnées de fichiers XML spécifiant les paramètres internes et externes des caméras correspondantes ;
- **Les annotations 2D** sous la forme de tracés vectoriels contenant les coordonnées image des points de contours.

En termes de visualisation 3D, ces deux dernières catégories correspondent toutes deux à des objets 3D dotés d'une géométrie solide et d'un matériau, au contraire des nuages de points qui se traitent de manières distinctes, habituellement sous la forme de particules ou bien, comme dans notre cas, d'octree (Chapitre 7).

5.2.1. Structure proposée pour spatialiser les données dans une même scène

Pour aborder ce travail, nous considérons que tous les éléments précités peuvent être chargés dans une scène 3D unique, pour laquelle nous proposons plusieurs niveaux hiérarchiques (Figure 92). Au sein de cette structure, chaque sous-groupe agit en quelque sorte comme une scène locale indépendante.

Les ressources à spatialiser y sont d'abord réparties selon leur nature en termes de visualisation : objets solides dans $Scene_{Objects}$, nuages de points dans $Scene_{Pointclouds}$, ou sources de lumière dans $Scene_{Lights}$. Nous proposons de distinguer en second lieu les ressources 2D des ressources 3D, afin de faciliter le passage d'un mode de projection à un autre, et incidemment, de proposer deux modalités d'interactions distinctes, tout en préservant la cohérence de la scène globale. Enfin, pour chaque ressource susceptible

de supporter des annotations, nous proposons des groupes reproduisant les niveaux hiérarchiques des calques de descriptions sémantiques, destinés à accueillir les régions.

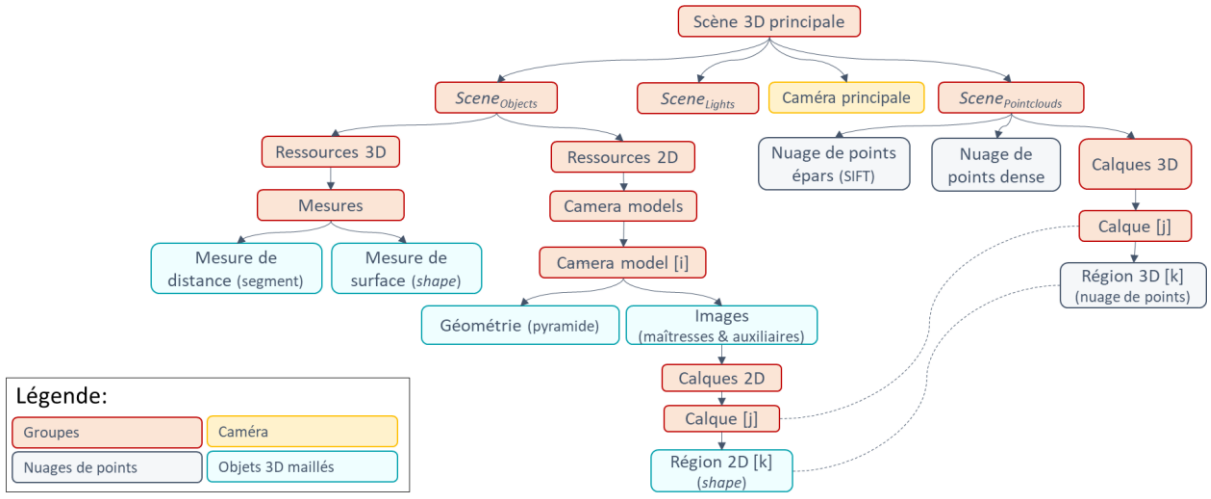


Figure 92 : Structure proposée pour le *viewer* hybride regroupant les ressources 2D et 3D au sein d'une même scène. Illustration : auteur.

Les niveaux hiérarchiques les plus élevés concernent donc le projet dans sa globalité. À l'inverse, les niveaux les plus bas s'intéressent à des éléments spécifiques de l'objet d'étude. Par exemple, le groupe $Scene_{Pointclouds}$ illustre bien ce principe. Il comprend trois éléments : les deux nuages de points de référence du projet, et un groupe contenant l'ensemble des calques, chacun étant structuré comme un groupe enfant indépendant contenant les nuages de points denses des régions 3D issues d'annotations.

Figure 93 : Une implémentation du *viewer* selon la structure annoncée. Le projet comprend un calque « figures » doté d'une annotation, visible en 3D sous forme de nuage de points, et en 2D sur chaque vue maîtresse.

Concernant les autres données du projet, les caméras utilisées pour produire les photographies maîtresses peuvent être modélisées par le modèle sténopé. En termes de visualisation, cela revient à les représenter dans la scène comme des pyramides orientées dont le sommet correspond au centre optique, la hauteur à la distance focale, et la base à la photographie elle-même (section 5.2.2.1). La structure que nous proposons pour ce *viewer* comprend donc dans la scène des objets $Scene_{Objects}$ un groupe enfant destiné à contenir chacune de ces pyramides de vue, elles-mêmes comprenant deux éléments : la géométrie du

modèle d'une part, et d'autre part les images associées au plan image (photographie maîtresse et images auxiliaires de caractérisation morphologique). Or, chaque image se comporte comme un potentiel support d'annotation. Pour cette raison, chaque plan image contient un groupe enfant contenant les calques 2D distincts. Chaque calque accueille des régions 2D correspondant à la reprojection sur l'image des annotations propagées, représentées sous forme de *shapes* (Figure 93).

5.2.2. Spatialisation des données 2D autour des nuages de points de référence

Comme nous l'avons souligné précédemment, les données produites par Aïoli à l'issue du calcul photogrammétrique ne sont pas géoréférencées. Les informations concernant les paramètres internes et externes des images maîtresses sont donc exprimées dans un repère arbitraire relatif au nuage de points.

En termes de processus il apparaît logique que ce dernier, qui joue le rôle de fédérateur, soit chargé en premier. Ensuite, il s'agit de produire les modèles de caméras et les positionner dans l'espace à partir des fichiers de calibration et d'orientation, et enfin, de charger le reste des ressources de la scène telles que les annotations. Les aspects techniques du chargement et rendu des nuages de points seront abordés dans le Chapitre 7 concernant l'implémentation informatique de notre approche.

5.2.2.1. Modélisation des pyramides de vue

Pour chaque image maîtresse, nous cherchons à produire un modèle 3D cohérent caractérisant ses propriétés projectives. Pour cela, nous pouvons exploiter les fichiers de calibration (Figure 94) produits lors des calculs photogrammétriques En effet, chaque fichier comprend plusieurs informations importantes pour modéliser l'appareil photographique utilisé.

```
<KnownConv>eConvApero_DistM2C</KnownConv>
<PP>1606.19888491608003 1240.18470048691256</PP>
<F>2807.02197163148185</F>
<SzIm>3264 2448</SzIm>
<CalibDistortion>
    <ModPhgrStd>
        <RadialePart>
            <CDist>1606.19888491608003 1240.18470048691256</CDist>
            <CoeffDist>5.67495972620652764e-09</CoeffDist>
            <CoeffDist>1.33633697815212024e-15</CoeffDist>
            <CoeffDist>-4.66542651700065114e-22</CoeffDist>
            <PPaEqPPs>true</PPaEqPPs>
        </RadialePart>
        <P1>-4.83830669739033967e-09</P1>
        <P2>-7.37232202513183631e-07</P2>
        <b1>-0.00185426453805267945</b1>
        <b2>0.0119783501159214847</b2>
    </ModPhgrStd>
</CalibDistortion>
```

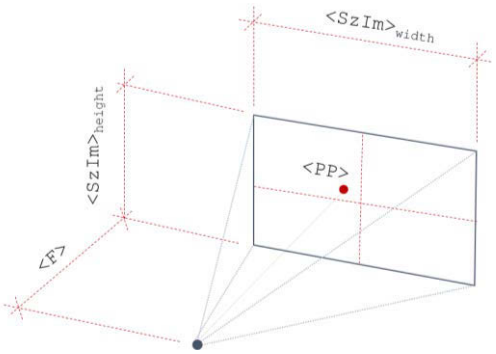


Figure 94 : Gauche : Extrait d'un fichier XML de calibration. Droite : Modélisation de l'appareil photographique selon le modèle sténopé, avec les paramètres du fichier de calibration.

Les balises <F> et <SzIm> donnent respectivement la focale et la taille de l'image, exprimées en pixels. Notons que les FOV horizontaux et verticaux peuvent être déduits à partir de la taille de l'image et de la distance focale. La balise <PP> donne les paramètres de la position du point principal de la prise de vue, exprimée en pixels dans l'image. Notons que même si de nombreux cas autorisent cette approximation, le point principal ne coïncide pas toujours avec le centre de l'image. Au niveau du modèle géométrique de la pyramide de vue, ce décalage du point principal se manifeste par un cisaillement que nous pouvons reproduire à l'aide d'une matrice 4 x 4 (Rodger, 1998) :

$$\begin{bmatrix} 1 & -\frac{PP_x - \frac{W}{2}}{F} & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -\frac{PP_y - \frac{H}{2}}{F} & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Où :

PP_x et PP_y correspondent respectivement aux coordonnées pixel du point principal telles qu'exprimées dans le fichier de calibration ;

W et H correspondent respectivement à la largeur et hauteur de l'image en pixels, telles qu'exprimées dans le fichier de calibration ;

F correspond à la distance focale.

Enfin, la balise `<CalibDistortion>` indique les paramètres de distorsions, qui dépendent du modèle utilisé, le modèle radial étant le plus courant. Il consiste à exprimer la distorsion à partir d'un centre `<CDist>` (qui coïncide avec le point principal), et de trois coefficients `<CoeffDist>`.

Pour tenir compte des distorsions des caméras dans le *viewer* afin de garantir la cohérence des points 3D avec les points des images, nous pouvons adopter deux approches. La première consiste à rendre l'ensemble de la scène 3D avec des distorsions similaires à celles présentées par les caméras, par exemple en utilisant des *shaders*. La seconde consiste à rectifier les images avant de les charger dans la scène, pour revenir à des modèles sténopés simples. C'est la démarche que nous proposons d'adopter, car elle simplifie considérablement la maîtrise temps réel des relations projectives entre les éléments affichés (section 6.2).

5.2.2.1. Orientation des pyramides de vue

Une fois la pyramide de vue modélisée en 3D, il reste à la spatialiser de manière cohérente dans la scène. Pour cela, nous pouvons exploiter le fichier XML d'orientation (Figure 95), qui comprend les paramètres externes de l'appareil photographique utilisé.

```
<TypeProj>eProjStenope</TypeProj>
<ZoneUtileInPixel>true</ZoneUtileInPixel>
<FileInterne>Ori-All/AutoCal_Foc-4200_Cam-iPhone_6.xml</FileInterne>
<RelativeNameFI>true</RelativeNameFI>
<Externe>
    <AltiSol>-18.0127720056169096</AltiSol>
    <Profondeur>15.5681640686412006</Profondeur>
    <Time>-1.0000000000000002e+30</Time>
    <KnownConv>eConvApero_DistM2C</KnownConv>
    <Centre>-1.59714936256893392 -0.115454028929885141 -1.64094439960836258</Centre>
    <IncCentre>1 1 1</IncCentre>
    <ParamRotation>
        <CodageMatr>
            <L1>0.998376395014286988 0.0315504306582554822 0.0474251431579814711</L1>
            <L2>0.031561594636618194 -0.999501676819557816 0.00051359406901777904</L2>
            <L3>0.0474177142244353381 0.00098405293095414668 -0.998874662816891634</L3>
        </CodageMatr>
    </ParamRotation>
</Externe>
```

Figure 95 : Extrait d'un fichier XML d'orientation produit par le module TAPAS lors du calcul photogrammétrique.

Trois paramètres nous intéressent en particulier. La balise <Centre> indique les coordonnées 3D du centre optique de la caméra. Cette information est nécessaire pour positionner la vue dans la scène. Le paramètre <CodageMatr> inclus dans <ParamRotation> correspond à sa pose dans l'espace. Celle-ci est exprimée sous la forme d'une matrice 3x3, divisée en trois lignes <L1>, <L2>, et <L3>. Les trois colonnes de la matrice ainsi reconstituée définissent trois vecteurs O_x, O_y, O_z. Les deux premiers, O_x et O_y, définissent le plan image. Le vecteur Oz indique quant à lui la direction de la caméra (Figure 96).

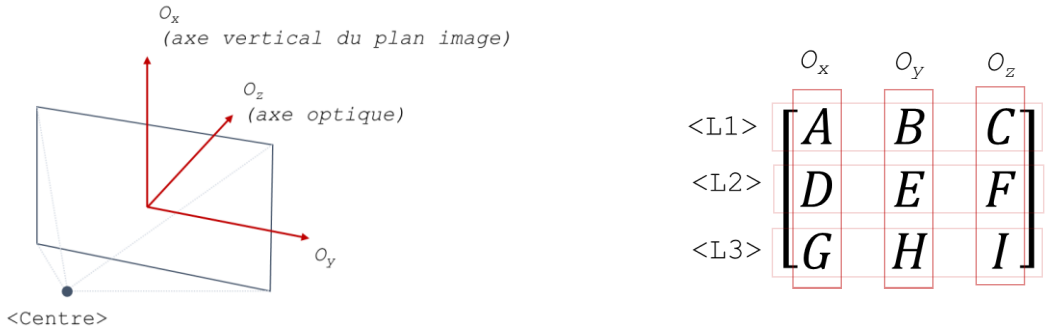


Figure 96 : Orientation de la caméra selon les paramètres extrinsèques calculés par TAPAS. Illustration : auteur.

5.2.2.2. Affichage des photographies maîtresses et auxiliaires dans le plan image

En termes de visualisation, les images peuvent être considérées simplement comme les textures des plans images des modèles géométriques des caméras correspondantes. Concrètement, cela signifie que pour afficher les images d'un projet dans le *viewer*, nous pouvons ajouter à chaque pyramide de vue une géométrie plane coïncidant avec le plan image, destinée à accueillir une ou plusieurs textures selon les données associées à ce point de vue (Figure 97).

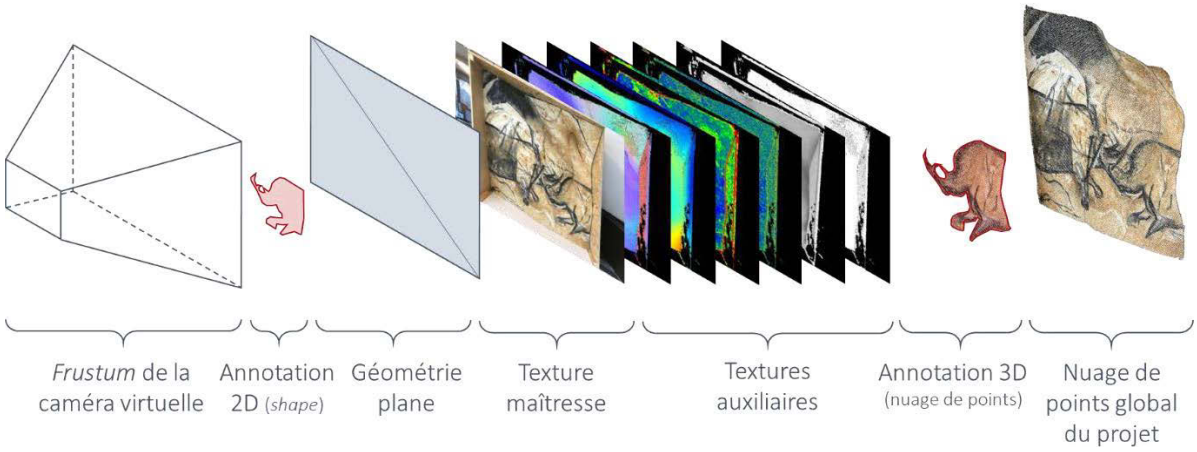


Figure 97 : Structure proposée de modèle de caméra à textures multiples permettant de visualiser les images maîtresses, les cartes de caractérisation morphologiques, et les annotations 2D. Illustration : auteur.

L'image maîtresse est alors considérée comme la *texture maîtresse* du modèle de caméra, et les images complémentaires (par exemple les images de caractérisation morphologique calculées lors de l'indexation) sont considérées comme des *textures auxiliaires*.

5.2.2.3. Chargement d'annotations 2D

À l'issue de la propagation d'une annotation, nous disposons de deux éléments : le nuage de points 3D correspondant à la région annotée (dont le chargement est similaire à celui des nuages de points principaux du projet), et pour chaque image sur laquelle la région est visible une liste ordonnée de points de contours 2D. Dans Aïoli, celle-ci est stockée sous forme de fichier vectoriel au format SVG, mais les coordonnées pixellaires de ces points suffisent à reconstruire les géométries des régions concernées sous forme de *shapes* (Figure 98, Annexes 6 et 7). En termes de positionnement dans la scène, ceux-ci se superposent au plan image de la caméra concernée. Il faut cependant tenir compte du facteur d'échelle utilisé pour la propagation, car la définition des images exploitées pour les calculs et celle des images utilisées pour la visualisation peuvent différer, selon des besoins d'optimisation. D'un point de vue pratique, cela ne représente pas d'enjeu particulier car ce facteur d'échelle est toujours connu.

Figure 98 : Construction d'une annotation sous forme de *shape* à partir des coordonnées pixellaires des points de contours. La structure du document JSON permet de tenir compte d'éventuelles discontinuités dans les contours en traçant des sous-régions.

Comme nous l'avons vu précédemment, les régions ne se limitent pas à des sélections tracées en 2D ou en 3D, elles font également l'objet de descriptions sémantiques explicitées par les différents acteurs du projet. Ainsi, elles disposent à la fois d'attributs géométriques, visuels, et sémantiques. Outre le fait de permettre la production d'un document synthétisant les observations du relevé, ce fonctionnement permet de réaliser des recoupements d'informations précieux pour la compréhension de l'objet d'étude, dans la lignée de l'approche de Messaoudi (2017) en matière de corrélation multidimensionnelle.

5.2.2.4. Vers la prise en charge de données 2D de lasergrammétrie

La prise en charge de données issues de lasergrammétrie (images panoramiques des stations) est plus délicate, car il n'existe pas de consensus concernant les modalités d'export ou de stockage des métadonnées, chaque constructeur proposant ses propres structures de fichiers, voire des formats spécifiques. Afin d'évaluer les possibilités et enjeux liés leur prise en charge, nous avons mené une expérience basée sur les données du relevé numérique de l'abri du Cap-Blanc, réalisé par l'entreprise Art Graphique et Patrimoine en 2012[36]. Elles comprennent les nuages de points et images

[36] Ces données ont été acquises avec un scanner laser Faro Focus 3D S 120.

équirectangulaires de chaque station, ainsi que le fichier global de travail[37]. Les stations sont déjà consolidées, c'est-à-dire spatialisées les unes par rapport aux autres.

Chaque station est représentée par une image panoramique, pour laquelle chaque pixel représente l'un des points mesurés par le laser lors de son balayage. Les dimensions de cette image dépendent donc du champ du balayage choisi par l'opérateur pour réaliser la station, 360° en horizontal, et 150° en vertical dans le cas des stations de ce jeu de données. Comme nous l'avons vu précédemment, les images équirectangulaires correspondent à des projections sphériques. Ces images peuvent être visualisées sans déformations sous la forme d'une texture appliquée à l'intérieur d'une sphère.

Figure 99 : Plaquage de la texture panoramique sur une sphère en tenant compte des limites du champ de balayage.

Dans le *viewer*, cela n'implique que la création d'un maillage sphérique et le chargement de l'image équirectangulaire pour le texturage. Au centre de la sphère doit être placée une caméra virtuelle, que l'utilisateur pourra faire pivoter pour naviguer. Cependant, étant donné que le champ vertical de balayage laser est réduit, la texture ne doit pas couvrir l'ensemble de la sphère. Dans notre cas, son hémisphère sud est tronqué d'un tiers[38]. Pour éviter d'impacter la géométrie, nous proposons simplement d'agir sur les coordonnées UV de la texture afin de laisser la zone non balayée exempte de texture (Figure 99).

Par ailleurs, les nuages de points issus de lasergrammétrie présentent la même structure que ceux produits par photogrammétrie et peuvent donc être chargés dans le *viewer* selon des méthodes similaires (voir Chapitre 7). Il reste donc uniquement à positionner les sphères texturées relativement au nuage de points affiché. Pour cela, à l'issue du processus de consolidation des stations, le logiciel FARO Scene permet l'export des paramètres de position et d'orientation des différentes stations sous forme d'un fichier texte (Figure 100).

```
"# iQscene Model Extract"
"iQLib" 6.2.4.30
References {
"References" -1549836.182622129 -4195723.327103934 -112.5352544959069
"GPS1" 8.987377871 -9.367103934 -1.245254496
"GPS2" 24.157377871 -16.127103934 -1.8652544959
}
Scans {
"Cap_blanc_int_000" -9.333122 1.443527 -0.269555 -0.0066724184 -0.00087758743 -0.99997735 147.0789 8.9.2012 9:58:43
"Cap_blanc_int_001" 3.184379 -0.71955 -0.374711 0.0015498522 -0.0026756514 0.99999522 145.16975 8.9.2012 10:16:17
"Cap_blanc_int_002" -10.518964 2.73784 -0.527799 0.0058647037 0.0021739779 -0.99998044 132.67443 8.9.2012 10:33:16
}
```

Figure 100 : Fichier texte issu de l'export des paramètres de position et d'orientation des stations depuis Faro Scene.

Chaque ligne de l'entrée `Scans` correspond à une station. La première colonne correspond à son nom, qui permet de retrouver le nuage de points correspondant. Les trois colonnes suivantes, séparées par des espaces, correspondent à la position de la station dans le repère global, c'est-à-dire à la position du centre des sphères. Les trois colonnes suivantes correspondent à la direction de l'axe d'orientation de la station

[37] Au format LSPROJ exploitable avec le logiciel Faro Scene.
[38] Pour accéder à ces informations de champ de balayage, nous avons dû passer par le logiciel Faro Scene, mais théoriquement, elles sont aussi communiquées par le prestataire au commanditaire en tant que métadonnées associées aux stations (Abergel et al., 2019). Nous pouvons donc considérer qu'un utilisateur souhaitant téléverser des données laser sur la plateforme Aïoli aurait la capacité de fournir les métadonnées en même temps que les fichiers.

par rapport au repère monde, et la quatrième à l'angle d'orientation autour de cet axe. Enfin, les dernières colonnes correspondent à la date et l'heure de l'acquisition. Ainsi, nous pouvons positionner les sphères de manière cohérente dans la scène en respectant ces paramètres, en leur appliquant une translation et une rotation adaptées. En réduisant l'opacité de la texture, nous pouvons constater que les sphères représentant les stations sont effectivement alignées de manière cohérente avec le nuage de points (Figure 101).

Figure 101 : Visualisation de données issues de lasergrammétrie (©CNP/AGP), comprenant le nuage de points total et deux stations. À gauche : image obtenue en se positionnant sur la première station, avec la sphère en opacité nulle. À droite : même positionnement et sphère à 50% d'opacité. La transparence montre que la texture panoramique de la sphère est positionnée de manière cohérente par rapport au nuage de points. Illustration : auteur.

Des problèmes peuvent survenir lorsque les données sont géolocalisées. Par exemple, dans le cas des données de l'abri du Cap-Blanc, les positions des stations selon les axes X et Y sont exprimées en millions de mètres et possèdent six décimales. Pour éviter les erreurs d'alignement que pourrait causer cette abondance de chiffres significatifs, une approche est de modifier préalablement les paramètres d'export des positions de stations afin qu'elles soient exprimées dans un repère plus proche. C'est ce que nous avons fait pour obtenir le résultat illustré en Figure 101. Nous pouvons constater que le fichier (Figure 100) possède une entrée nommée `References`. Celle-ci spécifie la translation opérée pour le changement de repère (ou le *shift* des coordonnées). Ainsi, nous pouvons conserver les informations relatives à la géolocalisation sans accumuler d'erreur, par le biais de la relation suivante :

$$\begin{pmatrix} X_{geo} \\ Y_{geo} \\ Z_{geo} \end{pmatrix} = \begin{pmatrix} X_{pos} - X_{ref} \\ Y_{pos} - Y_{ref} \\ Z_{pos} - Z_{ref} \end{pmatrix}$$

Où : X_{geo}, Y_{geo}, Z_{geo} sont les coordonnées géolocalisées ;

X_{pos}, Y_{pos}, Z_{pos} sont les coordonnées exprimées dans le fichier texte avec décalage du repère global ;

X_{ref}, Y_{ref}, Z_{ref} sont les coordonnées de référence spécifiant la valeur du décalage du repère global.

Cette perspective est intéressante dans la mesure où elle permettrait indifféremment de charger au sein du même *viewer* des données issues de photogrammétrie ou de lasergrammétrie. Nous avons vu en Chapitre 2 que ces types d'acquisitions ne répondent pas aux mêmes besoins, la lasergrammétrie permettant d'acquérir de grands volumes avec une précision homogène, et la photogrammétrie étant plus flexible et plus polyvalente. Cette dernière est prioritairement exploitée sur le relevé de panneaux, en particulier lorsque les exigences de précision sont élevées (par exemple avec des images macro).

L'alignement cohérent des sphères par rapport au nuage de points 3D permet d'envisager d'étendre les outils existants de propagation d'annotation 2D → 3D proposés par Aïoli, connaissant le système de projection géométrique exploité.

Pour permettre aux utilisateurs d'associer des projets réalisés dans Aïoli par photogrammétrie avec les données issues de leurs campagnes d'acquisition laser, une approche serait de repenser le module de photogrammétrie actuellement utilisé afin de permettre la géolocalisation des projets ou bien de contraindre les images autour d'un nuage de points préexistant. Une autre possibilité serait de ne permettre le téléversement de données de lasergrammétrie qu'une fois le projet créé, et d'exploiter un algorithme comme ICP afin de calculer la transformation nécessaire pour obtenir le recalage du projet dans le repère du nuage laser. Enfin, une autre approche, déjà utilisée dans NUBES, serait de proposer un outil de calage basé sur une sélection manuelle de plusieurs points de correspondances entre le nuage du projet et celui issu de lasergrammétrie.

5.2.3. Bilan

À ce stade, l'approche que nous proposons suffit à afficher au sein d'un même *viewer* des ressources variées telles que les nuages de points d'un projet, les photographies orientées accompagnées des images de caractérisation morphologiques, ainsi que les annotations issues de propagations (Figure 102, Annexe 9). Nous avons également montré qu'il est possible de prendre en compte des données issues de lasergrammétrie (nuages de points des stations et images panoramiques associées), bien que la fusion des données photogrammétriques et lasergrammétriques soulève encore certaines limites en termes de recalage.

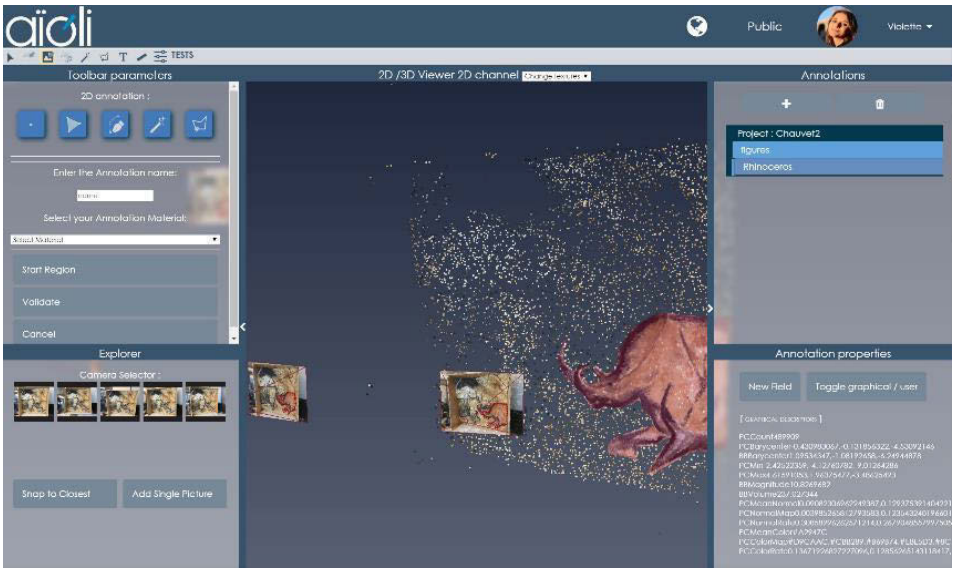


Figure 102 : Notre approche permet de visualiser les données d'un projet dans un *viewer* unique (nuage de points SIFT, images maîtresses, et une annotation 2D et 3D).

La diversité des représentations réunies dans ce *viewer* permet d'une part de limiter l'effet de dispersion des données, mais surtout de faciliter la compréhension de l'objet et rendre compte de sa complexité, tant géométrique que visuelle. **Cette approche est rendue possible par la maîtrise des différentes étapes du processus de spatialisation et d'annotation et la conservation des données de calcul, ce qui souligne bien l'intérêt d'une approche intégrée.**

À partir de ce travail – dont nous détaillerons l'implémentation informatique en Chapitre 7, nous pouvons maintenant nous intéresser à la définition d'une méthode permettant de réconcilier les environnements réels et virtuels par le biais de modalités d'interactions en réalité augmentée. Il s'agit là de permettre d'accéder *in situ* à l'ensemble des données spatialisées dans le *viewer* en réalité augmentée, afin de permettre aux utilisateurs de visualiser et interagir de manière contextualisée avec un environnement hybride comprenant à la fois l'objet réel et les informations numériques et sémantiques associées.

5.3. Méthode proposée pour le recalage en réalité augmentée des ressources d'un projet

Pour atteindre notre objectif, nous devons en premier lieu proposer une méthode de recalage pour assurer la cohérence de l'alignement des environnements réels (objet d'étude) et virtuels (scène 3D du projet correspondant). Dans cette section, nous formaliserons le problème que cela soulève au regard des choix technologiques réalisés, puis nous proposerons une approche semi-automatique pour y répondre. Nous détaillerons deux méthodes possibles pour aborder l'implémentation. Enfin, nous nous intéresserons aux apports techniques et méthodologiques introduits par la notion de persistance.

5.3.1. Formalisation du problème

Comme nous l'avons énoncé en Chapitre 4, le fait de nous inscrire dans un environnement web nous a conduits à privilégier l'API WebXR. Ainsi que nous le détaillerons en Chapitre 7, ce choix technologique permet de développer des applications web de réalité augmentée à partir d'une architecture client-serveur classique. Les navigateurs compatibles fournissent un certain nombre d'éléments dépendant des SDK pris en charge et de l'appareil utilisé pour l'affichage.

Néanmoins, ce choix soulève aussi certaines contraintes pratiques en termes d'implémentation, d'une part du fait du caractère expérimental des outils concernés, d'autre part car les environnements web soulèvent des enjeux importants en termes d'optimisation qui ont pour effet de limiter le nombre et la nature des données exposées, notamment celles issues du suivi AR. En effet, les points caractéristiques issus du suivi ne sont pas accessibles, et les images du flux vidéo de la caméra ne sont exposées qu'avec une définition réduite pour des raisons de latence. Nous ne disposons pas non plus de données relatives au suivi (repère arbitraire choisi lors de l'initialisation, images clés, optimisations, etc.) ni de la possibilité de le contraindre (par exemple à partir de données 3D).

Ainsi, pour doter notre système d'information de modalités d'interactions en réalité augmentée, la marge de manœuvre est plus limitée que pour une application native. De fait, cette configuration élimine d'office les méthodes de suivi basées sur l'utilisation d'un modèle 3D, qui auraient pu s'avérer intéressantes dans notre cas, puisque chaque projet comprend nécessairement un nuage de points dense de l'objet d'étude. Notre méthode de recalage doit non seulement être indépendante du contenu manipulé et de l'environnement réel (c'est-à-dire ne nécessiter aucune scénarisation préalable, ni connaissance ou maîtrise de l'environnement), mais elle doit également être exclusivement basée sur les données exposées par le WebXR, c'est-à-dire les poses successives des caméras du mobile exprimées dans un système de coordonnées arbitraire différent de celui du *viewer*.

En outre, comme nous l'avons vu précédemment, les données 3D des projets ne sont pas géoréférencées. Cela signifie que le nuage de points produit par photogrammétrie s'exprime dans un repère arbitraire, que nous nommerons R_{scene}, et que la transformation entre R_{scene} et un repère global R_{monde} défini par nos soins est inconnue. Rappelons par ailleurs que les données 2D sont spatialisées relativement au nuage de points dans ce même repère arbitraire au sein d'une scène 3D.

En parallèle, les poses de la caméra du mobile, obtenues par le suivi AR et communiquées au client par le navigateur, s'expriment elles aussi dans un repère arbitraire, qui dépend de la pose du mobile au moment de l'initialisation du suivi. Nous nommerons ce repère R_{AR}. Étant donné qu'il dépend de la pose initiale de la caméra, ce repère arbitraire change à chaque session d'utilisation, autrement dit, la transformation entre R_{AR} et R_{monde} est aussi inconnue, et elle n'est pas constante. **Ainsi, le problème de recalage doit être résolu distinctement à chaque session d'utilisation.** En supposant dans un premier temps que la mise à l'échelle du projet a déjà été effectuée par l'utilisateur ainsi que nous le détaillons en section 7.2.4 le problème peut être schématisé comme sur la Figure 103.

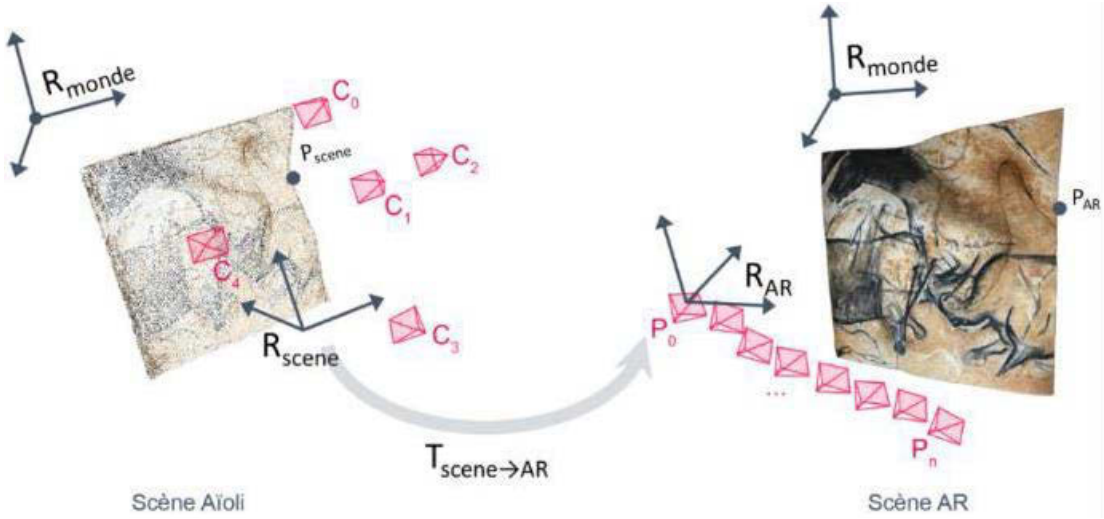


Figure 103 : Schématisation du problème posé par le recalage de deux scènes présentées dans des repères arbitraires. Illustration : auteur.

Pour assurer le recalage des deux scènes, nous cherchons donc la transformation $T_{scene \rightarrow AR}$ telle que :

$$P_{AR} = T_{\text{scene}\rightarrow\text{AR}} P_{scene}$$

$$\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

Où :

P est un point quelconque de l'espace ;

P_{AR} sont les coordonnées homogènes de ce point P exprimées dans le repère R_{AR} ;

P_{scene} sont les coordonnées homogènes de ce point P exprimées dans le repère R_{scene} ;

$T_{\text{scene}\rightarrow\text{AR}}$ est une matrice 4x4 composée d'une rotation [R] et une translation [t].

Étant donné que nous avons supposé que la scène de départ était à l'échelle, six paramètres sont nécessaires pour déterminer les coordonnées d'un point de R_{AR} en fonction des coordonnées du point de la scène R_{scene} : trois paramètres de rotation (autour des trois axes), et trois translations. Ainsi, réaliser le recalage par la recherche de la matrice de passage nécessite de connaître les coordonnées de trois points communs aux deux scènes exprimées dans chacun des deux repères. Ou encore, pour reformuler en termes d'infographie 3D, si M_{scene} est la matrice 4x4 de transformation d'un objet dans R_{scene} , et M_{AR} la matrice 4x4 de transformation du même objet dans $R_{AR,}$ on recherche $T_{\text{scene}\rightarrow\text{AR}}$ tel que :

$$M_{AR} = T_{\text{scene}\rightarrow\text{AR}}\, M_{scene} \qquad \text{soit :} \qquad T_{\text{scene}\rightarrow\text{AR}} = M_{AR}\, {M_{scene}}^{-1}$$

Autrement dit, pour résoudre ce problème nous devons déterminer les transformations locales d'un élément 3D commun aux deux scènes, exprimées dans chacun des deux repères.

5.3.2. Approche semi-automatique proposée pour le recalage d'un projet

Face à ce problème, nous proposons une méthode de recalage basée sur l'utilisation conjointe d'un calcul de résection spatiale et du suivi basé vision. En effet, comme nous l'avons vu précédemment, la méthode de suivi nous permet d'une part d'estimer la pose de la caméra en temps réel dans un système de coordonnées arbitraire non géoréférencé R_{AR}. D'autre part, toutes les données géométriques et sémantiques 2D et 3D d'un projet peuvent être représentées dans une seule scène cohérente et non géoréférencée de repère R_{scene}.

Il ne manque donc que la spatialisation d'une seule *frame* du flux vidéo[39] dans R_{scene} pour disposer de sa pose dans les deux repères : par photogrammétrie pour R_{scene}, par le suivi SLAM pour R_{AR}. Ainsi, nous pouvons réduire le problème à un calcul de résection spatiale, et en déduire $T_{\text{scene}\rightarrow\text{AR}}$.

Il reste alors à appliquer cette transformation à l'ensemble de la scène, afin qu'elle apparaisse positionnée de manière cohérente en réalité augmentée. Le processus envisagé est schématisé sur la Figure 104.

À T_0, c'est-à-dire lors de la création d'un projet, l'utilisateur téléverse ses photographies, qui sont traitées sur un serveur par un processus de référencement spatial basé sur la photogrammétrie et l'indexation des données obtenues (voir section 5.1.2). Au terme de ce processus, le projet comprend un nuage de points dense et des images indexées orientées, toutes représentées dans le repère R_{scene}. À l'issue de cette étape, la base de données 3D du projet est prête pour les annotations mais aussi pour la réalité augmentée.

À chaque T_n, c'est-à-dire à chaque session d'utilisation d'un projet, le mode de visualisation en réalité augmentée peut être activé. Dans ce cas, le navigateur communique au client en temps réel les poses successives de la caméra du mobile dans R_{AR}. Si l'utilisateur se positionne devant son objet d'étude et initie l'alignement, nous pouvons extraire la *frame* active du flux vidéo et l'envoyer au serveur pour un calcul de résection spatiale. À l'issue de ce calcul, nous disposons de sa pose, exprimée dans les deux repères R_{scene} et R_{AR}. Nous pouvons donc calculer la matrice de passage $T_{\text{scene}\rightarrow\text{AR}}$.

[39] Pour la suite de ce travail, cette *frame* sera désignée comme *frame* de référence.

Enfin, nous pouvons appliquer cette matrice à l'ensemble de la scène, c'est-à-dire à l'objet 3D global contenant tous les éléments du projet (voir section 5.2.1). Le recalage est alors terminé, et les données du projet sont théoriquement contextualisées de manière cohérente dans la réalité.

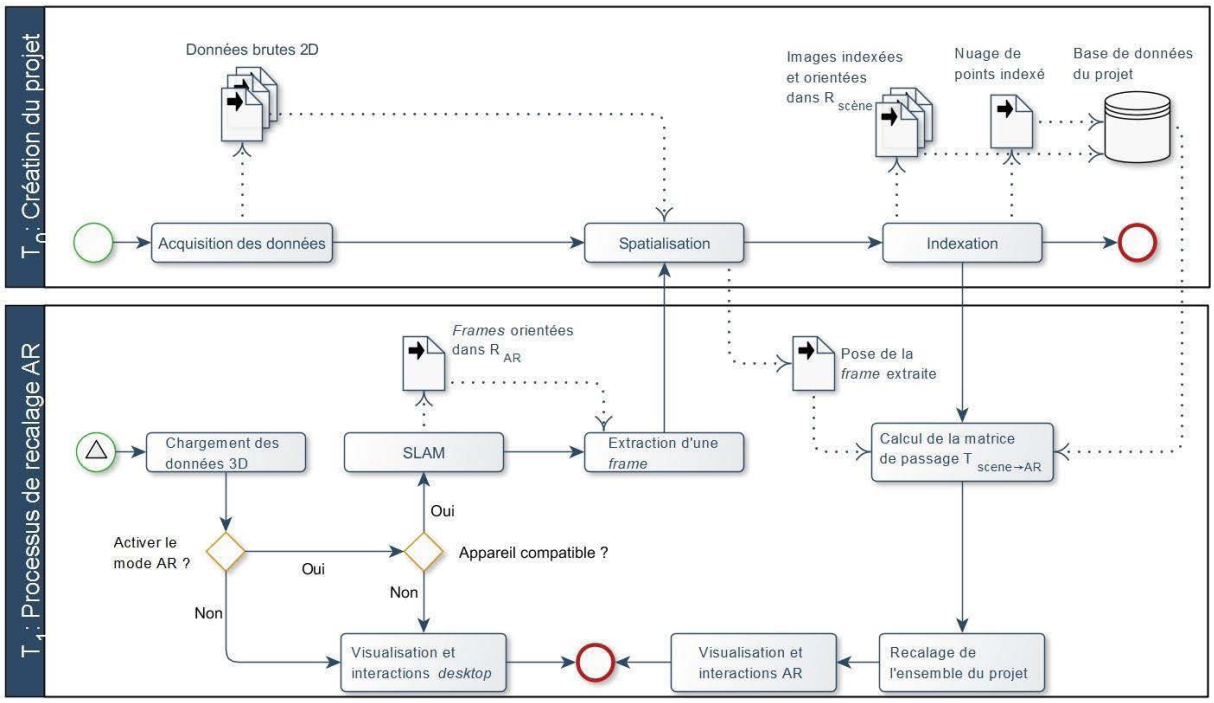


Figure 104 : Processus proposé pour le calcul de la matrice de passage permettant le recalage AR d'un projet. Flèches pleines : déclenchement de processus, flèches en pointillés : transfert de données. Illustration : auteur.

Cette approche repose donc essentiellement sur deux éléments : la précision du suivi des poses de la caméra, et la précision du calcul de résection spatiale. Sachant qu'une erreur de 1 degré représente un écart de plus de 17 centimètres à une distance de 10 mètres, nous devrons être vigilants car la qualité de l'alignement réel/virtuel pourrait être sérieusement compromise. De ce fait, nous pouvons supposer que cette méthode d'alignement sera plus adaptée aux scènes de dimensions contenues.

En outre, le choix de la *frame* de référence est décisif, car pour le calcul de résection spatiale l'image doit être appareillée avec les autres images utilisées pour la photogrammétrie. Ainsi, elle doit respecter les mêmes contraintes que lors d'une acquisition photogrammétrique : taux de recouvrement élevé, défauts d'images minimaux (flou, bruit, etc.), sujet non monochrome ni réfléchissant, etc.

5.3.3. Spatialisation d'une *frame* par photogrammétrie

Au déclenchement du processus de recalage, l'image active du flux vidéo peut être extraite, puis transmise au serveur sous forme de chaîne de caractères encodée en base64 par le biais d'une requête Ajax. Dans le même temps, les paramètres intrinsèques et extrinsèques de la caméra, exprimés dans R_{AR}, doivent être sauvés. Après un éventuel prétraitement des données, l'image peut être transmise au serveur de traitement pour être traitée dans le conteneur dédié aux processus de calculs photogrammétriques. Le calcul de la pose de cette image par rapport au nuage de points du projet et aux autres images peut être réalisé de deux manières (Pamart, Morlet, et al., 2019).

5.3.3.1. Méthode dite *par résection spatiale* (SR)

Cette méthode repose sur une connaissance préalable de couples de points 2D-3D. Or, dans notre processus, aucun point de contrôle au sol (GCP) n'est connu à l'avance. Pour cette raison, une première phase du processus s'intéresse donc à la recherche de correspondances 2D-3D.

Dans un premier temps, l'image candidate est utilisée pour extraire des points d'intérêt à l'aide de l'algorithme SIFT. À partir de ces points caractéristiques, elle est appareillée aux autres images du projet. Les points de correspondances 2D-2D ainsi établis sont nommés points homologues. Les coordonnées 3D correspondant aux points homologues sont obtenues par projection à partir des cartes de profondeur des images du projet. À l'issue de cette phase, nous disposons d'un ensemble de *n* points 3D de coordonnées connues et des *n* points image correspondants, de coordonnées également connues. Dans un second temps, la pose de l'image peut être calculée à partir de ces correspondances à l'aide de la méthode DLT ou *Direct Linear Transformation* (Abdel-Aziz, Karara, 2015).

À partir de l'équation du modèle de sténopé (section 3.3.6.2.1), nous pouvons poser [K] = [C][T] tel que :

$$m \sim [K]\, M$$

[K] est appelée **matrice de projection** de la caméra. Nous disposons donc pour chaque point d'une équation sous la forme :

$$\begin{bmatrix} u_i \\ v_i \\ 1 \end{bmatrix} \sim [K] \begin{bmatrix} X_i \\ Y_i \\ Z_i \\ 1 \end{bmatrix}$$

Ce qui nous donne deux équations linéairement indépendantes qui sont :

$$u_i = \frac{K_{11}X_i + K_{12}Y_i + K_{13}Z_i + K_{14}}{K_{31}X_i + K_{32}Y_i + K_{33}Z_i + K_{34}} \quad \text{et} \quad v_i = \frac{K_{21}X_i + K_{22}Y_i + K_{23}Z_i + K_{24}}{K_{31}X_i + K_{32}Y_i + K_{33}Z_i + K_{34}}$$

Ces équations peuvent être réécrites sous la forme $B_{2n\,x\,12}\, p_{12} = 0_{2n}$ où p est un vecteur composé des coefficients K_{ij}. La matrice B peut être calculée à partir des coordonnées des points homologues et le vecteur p contient les 12 inconnues de la matrice [K]. Chaque correspondance fournissant deux équations, 6 correspondances au minimum sont requises pour calculer p. Les paramètres internes et externes de la caméra peuvent ensuite être extraits à partir de p (Hartley, Zisserman, 2003). En pratique, les coordonnées 2D et 3D sont bruitées, les résultats de la méthode DLT doivent alors être affinés pour éliminer les valeurs aberrantes, ce qui est réalisé à l'aide de RANSAC (*RANdom SAmple Consensus*) (Fischler, Bolles, 1981).

Cette méthode est intéressante dans la mesure où elle permet d'estimer simultanément les paramètres intrinsèques et extrinsèques de la caméra, et peut donc s'appliquer à des images dépourvues de métadonnées. Cependant, l'utilisation d'une caméra calibrée permet en général d'obtenir des résultats plus fiables car en l'absence d'informations sur la focale de l'appareil cette méthode peut procurer une incohérence concernant la profondeur (Pamart, Ponchio, et al., 2019). Théoriquement, la méthode DLT peut aussi être utilisée lorsque les paramètres intrinsèques sont connus à l'avance. Néanmoins, dans ce cas le problème est surparamétrisé, les résultats sont alors moins stables et requièrent plus de mises en

correspondances de points que nécessaires (Tamaazousti, 2013). Ainsi, en l'absence de données de calibration cette méthode est à privilégier. Sinon, il faut privilégier la méthode dite *par ajustement de faisceaux*.

5.3.3.2. Méthode dite *par ajustement de faisceaux* (BBA)

La géométrie épipolaire décrit les contraintes reliant les observations par deux caméras Cam_1 et Cam_2 d'une même scène. Dans le cas d'une projection perspective, il existe un faisceau de points 3D qui se projettent sur le même point du plan image. La position 3D d'un point m_1 d'une image est donc située quelque part sur la droite passant par le centre optique de la caméra et m_1, que l'on appelle rayon de projection. La projection de cette droite sur la deuxième image donne lieu à une droite l_2 (Figure 105).

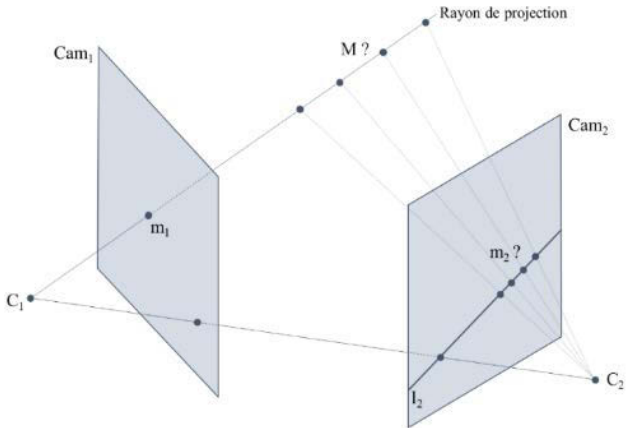


Figure 105 : Contrainte épipolaire. Illustration : auteur, d'après (Larnaout, 2014).

Étant donné que M, le point 3D correspondant à m_1, doit se trouver sur le rayon de projection, sa projection sur le plan image de Cam_2 se trouve nécessairement sur la droite l_2. On parle alors de contrainte épipolaire. Lorsque les paramètres internes des caméras sont connus, cette contrainte épipolaire peut se traduire par l'utilisation de la matrice fondamentale [F], qui permet de mettre en relation les correspondances de points entre les deux plans images avec la relation suivante :

$$m_2^T\,[F]\,m_1 = 0 \qquad \text{Où [F] est une matrice 3x3.}$$

Le déplacement entre les deux caméras peut alors être calculé, à un facteur d'échelle près, à partir de cinq correspondances 2D-2D (Tamaazousti, 2013). De manière globale, cette méthode est utilisée en photogrammétrie pour estimer de manière itérative la trajectoire d'une caméra à partir de correspondances 2D-2D et reconstruire ensuite, par triangulation, une représentation 3D initiale de la scène observée. Cependant, les poses et les points 3D ainsi calculés ne constituent qu'une estimation initiale, qu'il est recommandé de raffiner à travers une étape d'optimisation.

L'ajustement de faisceaux ou *Bundle Block Adjustment* est un processus d'optimisation dont l'objectif est d'optimiser un vecteur de paramètres pour réduire une fonction d'objectif donné. Dans notre cas, la fonction d'objectif vise à réduire les erreurs de reprojection, soit à minimiser la somme des carrés des écarts entre les points détectés dans les images et les reprojections obtenues à partir des points 3D calculés. Dans le module MicMac que nous utilisons, cette minimisation est réalisée par une procédure itérative, l'algorithme de Levenberg-Marquardt (Deseilligny, Cléry, 2011). Rappelons que nous cherchons à spatialiser une unique image par rapport à un jeu de données existant (acquisition maîtresse).

Étant donné que nous ne ce calcul ne vise que le recalage des scènes et non compléter le projet par de nouvelles données, les poses des caméras de l'acquisition maîtresse ainsi que les points 3D du nuage doivent être gelés pour éviter d'introduire des erreurs.

Cette approche est intéressante lorsque les paramètres internes de l'image candidate sont connus. Cependant, selon l'écart de résolution par rapport aux images de l'acquisition maîtresse, les différences de focales, de radiométrie, ou de contexte, cette méthode peut s'avérer délicate, voire échouer en raison d'un manque de correspondances (Pamart, Morlet, et al., 2019). De plus elle est plus coûteuse en temps de calcul.

Quelle que soit la méthode choisie pour la spatialisation de la *frame* référence, le temps de calcul – qui nécessite plusieurs minutes pour un projet moyen – constitue une limite importante de l'approche que nous proposons, d'une part car l'utilisateur devra attendre la fin du processus de recalage pour accéder à la visualisation en réalité augmentée de son projet, d'autre part parce que le SLAM peut provoquer une dérive progressive qui nous conduit à limiter le temps de calcul pour favoriser le temps d'utilisation effective.

5.3.4. Persistance AR des projets

Étant donné que le temps de calcul de résection spatiale ou d'ajustement de faisceaux est trop important pour envisager une utilisation systématique à chasse session AR sans nuire au confort d'utilisation et à la qualité du recalage, nous proposons une seconde approche basée sur la notion de persistance.

5.3.4.1. Précalibration des projets et sessions d'utilisation

La persistance est un concept désignant le fait de conserver l'état d'un environnement augmenté au-delà de la fin de la session d'utilisation. Ce concept est un élément clé pour permettre les expériences multitemporelles ou collaboratives. Dans notre cas, il nous permet d'envisager des scénarios opérationnels plus réalistes grâce à la réduction du temps de calcul, et des perspectives de collaborations intéressantes pour du relevé pariétal multiutilisateurs.

La persistance se base sur la possibilité de sauvegarder à la demande un objet contenant les informations de cartographie spatiale produites lors du suivi. Le chargement de cet objet lors d'une session d'utilisation ultérieure depuis la base de données permet alors d'accéder aux ancres construites lors de la session de référence. Concernant notre procédure de recalage, ce fonctionnement est très intéressant et nous permet de proposer une variante à notre approche, réduisant considérablement le temps de calcul (Figure 106).

Lors du processus de recalage tel que décrit dans les sections précédentes (T_1), nous proposons de créer une ancre correspondant à la pose P_{ref} de la *frame* référence exprimée dans R_{AR}. Cette ancre est donc suivie et actualisée au même titre que les autres ancres produites par le suivi AR. À l'issue du recalage, et plus précisément, en cas de succès, nous pouvons compléter le processus par une étape supplémentaire visant à sauvegarder l'objet contenant les informations de cartographie spatiale. Cette session peut alors faire office de ***précalibration***.

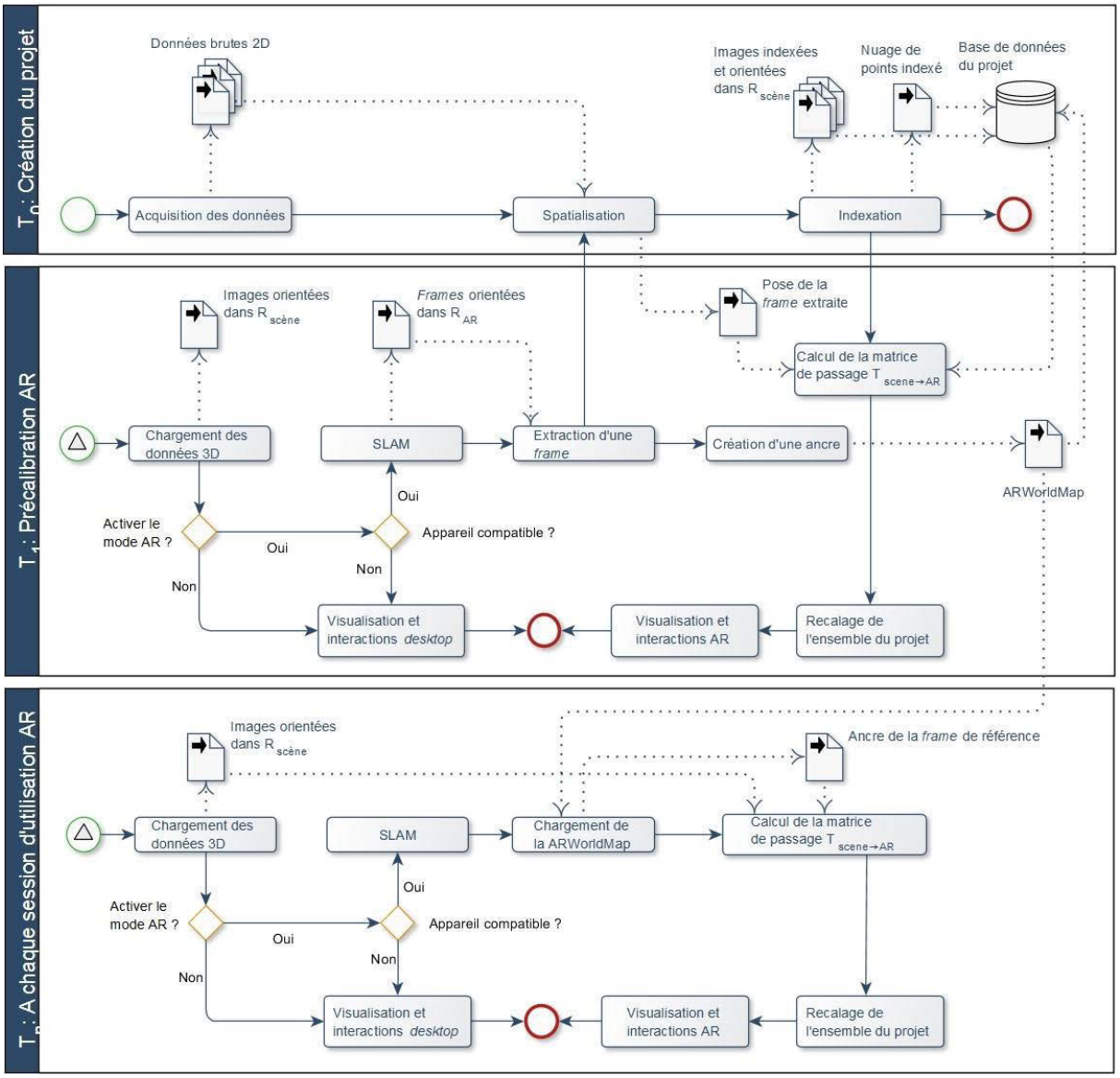


Figure 106 : Processus proposé pour diminuer le temps dédié au recalage. Illustration : auteur.

Lors d'une nouvelle session, le repère arbitraire établi pour le suivi n'est plus le même, mais nous pouvons charger les informations précédemment sauvegardées, et ainsi récupérer les ancres provenant de la session de précalibration. Parmi elles se trouve P_{ref}, l'ancre correspondant à la pose de la *frame* référence. Cette dernière a quant à elle bien été indexée lors de la précalibration. Elle est donc également chargée en tant qu'image du projet, bien que son identifiant la désigne comme image de recalage. Ainsi, sans avoir à réaliser de calcul de résection spatiale supplémentaire, nous disposons donc déjà de toutes les informations nécessaires : la pose P_{scene} de la *frame* référence dans R_{scene}, et sa pose P_{ref} dans R_{AR}. Il suffit alors de calculer la matrice de passage et de l'appliquer à l'ensemble de la scène. Dans cette configuration, le recalage ne prend que quelques secondes.

5.3.4.2. Potentiel collaboratif de la persistance

Nous avons vu que la notion de persistance permet de réduire considérablement le temps nécessaire au recalage. Ainsi, la précalibration demande plusieurs minutes, mais **toutes les sessions ultérieures n'exigent qu'une procédure d'une durée de l'ordre de la seconde**. Ce principe peut être exploité

pour proposer des outils collaboratifs. En effet, en permettant le partage des informations de cartographie spatiale de la session de précalibration aux différents acteurs travaillant sur le même projet[40], nous pouvons assurer que chaque utilisateur évolue dans la même scène augmentée (Abergel, 2019).

Dans le cas du panneau des chevaux de la grotte Chauvet-Pont d'Arc dont nous disposons d'un facsimilé, un relevé plastique a été exploité pour mener une étude stylistique visant à établir une chronologie des interventions anthropiques et à identifier les différents éléments du travail de composition des auteurs (Tosello, Fritz, 2004). Cette étude fait elle aussi intervenir des niveaux de lecture, s'intéressant aux techniques d'exécution (tracé au fusain, remplissage étalé au doigt, estompage, gravure, etc.), aux matériaux (charbon, argile, mélange de consistance pâteuse (sic), etc.) et à la chronologie des figures pour laquelle les auteurs de l'étude distinguent cinq phases principales. Dans ces contextes, les différents points d'intérêts qui apparaissent clairement interreliés pourraient tout à fait faire l'objet de calques. À ce titre, nous pensons que dans de telles situations, une session de relevé collaborative en réalité augmentée permettrait aux utilisateurs de bénéficier des observations des différents acteurs en temps réel, et ainsi, de faciliter le travail d'analyse.

En termes d'implémentation, cette perspective est rendue possible par la persistance permettant à plusieurs utilisateurs de partager les données de précalibration, et travailler sur le même projet. Elle nécessite seulement une actualisation régulière des contenus de la scène en cas de changement dans la base de données (ajout, édition ou suppression de calque ou de région, modification des descripteurs sémantiques, etc.). Pour cela, deux approches sont possibles : actualiser les contenus à intervalle régulier quelles que soient les activités sur la base de données, ou implémenter un système de notification d'évènements, qui communique au client un message lors d'une activité, le client pouvant alors déclencher une fonction d'actualisation le cas échéant. Cette seconde option semble la plus pertinente en termes d'optimisation.

5.4. Bilan

Dans ce chapitre, nous avons tout d'abord proposé une méthode permettant de spatialiser les différentes ressources impliquées dans le relevé d'art pariétal au sein d'un espace de représentation unique cohérent. La diversité des représentations réunies dans ce *viewer* ne tend pas seulement à limiter la dispersion des données, mais aussi à faciliter la compréhension des objets d'étude par l'analyse des relations spatiales et sémantiques entre leurs différentes représentations. Comme nous l'avions souligné dans le Chapitre 1, la possibilité d'un traitement analytique guidé par des critères sémantiques constituait bien l'une des exigences énoncées par Jean Airvaux, à laquelle nous nous sommes attachés à répondre à travers la structure par calques du système d'annotations sémantiques spatialisées et la caractérisation des aspects géométriques et visuels de l'objet d'étude.

À partir de ce travail, nous avons proposé une méthode de recalage dans l'optique de concilier les apports du relevé *in situ* et *ex situ*. Face aux précautions imposées par le caractère sensible des grottes ornées, nous avons proposé une méthode d'alignement réel/virtuel générique et sans marqueurs. En outre, nous nous sommes attachés à privilégier un environnement web en raison du potentiel collaboratif, ce qui

[40] Comme nous l'avons vu précédemment, l'étude de l'art pariétal fait intervenir des disciplines très variées et différents niveaux de lecture. Par exemple, l'étude de la grotte de Cussac fait intervenir des disciplines diverses telles que la karstologie, climatologie, microbiologie, géologie, anthropologie, paléogénétique, paléogénomique, etc.

nous conduit à utiliser une ébauche d'API, le WebXR. En dépit des contraintes posées par le caractère expérimental de ce choix technologique, l'approche proposée permet bien de visualiser et d'interagir avec les contenus du *viewer* en réalité augmentée. Pour l'étude de l'art pariétal, cela semble constituer une avancée importante, car en garantissant la continuité du travail réalisé sur site et hors site, nous proposons, à travers le *relevé multimodal* de concilier à la fois les capacités perceptives des observateurs et les capacités calculatoires de l'informatique.

Par la suite, nous avons proposé une amélioration de notre approche initiale, basée sur une précalibration du projet et la notion de persistance. Cette méthode a permis de diminuer considérablement le temps de calcul, et a ouvert des perspectives en matière de travail collaboratif. Elle pourrait également être améliorée en intégrant de nouveaux scénarios de collecte de données.

En effet, le flux vidéo pourrait tout à fait être exploité pour acquérir de nouvelles photographies en appréciant visuellement le taux de recouvrement des images du projet et garantir la complétude des données. Ces photographies extraites pourraient également être géolocalisées par le biais des capteurs du mobile. En outre, la sauvegarde d'images clés en provenance du flux vidéo lors de chaque session d'utilisation ouvre la voie au suivi multitemporel des objets d'étude.

Dans le chapitre suivant, nous étudierons les nouvelles possibilités apportées par ce système d'information multimodal en termes d'analyse et d'interactions.

Chapitre 6. Modalités d'augmentation géométrico-visuelle et d'annotation

6.1. Augmentation géométrico-visuelle des ressources du projet

Dans le chapitre précédent, nous avons présenté une approche permettant d'allier les moments de relevés sur site et hors site, et de manipuler les différentes ressources 2D ou 3D impliquées dans les processus d'analyse. Nous pouvons donc maintenant nous intéresser à l'élaboration d'outils d'enrichissement sémantique et en particulier d'annotation, adaptés à ces nouvelles modalités de relevé.

Comme nous l'avons énoncé en Chapitre 4, notre objectif est d'exploiter cette hybridation pour faciliter l'observation d'altérations sur les parois et de traces anthropiques. Un certain nombre d'informations quantitatives peut être extrait automatiquement des données 3D (boîte englobante, barycentre, position, amplitude, volume, orientation moyenne, etc.). Cependant, bien qu'elles permettent de décrire certains aspects géométriques des ressources, elles n'en facilitent pas nécessairement l'interprétation. Certains attributs géométriques peuvent être utilisés pour produire des caractérisations visuelles d'attributs morphologiques. Comme nous l'avons vu en Chapitre 5, cette approche est déjà exploitée pour produire des images 2D par projection comme des cartes de normales, qui nous ont conduits à proposer un modèle de caméra multicalques.

Nous verrons dans ce chapitre que ce principe peut être étendu pour permettre la prise en charge d'une plus grande variété de ressources 2D, et que celles-ci peuvent profiter de méthodes de traitement d'images facilitant l'analyse. Nous verrons également qu'une approche semblable peut être exploitée en 3D pour produire des textures faisant ressortir certaines caractéristiques de l'objet étudié, ce qui présente des potentiels intéressants dans le cas d'un affichage en réalité augmentée. Nous verrons également que certains traitements peuvent être appliqués au flux vidéo de la caméra du dispositif de réalité augmentée, notamment pour faciliter la détection de stigmates ou visualisation de peintures. Enfin, sur la base de ces méthodes d'accentuation visuelle, nous proposerons des outils d'annotation permettant d'enrichir sémantiquement les ressources du projet par la description des faits observés.

6.1.1. Stratégies d'augmentation basées sur les données 2D

Comme nous l'avons vu précédemment, dans le cas d'une utilisation hors site de la plateforme, les images sont des supports privilégiés pour l'annotation. Les photographies en lumière visible rendent compte de l'aspect de l'objet d'étude et contribuent à l'identification des matériaux, des dimensions, des phénomènes d'altération des surfaces, etc. Certains phénomènes sont plus difficiles à percevoir, notamment les gravures qui ne se manifestent sur les photographies que par de légers changements de ton. Dans ce cas, il peut être utile d'accéder conjointement à l'information colorimétrique et à des attributs morphologiques, pour faciliter le processus de sélection des régions. Ce constat nous a conduits à proposer un modèle de caméra multicalques permettant d'intégrer, pour une vue donnée, les images de caractérisation morphologiques issues de projection 3D → 2D en plus des images maîtresses (Figure 107).

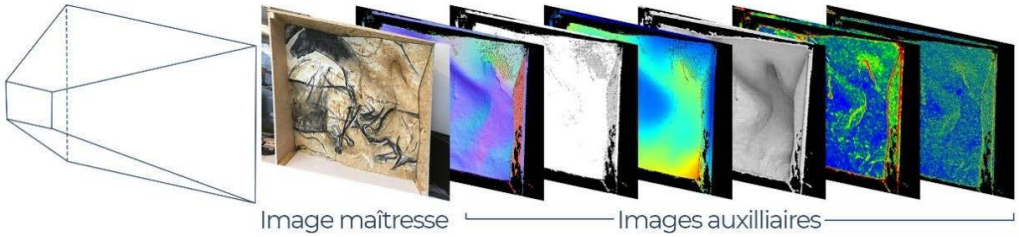


Figure 107 : Modèle de caméra permettant de prendre en compte des images de caractérisation géométrique. Illustration : auteur.

En pratique, cette approche ne se limite pas aux cartes d'attributs morphologiques. Comme nous l'avons vu précédemment (section 2.2.1), en termes de représentation, les images du spectre visible et images de caractérisation morphologiques peuvent être insuffisantes pour visualiser certaines caractéristiques de l'objet étudié, notamment dans le cas des sites d'art pariétal où les conditions environnementales et la dégradation naturelle des parois peuvent rendre les conditions d'observation très difficiles. Plusieurs autres images produites à partir d'analyses scientifiques complètent efficacement ces données pour contribuer la compréhension de l'objet étudié.

Par exemple, nous avons vu que les photographies multispectrales pouvaient permettre de faire apparaître des figures invisibles à l'œil nu, mais aussi que les images RTI permettaient de faire ressortir de fines variations du relief, ou encore que certains traitements appliqués aux images comme le *decorrelation stretching* facilitaient la visualisation d'éléments peints. Le modèle de caméra que nous proposons est justement pensé pour permettre la gestion de cette grande variété d'images, à condition que celles-ci puissent être orientées par photogrammétrie. Au niveau du *viewer*, cela implique néanmoins de préciser le processus de chargement d'images en distinguant deux cas :

- L'image candidate permet une orientation directe par photogrammétrie (Figure 108) ;
- L'image candidate nécessite une orientation indirecte à l'aide d'une image intermédiaire (Figure 109).

6.1.1.1. Images compatibles avec le processus d'orientation directe

Si l'image à orienter offre suffisamment de points communs avec les autres images du projet, sa pose relative au nuage de point peut être calculée à l'aide du module TAPAS. Dans ce cas, dans le *viewer* l'image orientée peut être considérée comme une image maîtresse.

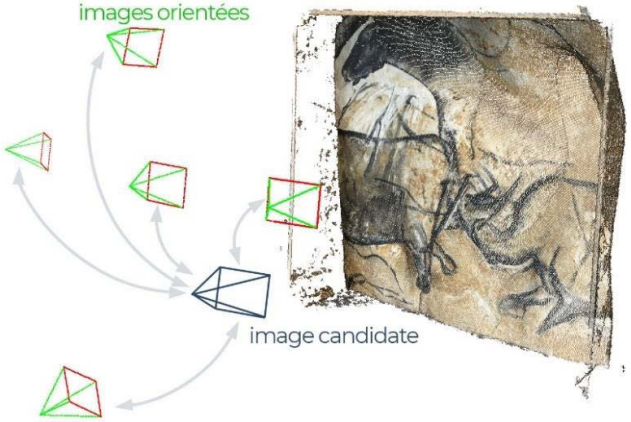


Figure 108 : Orientation directe d'une image par rapport à un jeu d'images déjà orientées. Illustration : auteur.

Cela concerne le plus souvent les photographies numériques en lumière visible mais la pratique montre que le module TAPAS basé sur l'algorithme SIFT est suffisamment robuste pour permettre d'orienter des images présentant de plus importantes variations, comme des images IR ou UV.

6.1.1.2. Images nécessitant une orientation indirecte

Dans certains cas, l'image à orienter ne possède pas suffisamment de points communs avec les autres images du projet. C'est notamment le cas pour les images thermographiques, ou encore pour le cas particulier des images RTI. Dans cette configuration, il est possible d'utiliser une image intermédiaire compatible avec le processus d'orientation directe et dont on connaît la pose relative par rapport à l'image candidate.

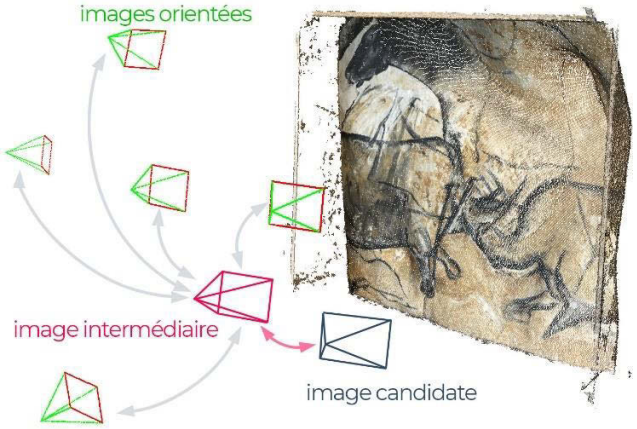


Figure 109 : Orientation indirecte d'une image. L'image candidate est orientée de manière relative par rapport à une image intermédiaire. L'image intermédiaire est orientée de manière directe. Illustration : auteur.

Ainsi, par transitivité, la pose de l'image candidate peut être calculée par rapport au nuage de point. Dans cette configuration, nous devons encore distinguer deux possibilités.

Dans la majorité des cas, l'image intermédiaire et l'image candidate ne partagent pas la même pose. Dans cette situation, à l'issue du calcul photogrammétrique, les poses de ces deux images sont exprimées indépendamment dans des fichiers XML d'orientation distincts. Ainsi, bien que l'image candidate soit calculée de manière indirecte, sa pose est stockée exactement de la même manière que pour les images issues d'orientation directe. Les deux images peuvent donc être chargées dans le *viewer* comme deux images maîtresses distinctes.

Dans d'autres cas, il est possible que les deux images partagent exactement la même pose. L'intérêt de l'image intermédiaire réside alors dans le fait de posséder des caractéristiques plus proches des autres images du projet, ce qui facilite l'extraction de points homologues.

Par exemple, dans le cas des images RTI, c'est la nature dynamique de l'image qui impose de passer par une image intermédiaire pour des raisons d'interopérabilité. Le calcul des paramètres intrinsèques et extrinsèques repose alors principalement sur la pertinence de ce choix : les images brutes ayant servi à la production de l'image RTI disposent d'éclairages trop marqués qui risquent de causer des problèmes de corrélation. Il est préférable de se baser sur l'image albédo produite à partir de l'image RTI, ou encore sur l'image médiane ou moyenne (Pamart, Ponchio, et al., 2019).

Dans cette configuration, l'image intermédiaire doit être considérée comme image maîtresse, et l'image candidate comme l'une de ses textures auxiliaires.

6.1.1.2.1. Images RTI

Outre le calcul de leur pose, la prise en charge d'images RTI dans notre environnement numérique suppose deux problèmes distincts : la création de l'image RTI à partir des données brutes d'acquisition, et l'affichage dans le *viewer*. Concernant le premier problème, deux stratégies sont envisageables :

- Intégrer l'ensemble du processus sur la plateforme, afin de générer l'image RTI directement sur le serveur de traitement, ainsi que son image médiane spatialisable par photogrammétrie ;
- Considérer la création d'une image RTI comme un processus individuel externe à la plateforme, et proposer des modalités d'import des fichiers résultant (généralement au format PTM). Dans ce cas, l'image médiane doit être produite *a posteriori*.

Dans la logique de notre approche intégrée, nous favorisons la première approche, sans occulter la seconde qui peut présenter un intérêt dans les cas où la création de l'image RTI exige des paramétrages ou manipulations spécifiques. Dans les deux cas, Ponchio, Potenziani, et al., (2019) offrent une solution fonctionnelle permettant la création d'images RTI et le prétraitement de fichiers PTM en vue d'une visualisation web. Concernant l'affichage, une fois la pose de l'image médiane calculée, nous pouvons charger l'image RTI dans le *viewer* pour bénéficier de son apport en termes d'analyse. L'utilisation de ce type d'image comme support efficace d'annotation suppose l'intégration de deux processus supplémentaires, dont nous détaillons l'implémentation dans le Chapitre 7 :

- La création d'éléments de GUI permettant aux utilisateurs de varier l'orientation de la source lumineuse. Dans les logiciels de visualisation de RTI les plus répandus, cela se présente sous la forme d'une sphère colorée que l'utilisateur peut faire tourner sur elle-même ;
- Le chargement du fichier RTI comme texture de la pyramide de vue et l'actualisation de celle-ci à chaque changement lumineux ;

La prise en charge de ces images est particulièrement intéressante dans le cas de gravures ou de stigmates se traduisant par de fines variations du relief. Ces éléments sont difficiles à percevoir à l'œil nu, et leur visualisation sur les nuages de points 3D nécessiterait une résolution spatiale élevée qui n'est pas toujours envisageable selon les contextes d'acquisition. La possibilité d'annoter une image RTI tout en faisant varier les directions du faisceau lumineux facilite alors considérablement le travail d'interprétation et de documentation.

6.1.1.2.2. Decorrelation stretching d'images maîtresses ou auxiliaires

Comme nous l'avons vu en section 2.3.1.3, le *decorrelation stretching* est un traitement très courant dans le domaine de l'art pariétal, car il facilite la visualisation d'éléments peints, ce qui peut s'avérer particulièrement délicat sur des photographies non traitées lorsque les pigments sont proches de la couleur de la roche. L'idée générale est de considérer que les composantes des couleurs des différents éléments apparaissant sur une image sont *corrélées* entre-elles, par exemple que chaque élément est associé à une proportion spécifique de rouge, vert, et bleu dans le cas de l'espace colorimétrique RGB. La *décorrélation* a pour objectif de mettre en évidence ces combinaisons, c'est-à-dire déterminer des familles de pixels dont les composantes sont réparties dans les mêmes proportions. Ensuite, il s'agit de remplacer les couleurs de l'image originale par d'autres couleurs plus éloignées les unes des autres.

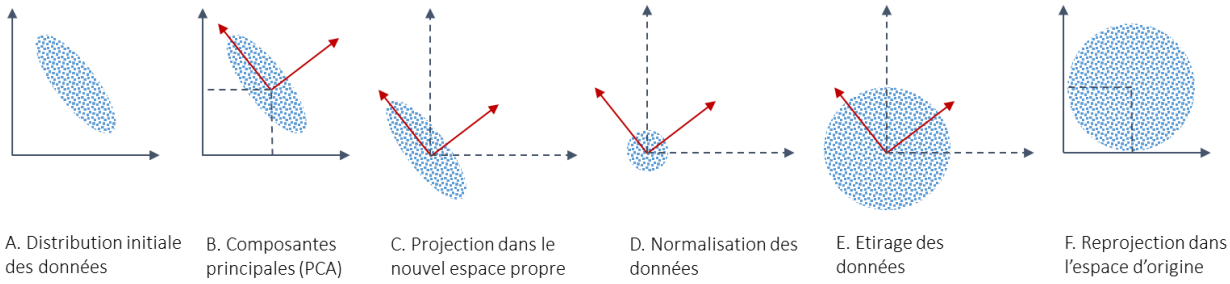


Figure 110 : Principe de l'algorithme de *decorrelation stretching*. Illustration : auteur, d'après (Dangampola, 2015).

Cette opération est basée sur une analyse statistique appelée Analyse en Composantes Principales (PCA) permettant de déterminer les axes exprimant le mieux la dispersion des données (ici, les pixels, considérés comme des points 3D de l'espace colorimétrique). Ces axes constituent un nouveau système de coordonnées dans lequel les données s'expriment de manière décorrélée. Les données sont ensuite mises à l'échelle et étirées puis reprojetées dans l'espace original (Figure 110). Cette opération peut être exprimée comme une transformation linéaire réalisée pour chaque pixel, dont les paramètres dépendent des valeurs statistiques extraites de l'analyse de l'image d'entrée et des valeurs attendues en sortie (Dangampola, 2015 ; MathWorks, 2019).

$$b = [T]\ x\ (a - m) + m_{target}$$

Où :

a est un vecteur 3 contenant les valeurs RVB du pixel concerné de l'image d'entrée ;

b est un vecteur 3 contenant les valeurs RVB du pixel correspondant sur l'image de sortie ;

$[T]$ est une matrice 3 x 3 ;

m est un vecteur 3 contenant la moyenne de chaque bande de l'image d'entrée ;

m_{target} est un vecteur 3 contenant la moyenne attendue pour chaque bande de l'image de sortie (si celle-ci n'est pas spécifiée, sa valeur par défaut est égale à m).

La matrice [T] peut être calculée à partir de la matrice de corrélation (ou aussi à partir de la matrice de covariance). Celle-ci est obtenue par PCA à partir de l'ensemble des données d'entrée ou d'une sous-sélection définie par l'utilisateur pour désigner une portion significative de l'image. Elle s'exprime de la façon suivante :

$$[T] = s_t\ R^T\ s_c\ R$$

Où :

s_t est la matrice d'étirage, ce paramètre permet de régler la variance des données de l'espace d'entrée sur la variance souhaitée en sortie. Il s'agit d'une matrice diagonale 3x3 dont les éléments de la diagonale sont les valeurs de l'écart-type souhaité.

s_c est la matrice de mise à l'échelle. Ce paramètre permet de normaliser les données dans l'espace propre. Il s'agit d'une matrice diagonale 3x3. Les éléments de la diagonale correspondent à l'inverse des valeurs d'écart-type dans l'espace propre.

Comme nous venons de le voir, le *decorrelation stretching* est un post-traitement appliqué à une image numérique n'ayant une incidence que sur la couleur des pixels. Les deux images (entrée et sortie) possèdent donc bien les mêmes poses. Dès lors, l'image d'entrée est celle utilisée pour le calcul photogrammétrique, dans le *viewer* elle est donc considérée comme image maîtresse. L'image décorrélée peut alors être affichée comme l'une de ses textures auxiliaires (Annexe 3). L'intégration de cette méthode sur Aïoli peut être réalisée en conteneurisant l'application sur le serveur de traitement.

Notons enfin qu'il est possible d'appliquer ce traitement à d'autres images, par exemple une carte de normales. Dans ce cas, son apport est moins flagrant que sur les peintures mais permet tout de même de faciliter l'identification de groupes de points dont les orientations sont semblables (Figure 111).

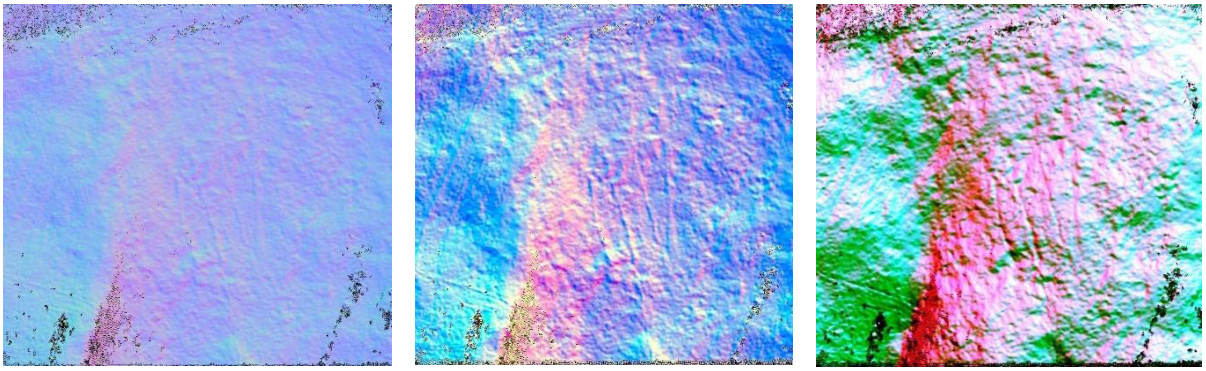


Figure 111 : *Decorrelation stretching* d'une carte de normales provenant d'un projet Aïoli. À gauche : image originale. Au centre : décorrélation dans l'espace colorimétrique LDS. À droite : décorrélation dans l'espace colorimétrique YYE. Illustration : auteur.

6.1.2. Stratégies d'augmentation basées sur les données 3D

6.1.2.1. Stockage des attributs de points

Nous avons vu en section 5.1.2 que les nuages de points de projets Aïoli étaient utilisés lors de l'indexation pour calculer des descripteurs géométriques et produire des images 2D représentant ces attributs. Concernant l'affichage de ces attributs en 3D, le problème principal n'est pas la visualisation mais bien le stockage des données, car les formats usuels de nuage de points ne prennent pas en charge une telle variété et se limitent habituellement aux coordonnées, couleurs, et normales.

Le format PLY est un format initialement développé par l'université de Stanford pour stocker l'information 3D. Ce format de fichier a été pensé pour offrir une solution souple adaptée à la diversité des données 3D. L'information de base contenue dans un fichier PLY est une liste de triplets (*x, y, z*) correspondant aux coordonnées des sommets, suivie éventuellement d'une liste d'indices de faces dans le cas d'un modèle 3D maillé. Chaque point possède un nombre fixe de propriétés dont la nature est spécifiée dans le *header* du fichier. Outre les coordonnées, les plus courantes sont les propriétés RGB. L'intérêt majeur de ce format réside dans la possibilité de créer de nouvelles propriétés rattachées aux points, sans mauvaise conséquence pour les applications n'ayant pas connaissance des nouvelles propriétés.

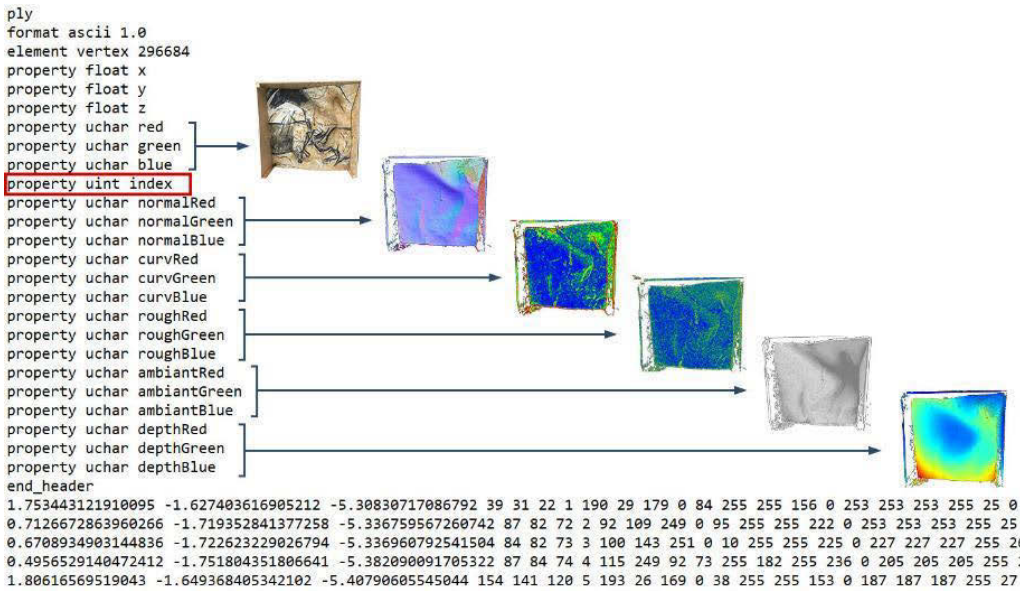


Figure 112 : Exemple d'extrait du fichier PLY d'un nuage de points doté d'attributs stockés comme des couleurs.

Dans notre cas, les attributs géométriques calculés lors de l'indexation ont déjà été traduits en attributs visuels à l'aide de dégradés de couleurs pour générer les images auxiliaires des images maîtresses. Ainsi, pour conserver au sein du *viewer* une homogénéité dans les représentations des différents attributs, il est préférable d'utiliser les mêmes codes couleur que pour ces images. Pour exploiter ces données, et plus particulièrement faciliter la compréhension des objets d'études et leur interprétation par le biais d'outils d'annotation, notre approche est donc de produire à l'issue de l'indexation et pour chaque nuage de points, un fichier PLY doté de nouvelles propriétés colorimétriques précalculées. Celles-ci peuvent être directement encodées comme des couleurs, c'est-à-dire exprimées par des triplets d'entiers (Figure 112).

Nous pouvons réaliser une exception pour l'attribut *position des indices* (Annexe 2). En 2D, celui-ci se manifeste par un masque dans lequel les pixels blancs traduisent la présence d'un point 3D correspondant. En 3D, une telle texture reviendrait à colorer le nuage de points en blanc, ce qui ne présente pas d'intérêt. En revanche, la connaissance des indices de chaque point 3D est intéressante à conserver, particulièrement pour faciliter la procédure de sélection lors d'annotations 3D (voir section 6.2.2).

6.1.2.2. *Decorrelation stretching* d'attributs de points 3D

Comme nous venons de le voir, dans le fichier PLY, les couleurs des points sont stockées comme des caractères non signés à l'aide de trois propriétés *R*, *G*, et *B* dont les valeurs sont comprises entre 0 et 255. De fait, les données colorimétriques des nuages de points sont bien de la même forme que celles d'une image couleur. Dès lors, nous pouvons proposer une adaptation simple de notre application de *decorrelation stretching* afin qu'elle permette de traiter les attributs colorimétriques d'un nuage de points plutôt que les pixels d'une image. Comme les données d'entrée sont de la même nature, l'algorithme en lui-même est inchangé. Ce travail n'implique que la création de fonctions permettant :

- De lire un fichier PLY en entrée pour récupérer les attributs colorimétriques ;
- D'écrire un fichier PLY comprenant toutes les données du fichier d'entrée ainsi que des propriétés supplémentaires liées au(x) traitement(s) des attributs colorimétriques.

En termes d'implémentation, cela ne présente pas d'enjeu particulier dans la mesure où nous travaillons avec des fichiers que nous produisons nous-mêmes lors de l'indexation et dont nous maîtrisons la structure. Comme nous pouvons le voir sur cette illustration (Figure 113), le *decorrelation stretching* appliqué à la texture du nuage de points facilite la visualisation des éléments peints ou de changements de matière sans perte d'informations géométriques. Comme dans le cas des images 2D, nous pouvons exploiter différents espaces colorimétriques selon l'accentuation recherchée.

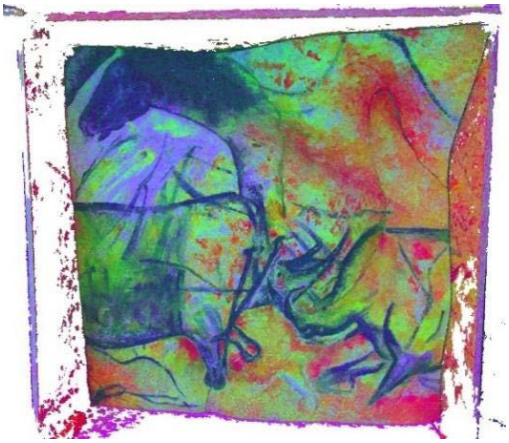


Figure 113 : Rendu du nuage de points avec une texture issue d'un traitement par *decorrelation stretching*.

6.1.2.3. Colorisation des points par les *shaders*

Les méthodes de colorisation des points présentées précédemment concernent toutes des propriétés précalculées. Au niveau du *viewer*, certains *shaders* peuvent également contribuer à l'intelligibilité des données manipulées en intervenant sur certains paramètres de rendu.

Le texturage par **dégradé d'élévation**, par exemple, permet de colorer les points du nuage en fonction de leur hauteur afin de mieux percevoir les dénivelés dans le cas de terrains. Une modification simple de ce *shader* permet de changer l'axe de référence afin l'adapter à l'étude de panneaux dans le cadre de relevé pariétal. Ainsi, le dégradé d'élévation (ou dans ce cas, dégradé de profondeur) peut contribuer à la compréhension des morphologies des parois (Figure 114). Le paramétrage de l'intervalle permet en outre de modifier la plage d'application du dégradé et ainsi, de l'adapter à différentes configurations pariétales de reliefs plus ou moins prononcés.

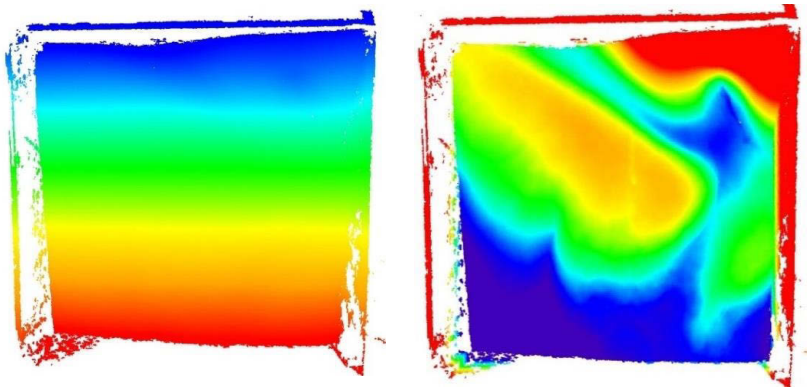


Figure 114 : À gauche : coloration du nuage de points à partir d'un *shader* de dégradé d'élévation. À droite : Modification de l'axe du dégradé pour révéler les variations morphologiques de la paroi.

Une autre méthode de colorisation très intéressante pour l'art pariétal est le **texturage composite**. Le principe de ce *shader* est de permettre la superposition des différentes méthodes de colorisation à la manière de calques dont l'opacité serait réduite. Sur des nuages de points classiques, cette méthode de rendu est surtout intéressante pour la superposition des couches RGB et dégradé de profondeur. Par exemple, sur le facsimilé du panneau des chevaux représenté en Figure 115 ce mode de rendu nous permet d'observer que l'encolure du cheval semble se confondre avec une concavité de la paroi.

Figure 115 : Rendu composite du nuage de points du facsimilé, basé sur l'utilisation conjointe des couches RGB et du dégradé de profondeur.

Cependant, dans Aïoli, la présence des nombreux attributs morphologiques de points calculés lors de l'indexation ou à l'issue de celle-ci, ouvrent de nouvelles perspectives au texturage composite. Celui-ci permet alors d'associer dans le rendu différentes méthodes de colorisation qui renvoient toutes à des propriétés géométriques ou visuelles bien spécifiques de l'objet d'étude. Ainsi, le mélange des couches facilite grandement la compréhension de l'objet, particulièrement en permettant l'identification d'occurrences ou de liens éventuels entre les propriétés. Par exemple, l'utilisation conjointe de la texture décorrélée (qui accentue les peintures) et de la texture par normales (qui met en valeur les orientations de la surface) peut révéler des jeux éventuels entre les figures peintes et leur support.

Enfin, le *shader* **EDL** ou *Eye-Dome Lighting* (voir section 2.3.2.5) offre la possibilité de tenir compte d'ombres produites par différentes sources de lumière, notamment les lumières directionnelles (Figure 116). Dans le cas de l'art pariétal, il s'agit d'une fonctionnalité très importante car, comme nous l'avons vu en Chapitre 2, la simulation de lumières rasantes est très efficace pour faciliter l'observation et l'identification d'éléments gravés ou sculptés. Comme dans le cas des images RTI, cela implique de proposer des éléments d'interface utilisateur dédiés au paramétrage des sources lumineuses dans la scène 3D (position, orientation, intensité, et couleur).

Figure 116 : Détail du panneau des chevaux selon différents modes de rendu. À gauche : RGB. Au centre : *shader* EDL. À droite : *shader* EDL et lumière directionnelle rasante. Ce dernier mode de rendu révèle bien les traces de griffures.

6.1.3. Traitement temps-réel du flux vidéo

Dans les sections précédentes, nous avons présenté des traitements appliqués aux ressources 2D ou 3D chargées dans la scène du projet. Dans le cas d'une utilisation *in situ* en réalité augmentée, des traitements supplémentaires peuvent être envisagés pour contribuer à la compréhension de l'objet d'étude. En effet, dans ce contexte, l'observateur perçoit son environnement à travers les données captées par la caméra de son dispositif d'affichage. Nous pouvons donc exploiter les images du flux vidéo pour accentuer certaines caractéristiques géométriques ou visuelles, à travers l'utilisation de filtres. Le processus consiste à appliquer un algorithme de traitement d'image à chaque *frame* du flux vidéo (par exemple un algorithme de détection de contour comme Canny), et à restituer l'image traitée en arrière-plan du rendu AR plutôt que l'image originale (Figure 117).

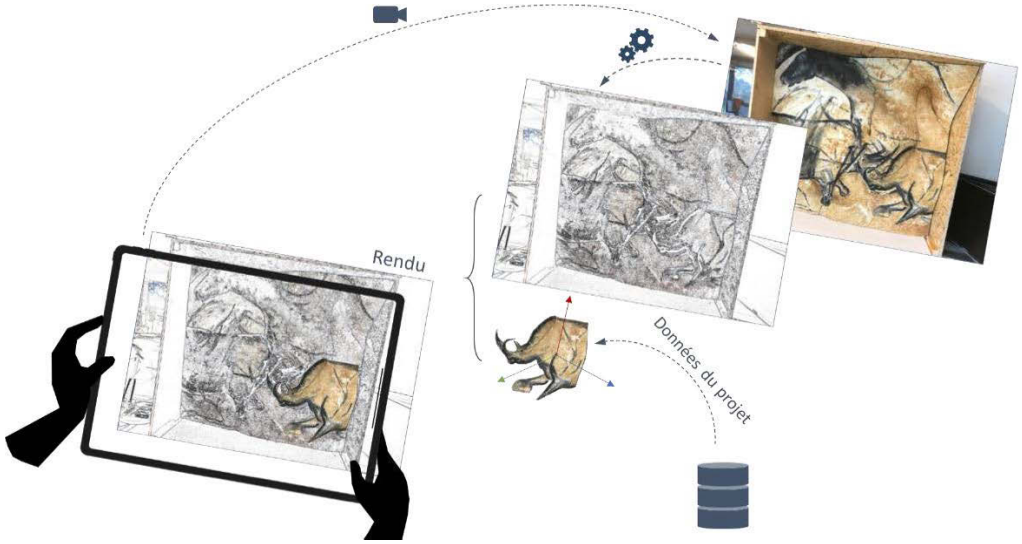


Figure 117 : Schéma du traitement envisagé. Le traitement du flux vidéo pourrait faciliter la compréhension de l'environnement en AR, par exemple à l'aide d'algorithmes de détection de contours. Illustration : auteur.

Il existe un grand nombre d'algorithmes de traitement d'images[41]. Dans le domaine de l'art pariétal, les principales difficultés sont de visualiser les reliefs les plus fins et les éléments peints dont la couleur est souvent dégradée ou trop proche de celle de la paroi. Dans le premier cas, les algorithmes de détection des contours ou de seuillage peuvent fournir des images très intéressantes, éloignées de la réalité du point de vue de la perception mais révélant des artefacts difficiles à observer à l'œil nu (Figure 118).

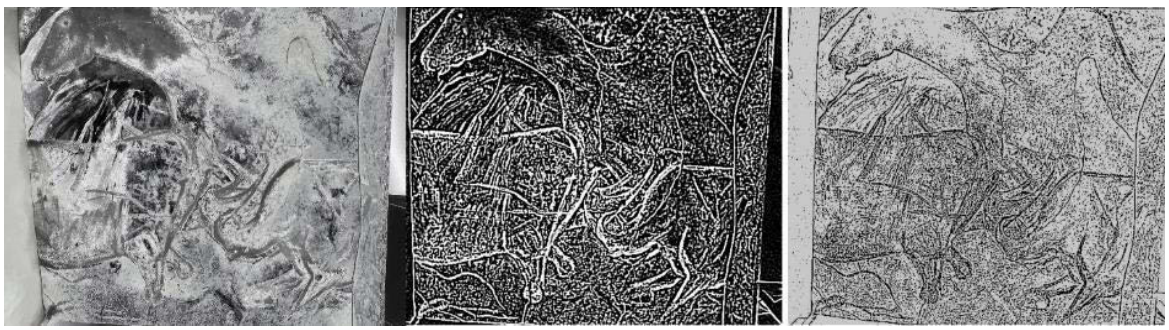


Figure 118 : Exemples de traitements d'images modifiant la perception de l'objet étudié. À gauche : solarisation de l'image (inversion partielle des densités). Au centre : détection des contours par passage par zéro du Laplacien. À droite : seuillage.

[41] Il ne conviendrait pas de détailler dans ce document. Le lecteur intéressé pourra se référer à Gonzalez et Woods (2018).

Dans le second cas, l'application de l'algorithme de *decorrelation stretching* aux *frames* successives du flux vidéo constitue un outil d'**exploration** particulièrement intéressant. En effet, les versions actuelles de l'algorithme ne s'appliquent qu'aux images fixes, ce qui impose dans une certaine mesure de savoir préalablement où se situent les éléments d'intérêt. Une version vidéo temps réel pourrait permettre d'offrir aux observateurs un outil de détection efficace exploitable lors des déambulations sur site (Figure 119).

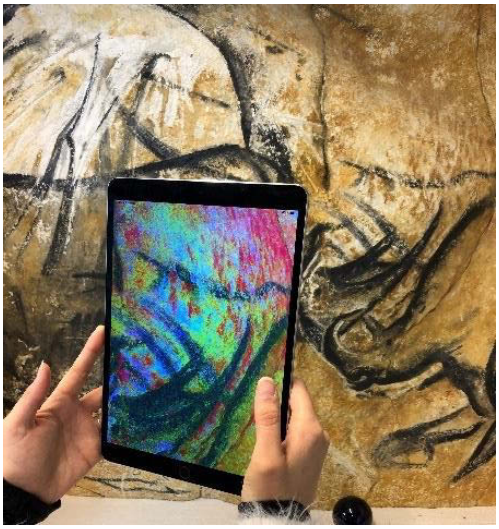


Figure 119 : Utilisation de notre application de *decorrelation stretching* vidéo pour apprécier visuellement les variations de matériaux sur le facsimilé. Illustration : auteur.

Si l'approche s'est avérée prometteuse pour révéler les peintures et faciliter la compréhension instantanée de l'environnement, elle soulève cependant des effets de latence influencés par la fréquence des calculs et la définition des images traitées. Pour les limiter, il conviendrait donc d'adapter notre approche en travaillant avec des images de définition réduite ou de limiter le nombre de calculs, par exemple en n'appliquant la PCA qu'aux images clés du suivi et en conservant les paramètres subséquents entre deux images clés (Figure 120).

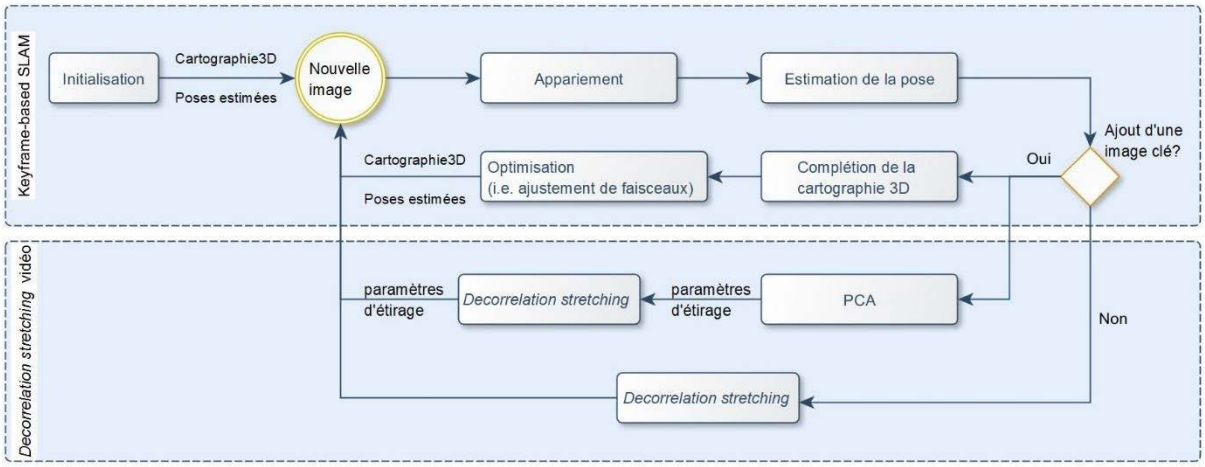


Figure 120 : Processus proposé pour fluidifier le traitement vidéo en exploitant les images-clés. Illustration : auteur.

En ce qui concerne l'intégration de méthodes traitements vidéo temps réel dans la plateforme Aïoli, trois approches sont envisageables.

La première consiste à réaliser les traitements vidéo sur le serveur, mais cela soulève des contraintes importantes car la volonté d'assurer un affichage temps-réel au client suppose des temps de calculs très

courts, lésés par le temps consacré au transfert de données. Une autre approche consiste à consiste à considérer que le navigateur doit gérer par lui-même l'affichage vidéo – comme c'est le cas avec l'API WebXR – et qu'il pourrait par conséquent proposer nativement certains algorithmes de traitement, ce qui limiterait les transferts de données. Néanmoins, cela soulève également plusieurs problèmes, car le développement d'algorithmes spécifiquement dédiés au relevé pariétal suppose alors de proposer notre propre navigateur, en plus de la plateforme *cloud*, ce qui est beaucoup trop contraignant en termes de maintenabilité. Enfin, une troisième approche consiste à implémenter les algorithmes de traitement directement côté client. Bien que plus simple d'un point de vue pratique, elle peut s'avérer très délicate en termes d'optimisation, car la nécessité d'accéder à chaque *frame* et les calculs nécessaires pour appliquer les filtres peuvent causer d'importants problèmes de latence.

6.2. Modalités d'annotation

Nous avons présenté dans les sections précédentes notre approche permettant de spatialiser de manière cohérente les ressources 2D et 3D du projet. Par la suite, nous avons présenté différentes méthodes d'accentuation d'attributs géométrico-visuels afin de faciliter l'analyse et la compréhension des objets d'étude. Pour être véritablement exploitables, ces méthodes doivent être associées à des outils permettant d'interagir avec les éléments de la scène.

Pour cela nous proposons dans les sections suivantes différentes méthodes d'annotation permettant d'enrichir sémantiquement les ressources du projet sur la base des représentations de leurs attributs géométriques ou visuels. Dans le cadre du relevé d'art pariétal, cette démarche s'inscrit dans le moment de figuration – c'est-à-dire l'extraction d'éléments signifiants et leur qualification. Il s'agit d'une étape très importante, particulièrement dans le cadre d'études pluridisciplinaires où la capacité à concentrer les informations autour d'un intérêt commun est déterminante dans la compréhension de la complexité de l'objet étudié.

6.2.1. Annotation de données 2D

6.2.1.1. Saisie de tracés contraints sur une image orientée

Comme nous l'avons vu précédemment, la définition d'une annotation 2D passe par le choix d'un calque et d'une image de référence. Le processus de propagation utilise comme données d'entrée les coordonnées UV des points de contours de la région tracée. Lorsque la pose de la caméra du *viewer* coïncide avec celle d'une vue maîtresse, nous pouvons proposer des modalités d'interactions en 2D, notamment des outils d'annotation de type *polyligne* ou *main levée*.

L'approche que nous proposons se base sur la technique du lancer de rayon. Celle-ci consiste à extraire la position de l'intersection d'une droite passant par le centre optique de la caméra et la position 2D du curseur avec un objet de la scène : dans notre cas le plan image de la vue maîtresse. Les positions extraites sont exprimées en coordonnées UV. À partir des déplacements du curseur guidé par l'utilisateur, il est alors possible de récupérer les points d'intersection successifs et les exploiter comme points de contrôle du tracé d'annotation (Figure 121).

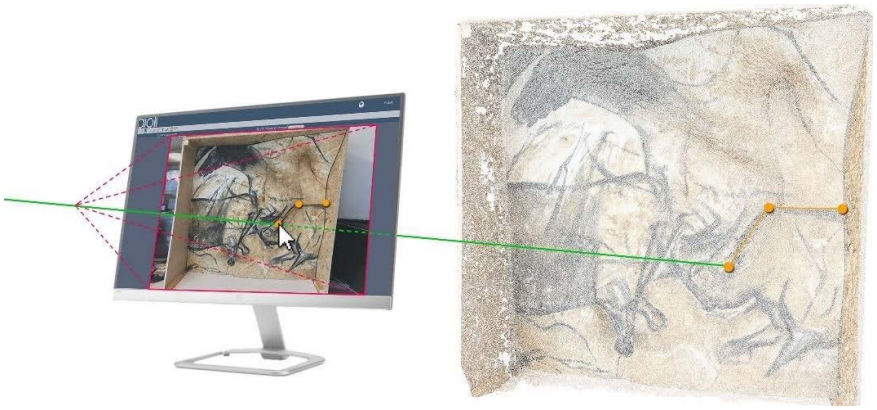


Figure 121 : Schéma du processus d'annotation 2D d'une image orientée par lancer de rayons, et la prévisualisation 3D du tracé subséquent (voir section suivante). Illustration : auteur.

Ce processus schématisé sur la Figure 122 présente l'avantage d'être indépendant du type de texture affiché sur le plan image. Ainsi, l'utilisateur peut choisir de baser son annotation sur une photographie classique, ou bien choisir d'utiliser l'une des textures auxiliaires, que ce soient les images de caractérisation morphologique, les images décorrélées, ou même une image RTI.

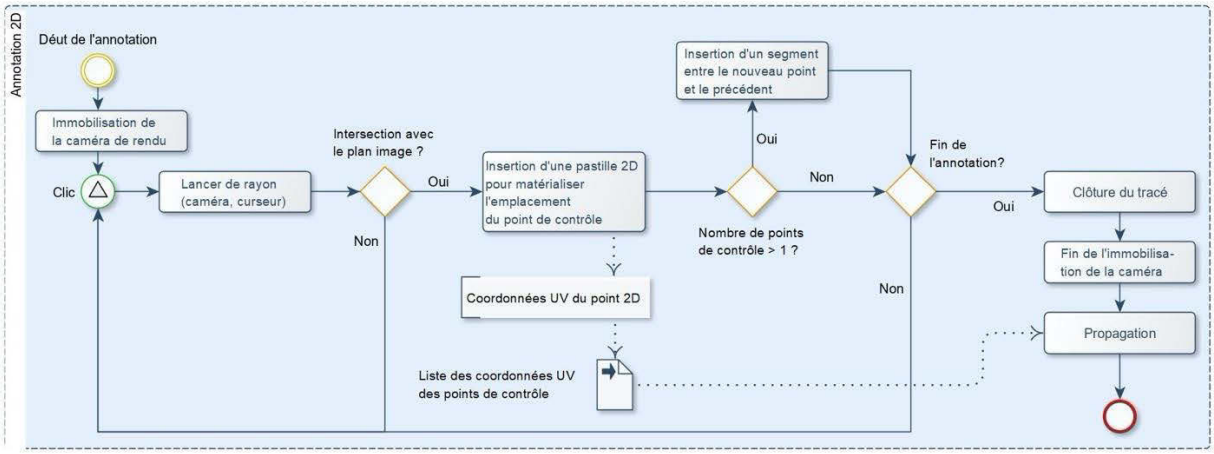


Figure 122 : Processus proposé pour l'annotation d'une ressource 2D d'un projet. Illustration : auteur.

Plus particulièrement, la possibilité de conserver à chaque instant les contrôles sur les textures signifie qu'une unique annotation peut résulter de l'interprétation de plusieurs images différentes, ou de plusieurs configurations lumineuses dans le cas d'une image RTI. En effet, cette méthode, associée aux caméras multicalques nous a bien permis de produire une annotation 3D basée sur un tracé réalisé en 2D sur une image RTI. Il s'agit d'une avancée significative dans le cadre du relevé, et particulièrement du relevé d'art pariétal, car jusqu'alors, malgré leurs potentiels indéniables, la majorité des données produites lors des phases d'acquisition demeurait sous-exploitée faute d'outils adéquats.

6.2.1.2. Prévisualisation 3D d'annotations 2D

Les images sont orientées de manière cohérente par rapport aux nuages de points du projet. Cela signifie que lorsqu'un point est sélectionné sur une image, nous pouvons connaître instantanément son équivalent 3D s'il existe, par la prolongation du rayon exploité pour les tracés 2D. Évidemment, le nuage offre une couverture de l'objet d'étude inférieure à celle des photographies, notamment parce que tous les pixels ne sont pas exploités pour produire des points 3D lors de la reconstruction photogrammétrique.

Ainsi, si ce principe demeure moins précis que le calcul des régions 3D par la propagation basée sur l'indexation des ressources, nous pouvons tout de même l'exploiter pour générer en temps réel la prévisualisation 3D des annotations 2D en cours. Par rapport au processus d'annotation présenté précédemment, cette méthode n'implique qu'une étape supplémentaire visant à récupérer, à chaque nouveau point de contrôle, les coordonnées du point 3D équivalent et à insérer dans la scène des pastilles et segments destinés à l'affichage de la région ainsi esquissée (Figure 121).

Selon le nombre de photographies maîtresses, leur recouvrement, et les propriétés de l'objet étudié, il est possible que le nuage de points comporte des zones non couvertes qui se manifestent par des densités très faibles. Ainsi, il est probable qu'un point sur une image maîtresse n'ait pas d'équivalent 3D dans un seuil de tolérance raisonnable. Dans cette configuration, l'ajout de point de contrôle 2D reste utile car le processus de propagation basé sur l'indexation permet de vérifier tous les points compris dans la région correspondante. En revanche, le lancer de rayon pour ce point de contrôle ne renverra aucun résultat 3D. Dans une telle situation, nous proposons de conserver le point 2D correspondant sans insérer de point de contrôle 3D dans la scène. Ainsi, la prévisualisation de la région en cours de propagation s'avèrera bien moins précise que le résultat attendu, mais donnera tout de même une première indication de la morphologie globale de la région annotée (Figure 123).

Figure 123 : Prévisualisation 3D d'une annotation 2D. La prévisualisation ne correspond pas tout à fait au tracé du fait du manque d'informations 3D dans la zone du cheval. Illustration : auteur.

6.2.1.3. Cohérence de la prévisualisation 3D en cas d'édition du tracé

Dans le cas de l'édition d'un tracé d'annotation avant sa propagation, par le biais de la manipulation de ses points de contrôles, nous devons actualiser la prévisualisation 3D de manière cohérente. Plus précisément, lorsqu'un point de contrôle 2D est édité, la position de la pastille 3D correspondante doit être actualisée en continu pour suivre ses déplacements. Cela n'est pas trivial, car comme nous l'avons souligné, le nuage de points comporte moins d'informations que les photographies. Pour ce faire, nous pouvons répéter la méthode de lancer de rayon à chaque *frame* jusqu'à la fin de l'édition, afin de déterminer les coordonnées successives des points intermédiaires selon la position actuelle du point 2D. Nous devons alors distinguer trois cas :

- **Cas 1 : Le point de contrôle 2D possède un équivalent 3D, au départ et à l'arrivée :**

Il s'agit du cas le plus simple. Si la fluidité du déplacement dépend de la présence de points intermédiaires, le fait que le point de contrôle 2D possède un équivalent 3D au départ comme à l'arrivée garantit quoi qu'il arrive la cohérence de la prévisualisation à l'issue de l'édition.

- **Cas 2 : Le point de contrôle 2D ne possède pas d'équivalent 3D mais est déplacé dans une zone plus dense :**

Dans cette configuration, il peut arriver qu'un point de contrôle qui ne disposait pas d'équivalent 3D initialement soit déplacé dans une zone de densité supérieure. Lorsque l'un de ces points est en cours d'édition, nous pouvons tout de même recommencer à tester l'intersection avec le nuage de points 3D. Ainsi, lorsque le lancer de rayon renvoie pour la première fois un résultat, un point de contrôle 3D peut être inséré dans la prévisualisation, et le tracé 3D peut être actualisé en conséquence. Cela nous renvoie alors au cas 1 ou au cas 3.

- **Cas 3 : Le point de contrôle 2D possède un équivalent 3D mais est déplacé dans une zone non couverte :**

Dans cette configuration, le point de contrôle 2D en cours d'édition possède déjà un équivalent 3D, mais est déplacé dans une zone moins dense du nuage de points. Dans cette situation, le lancer de rayon exécuté à chaque *frame* permet de déterminer des positions intermédiaires. Nous pouvons donc placer le point de contrôle 3D correspondant sur sa dernière position connue. Cette situation est la plus susceptible de générer des erreurs importantes de prévisualisation.

6.2.2. Annotation de données 3D

En plus des images 2D, le nuage de points 3D peut lui aussi être exploité pour la saisie de nouvelles annotations. Dans certaines configurations, il peut s'avérer plus approprié car il rend mieux compte des dimensions morphologiques de l'objet étudié. Dans ce cas, l'utilisation d'un outil d'annotation n'implique pas le gel des contrôles de la caméra. Au contraire, ceux-ci sont nécessaires pour permettre à l'utilisateur de naviguer dans la scène et manipuler l'objet 3D comme bon lui semble.

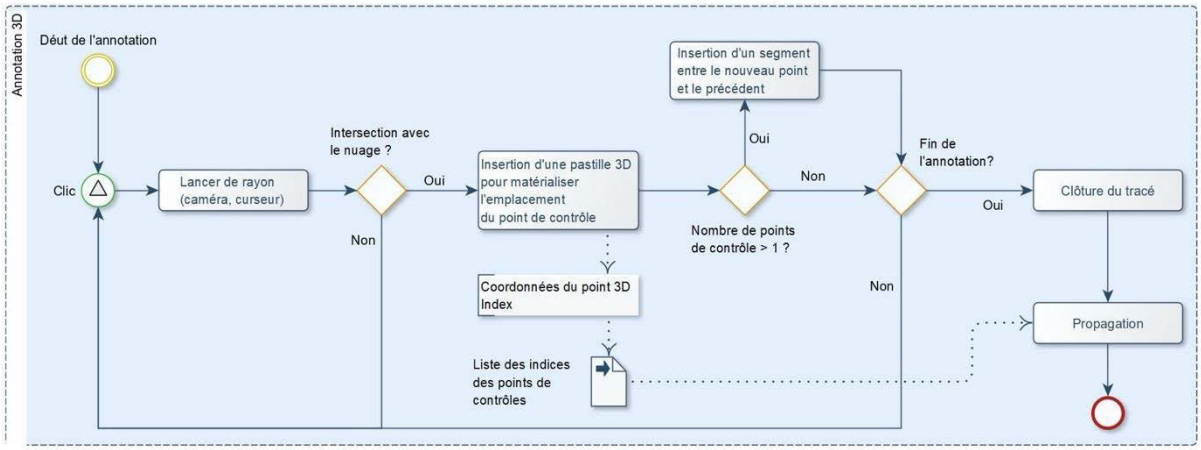


Figure 124 : Processus proposé pour l'annotation d'un nuage de points 3D d'un projet. Illustration : auteur.

Le processus proposé pour l'annotation 3D est similaire (Figure 124), il repose lui aussi sur le lancer de rayon. À chaque clic – ou chaque déplacement dans le cas de l'outil *main-levée* – nous cherchons à tester l'intersection avec le nuage de points, afin de vérifier si un point 3D existe à cet emplacement. Si c'est le cas, un point de contrôle 3D peut être inséré dans la scène.

Cependant, contrairement à l'annotation 2D, dans cette situation nous ne cherchons pas à obtenir des coordonnées UV, mais bien les nombreux attributs de points stockés dans le nuage, en particulier la

position et l'index, qui sont à la base des calculs de propagation. La difficulté repose principalement sur la capacité à conserver l'ordre des points de contrôle lors de la propagation, afin de garantir la cohérence des contours des régions 2D et 3D résultantes.

6.2.3. Perspectives : vers de nouvelles modalités d'annotation

6.2.3.1. Annotation en réalité augmentée

Dans le cas d'une utilisation en réalité augmentée, l'analyse de l'objet d'étude peut passer par l'assignation d'éléments de l'environnement qui n'auraient pas encore fait l'objet d'une acquisition 2D ou 3D. En outre, les approches actuelles en termes de numérisation s'intéressent principalement à l'élaboration de représentations des éléments de la surface de l'objet étudié. Or, il serait très restrictif de limiter un objet d'étude à son enveloppe physique, particulièrement dans le cas du patrimoine architectural, pour lequel la spatialité est une question majeure qui impose de raisonner en volumes. Dès lors, il peut être utile de proposer des outils d'annotation permettant de réaliser des tracés librement positionnés dans l'environnement et associés à des descriptions sémantiques[42]. Nous pouvons décomposer ce problème en deux aspects à traiter :

- **Création d'une géométrie et son positionnement dans l'environnement :** dans la mesure où l'annotation AR ne peut s'appuyer sur les ressources du projet qui ne couvrent pas nécessairement l'intégralité de l'environnement de l'objet d'étude, il s'agit d'étudier les modalités permettant de créer une géométrie sans exploiter ces supports, et de garantir leur positionnement dans le repère du projet.
- **Description sémantique de l'objet généré :** le fait que les géométries générées ne s'appuient pas sur les ressources du projet soulève des questions quant à leur description sémantique. En effet, pour garantir la possibilité de croiser efficacement les niveaux d'observation selon l'approche de Messaoudi (2017), nous devons préserver une certaine homogénéité dans les descriptions géométriques et sémantiques. Or, la nature des objets supportant les concepts décrits diffère, puisqu'il s'agit de numérisations du réel dans le cas usuel, et de géométries isolées dans le cas AR.

6.2.3.1.1. Création d'une géométrie et positionnement dans l'environnement

Vis-à-vis de ce problème, nous pouvons proposer différentes stratégies. Une première approche consisterait à permettre à l'utilisateur de réaliser une annotation en 2D, contrainte sur le plan image de la pose active de son dispositif d'affichage. Ainsi, lors de la validation de l'annotation, nous pourrions récupérer la *frame* active ainsi que le tracé, et réaliser un calcul de résection spatiale sur l'image extraite afin de l'orienter par rapport aux autres images du projet, puis propager l'annotation de manière classique comme tracé 2D (Figure 125).

[42] Il s'agit là d'une fonctionnalité prospective à l'étude, qui n'est pas implémentée sur la plateforme.

Figure 125 : À gauche : principe de l'annotation 2D basée sur la *frame* active du flux vidéo. À droite : exemple d'implémentation, basé sur le fonctionnement actuel d'Aïoli. Illustration : auteur.

Cette approche permet de garantir l'homogénéité des annotations et ainsi, leur comparabilité. Cependant, elle présente une limite notable. En effet, si la *frame* extraite est trop éloignée du jeu d'images maîtresses du projet, l'annotation ne pourra être propagée. Par ailleurs, elle peut s'avérer délicate à implémenter puisqu'elle exige de prendre garde aux déplacements inopinés de l'utilisateur.

Une seconde approche consisterait à créer un objet de type `Spline` suivant les positions successives du dispositif d'affichage (Figure 126). Cette approche semble plus adaptée à une utilisation en réalité augmentée car elle profite davantage de la liberté de mouvement permise par ce mode de visualisation. Cependant, elle risque de manquer considérablement de précision et de s'avérer contraignante en termes d'expérience utilisateur. De fait, la déambulation dans les grottes peut être restreinte pour des raisons de conservation. Les déplacements dans la grotte Chauvet Pont-d'Arc par exemple doivent nécessairement suivre des passerelles installées à cet effet et ne peuvent s'en écarter.

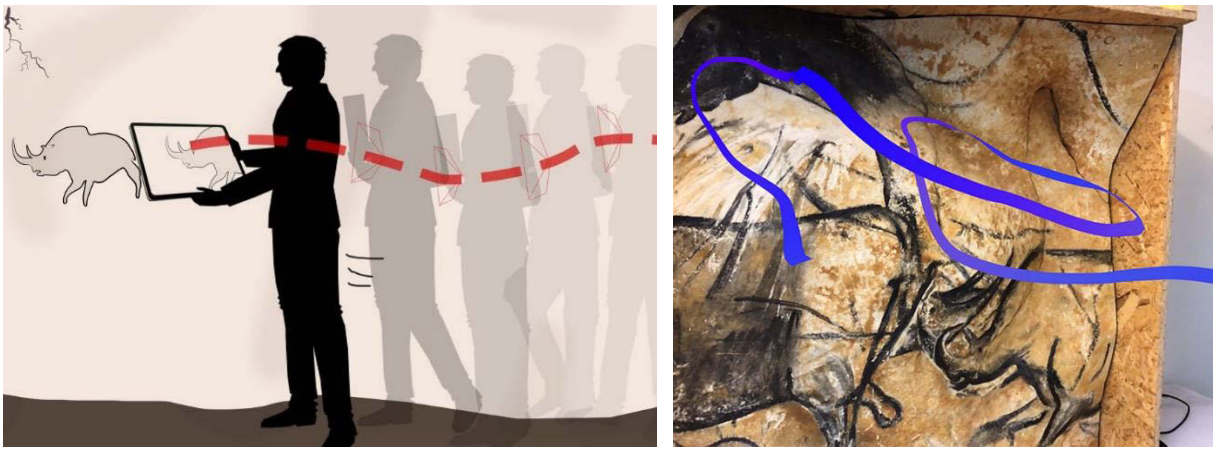


Figure 126 : À gauche : principe de l'annotation dépendant des déplacements de l'utilisateur à l'aide d'une Spline. À droite : exemple d'implémentation, à partir des démonstrations de WebARonARKit. Illustration : auteur.

Enfin, une dernière approche serait de s'appuyer sur des ancres de l'environnement pour créer et positionner la géométrie (Figure 127). Le placement d'objets virtuels dans le monde réel est un cas commun d'utilisation de la réalité augmentée. Pour positionner un objet de manière cohérente par rapport à la géométrie de l'environnement réel, à un endroit choisi par l'utilisateur, l'approche la plus courante est d'utiliser le lancer de rayon. La recherche d'ancres via cette technique est plus communément nommée *hit-test*. Étant donné que les systèmes de réalité augmentée n'exploitent pas

tous les mêmes éléments de compréhension de l'environnement (points 3D, plans, maillages, etc.), le WebXR se propose de considérer le *hit-test* comme une fonctionnalité interne de l'API. Ce fonctionnement permet d'harmoniser et simplifier le développement d'applications AR. Concrètement, cette approche consiste à lancer un rayon lorsque l'utilisateur touche l'écran tactile. Le rayon a pour origine le centre optique de la caméra, et sa direction est calculée à partir des coordonnées écran du point touché par l'utilisateur. Si ce rayon intersecte un ou plusieurs éléments issus du suivi AR, la fonction retourne un tableau contenant les ancres concernées, ordonnées par distance croissante.

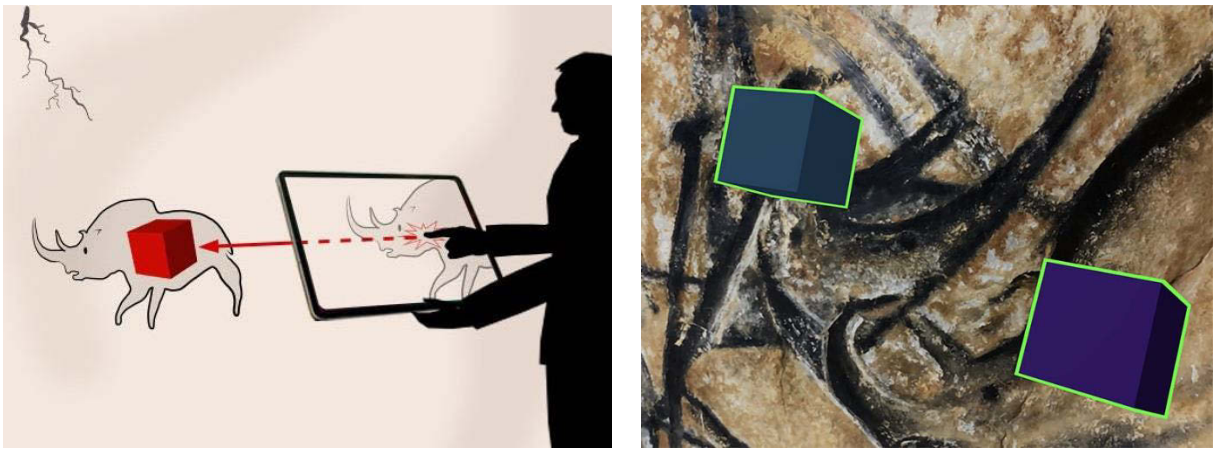


Figure 127 : À gauche : Principe de l'annotation par *hit-test*, dépôt d'une géométrie solide sur une ancre issue du suivi. À droite : Exemple d'implémentation, à partir des démonstrations de WebXR Viewer - polyfill. Illustration : auteur.

Globalement, cela revient à considérer la géométrie du monde réel comme un support magnétique contraignant le tracé, comme dans les logiciels usuels de dessin ou de modélisation 3D. Cette approche est intéressante dans la mesure où elle nous permet d'envisager une grande variété formelle pour les annotations, que ce soient des géométries solides simples (par exemple un cube ou sphère) pour lier une description à un unique point 3D, ou des tracés de type polyligne ou polyèdre pour des éléments plus complexes.

6.2.3.1.2. Description sémantique de l'objet généré

En ce qui concerne la description sémantique de l'objet généré, nous pouvons une fois de plus considérer deux cas. Si la géométrie permet d'une manière ou d'une autre la propagation, alors sa description pourrait être réalisée comme pour n'importe quelle région 2D ou 3D à l'aide des descripteurs sémantiques du volet propriétés. Dans le cas contraire, nous devons associer des descriptions aux tracés, sans pouvoir garantir que ces éléments pourront être comparés aux autres ressources du projet.

Ce cadre implique de penser une structure de données spécifique aux contenus générés via la modalité d'annotation AR, mais offre en contrepartie une certaine liberté quant à la nature des descriptions. Par exemple, en plus des modèles sémantiques usuels basés sur du texte (tags, attributs, ontologies), nous pouvons nous intéresser plus profondément au modèle par relations, en offrant par exemple la possibilité à l'utilisateur de *déposer* un commentaire audio, c'est-à-dire d'enregistrer une description vocale à localiser dans l'espace.

6.2.3.2. Annotation semi-automatique basée sur une présegmentation

Comme nous l'avons vu tout au long de ce travail, l'intérêt majeur du fonctionnement de la plateforme Aïoli repose sur la capacité à exploiter la relation projective établie entre les ressources 2D et 3D. Jusqu'à présent, nous l'avons utilisée pour propager automatiquement sur différentes ressources des annotations réalisées manuellement sur une donnée d'entrée. En pratique, ces relations géométriques ouvrent de nombreuses autres perspectives, notamment parce qu'elles permettent également d'établir des liens entre la segmentation des données et la création de calques d'annotations.

6.2.3.2.1. Processus proposé

Les méthodes de segmentation visent à subdiviser des données numériques, que ce soient des images, des nuages de points, ou des maillages. Il s'agit de dégager les différents éléments constitutifs d'une donnée d'entrée pour permettre leur traitement, analyse, ou visualisation, indépendants. Le principe se base sur le regroupement d'éléments partageant une propriété commune, habituellement colorimétrique ou géométrique. Cependant, s'il existe aujourd'hui des méthodes de segmentation efficaces pour les images 2D, la segmentation automatique de nuages de points 3D constitue une opération plus délicate, en partie en raison des difficultés liées à la nature non structurée des données mais également de leur précision (irrégularité, bruit, aberrations, etc.). Dans ce cadre, la relation projective 2D/3D constitue une piste intéressante, car les éléments issus d'une segmentation peuvent tout à fait faire l'objet de régions, ou à l'inverse, des régions issues d'annotations manuelles peuvent servir de base à l'élaboration d'un modèle segmenté.

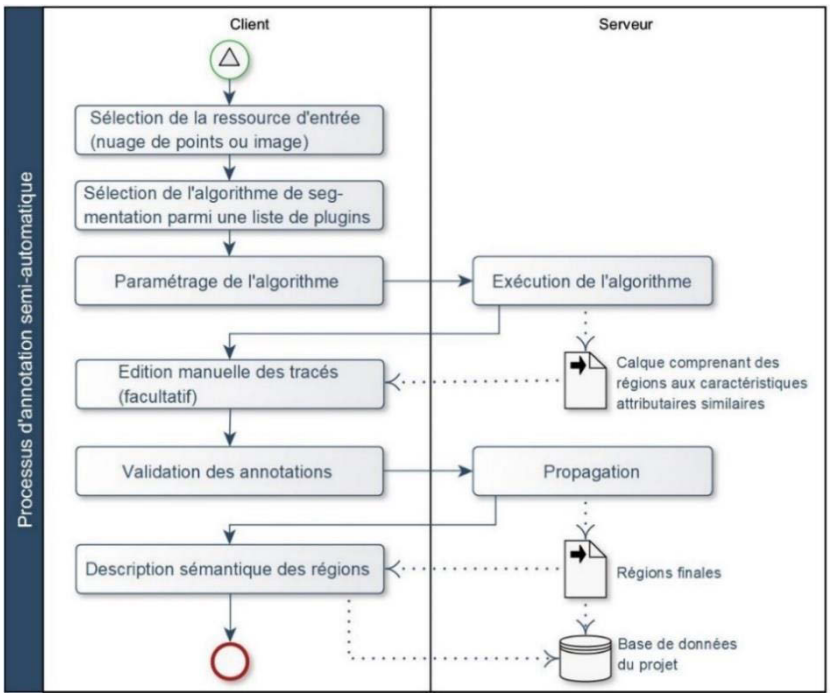


Figure 128 : Processus proposé pour l'annotation semi-automatique basée sur des algorithmes de segmentation. Illustration : auteur.

L'idée est d'appliquer un algorithme de segmentation 2D sur une image maîtresse, et de projeter les régions obtenues pour subdiviser le nuage de points correspondant. Les éléments issus de cette segmentation peuvent être exploités pour générer automatiquement des régions réparties dans un calque.

Les régions 3D correspondantes sont stockées sous la forme de nuages de points indépendants, qui peuvent donc être traités ultérieurement de manière isolée. À l'inverse, une segmentation opérée sur le nuage de points 3D peut aussi être reprojetée sur les ressources 2D du projet afin de produire rapidement de grandes quantités d'images segmentées et annotées.

En pratique, la flexibilité de l'architecture d'Aïoli basée sur l'utilisation de conteneurs permet d'envisager l'implémentation d'une grande diversité d'algorithmes de segmentation, à la manière de *plugins*. Cela nous permettrait de proposer une méthode générique d'annotation semi-automatique paramétrable, telle que schématisée sur la Figure 128.

Pour vérifier la faisabilité de cette approche, nous l'avons mise en pratique en exploitant un algorithme de segmentation de nuages de points en éléments plans. Le but de l' expérimentation, décrite par Manuel et al. (2018) est de vérifier la possibilité d'exploiter un algorithme de segmentation pour générer un calque d'annotations de manière automatique. D'un point de vue mathématique, le plan est un élément géométrique relativement simple à identifier[43]. Cette méthode de segmentation est basée sur une approche de type *split-and-merge* ou « séparation et fusion ». Elle consiste en deux étapes itératives :

- ***Split*** : Cette étape consiste à structurer les points du nuage de points d'entrée sous la forme d'un octree, dans lequel les voxels contiennent les points satisfaisant un même critère de raffinement. Dans notre cas, le critère de raffinement est satisfait si la distance moyenne des points par rapport à leur meilleur plan est inférieure à un seuil préalablement défini par l'utilisateur. À partir d'un nœud contenant tous les points, chaque itération subdivise les différents nœuds en testant, pour chacun d'eux, le critère de raffinement jusqu'à ce que tous les nœuds le satisfassent. Ce processus permet de construire un octree dont chaque voxel contient les points identifiés comme coplanaires.

- ***Merge*** : Cette étape consiste à regrouper les points coplanaires en régions. Les régions sont construites de manière itérative à partir des points d'un voxel en ajoutant progressivement les éléments des voxels voisins tant que le critère de fusion est satisfait. Dans cette étape, le critère évalué est l'angle entre la normale de la région actuelle et la normale moyenne des points du voxel testé. Si cet angle est inférieur à un seuil préalablement défini, les points du voxel testé sont ajoutés à la région.

Le résultat de cette opération donne donc des points 3D répartis en régions différentes correspondant aux plans identifiés dans le nuage de points d'entrée. Ces régions peuvent être reproduites sur chaque image maîtresse du projet par le biais de l'indexation (Annexe 11). À l'issue de ce processus, nous pouvons générer automatiquement un calque d'annotation contenant les différentes régions détectées lors de la segmentation. En ce qui concerne l'enrichissement sémantique, la responsabilité incombera tout de même à l'utilisateur de décrire les régions identifiées.

Cette expérimentation nous a donc bien permis de valider la faisabilité de notre approche. Cependant, comme nous l'avons constaté dans les chapitres précédents, dans le cas de l'art pariétal la segmentation automatique est une opération délicate du fait de la complexité des objets d'étude et de leur unicité. Pour cette raison, nous nous sommes également intéressés à la définition d'un outil d'annotation semi-

[43] Dans cette expérimentation, les critères de raffinement ont été choisis pour détecter des éléments coplanaires, mais leur modification pourrait permettre de détecter d'autres primitives géométriques.

automatique basé sur l'exploitation des attributs géométriques et visuels, calculés lors du processus d'indexation.

6.2.3.2.2. Annotation par baguette magique

Il existe dans des logiciels d'édition d'image des outils très répandus d'assistance à la sélection qui pourraient s'avérer très pertinents pour l'annotation dans le cadre du relevé d'art pariétal. C'est notamment le cas des outils de type *baguette magique* ou des outils basés sur le magnétisme comme le lasso. En effet, nous avons vu dans les chapitres précédents que les éléments observables sur les parois présentent souvent des caractéristiques communes que l'on peut mettre en valeur à travers différents traitements. Des peintures pariétales réalisées avec les mêmes pigments sont ainsi caractérisées par des pixels de mêmes couleurs sur des images décorrélées, et des points coplanaires se traduisent par des pixels de mêmes couleurs sur des cartes de normales.

Or l'outil *baguette magique* permet de sélectionner un ensemble de pixels sur une image à partir d'une valeur de référence et d'un seuil de tolérance, c'est-à-dire un ensemble de données aux propriétés constantes au regard d'un paramètre. Appliqué aux ressources d'un projet, cet outil de sélection basé sur un simple critère colorimétrique pourrait en réalité permettre de sélectionner rapidement des points aux caractéristiques géométriques similaires. Pour reprendre l'exemple précédent, cet outil sélectionne automatiquement un plan lorsqu'il est appliqué sur une carte de normales, avec un seuil bien paramétré (Figure 129).

En outre, comme le montre Adeline Manuel (2016) dans sa thèse, nous pourrions aller plus loin, en complexifiant légèrement les paramètres. Si au lieu d'un seuil nous fournissons une fonction à l'outil de sélection, nous pourrions non seulement extraire des données constantes, mais aussi une évolution prévisible décrite par des profils mathématiques. Par exemple, à partir d'une carte de normales, l'utilisation d'une fonction de la forme $f(x) = Seuil + x$, appliquée à l'axe X, nous permet de sélectionner automatiquement des courbes d'un simple clic. Dans ce cadre, les différents attributs morphologiques calculés lors de l'indexation révèlent tout leur potentiel.

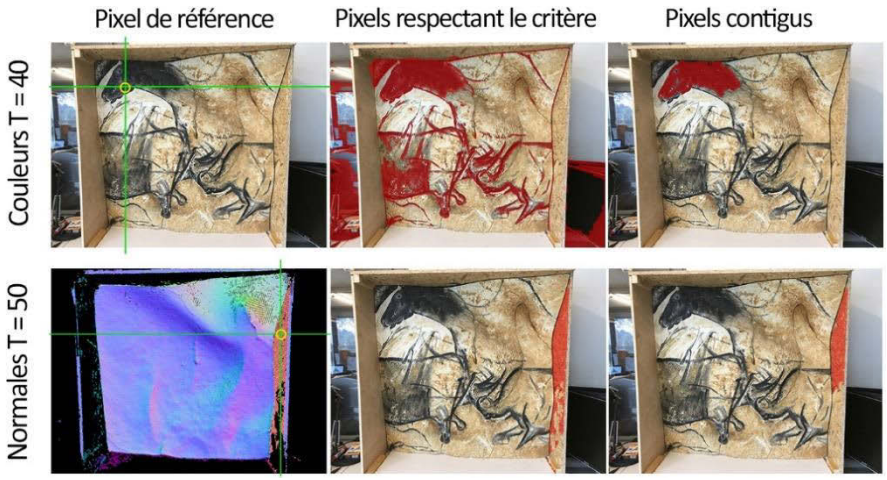


Figure 129 : Potentiel d'un outil de sélection par seuillage appliqué aux attributs colorimétriques ou aux normales. T correspond à la valeur de tolérance. Illustration : auteur, d'après (Manuel et al., 2018).

Si ce principe est initialement pensé pour des images, nous pouvons proposer une méthode similaire pour sélectionner des points d'un nuage 3D. Cela suppose en premier lieu de disposer d'un outil

permettant de spécifier un point 3D dont les valeurs attributaires constituent la référence de la sélection, ce que nous pouvons faire par un lancer de rayon. La valeur de référence peut alors être choisie parmi les différents attributs colorimétriques du point (voir section 6.1.2.1), voire même être définie comme une combinaison de plusieurs attributs. Le seuil de tolérance, compris entre 1 et 100, doit être spécifié par l'utilisateur avant la sélection. Pour sélectionner automatiquement l'ensemble des points 3D dont les valeurs attributaires sont égales à la valeur de référence plus ou moins le seuil, nous pouvons envisager deux approches.

- **Par la fonction de coloration du *shader* :**

La première consiste à exploiter la fonction de coloration des points dans le *shader*, car tous les points du nuage rendus dans le *viewer* passent nécessairement par celle-ci (voir section 7.2.3.2, Annexe 12). L'avantage de cette approche est qu'elle permet de tester rapidement chaque point, et sauvegarder l'index du point s'il satisfait le critère de sélection (Figure 130).

En contrepartie, cette approche exige un travail de post-traitement. En effet, les points sélectionnés ne sont pas structurés, ce qui ne permet pas de limiter simplement la sélection aux points contigus. Cela impose également de passer par une fonction supplémentaire pour transformer les éléments sélectionnés en régions et en extraire les différents contours. Une autre limite est que cette approche impose de tester systématiquement tous les points affichés à l'écran, même lorsque la sélection désirée ne concerne qu'une portion restreinte du nuage. Ainsi, si le test du critère d'attribut peut être réalisé relativement rapidement, le fonctionnement global de cette approche peut nuire aux performances d'affichage en augmentant la latence.

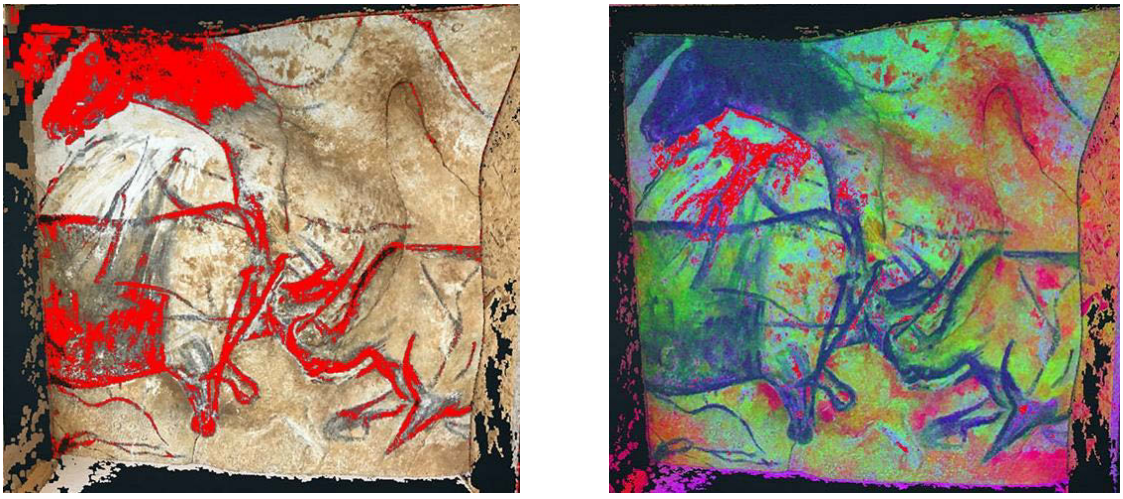


Figure 130 : Outil baguette magique 3D via la fonction de coloration du *shader* (les points sélectionnés sont colorés en rouge). À gauche : sur la texture RGB du nuage de points. À droite : sur la texture décorrélée. Illustration : auteur.

- **Par la relation projective 3D – 2D :**

Pour concevoir une baguette magique 3D sans nuire aux performances du client, une deuxième approche consiste à réduire le problème à un problème 2D. En effet, quelle que soit la modalité de visualisation (*desktop* ou AR), la scène est toujours rendue en 2D. Nous pouvons donc réaliser simplement des sélections à partir de l'image de rendu du nuage de points plutôt que sur le nuage en lui-même (Annexe 13). L'approche est la suivante (Figure 131) :

- Lors de la sélection du pixel de référence par un clic sur l'écran ①, nous pouvons procéder au lancer de rayon 2D → 3D ② pour récupérer l'index du point 3D correspondant ③ ;
- À partir de la valeur colorimétrique du pixel, tel qu'affiché à l'écran, nous procédons à la sélection de type *baguette magique* directement sur l'image de rendu ④ ;
- Si le point de vue change (déplacement de la caméra de rendu), nous pouvons repartir du point 3D de référence ⑤ pour réaliser un lancer de rayon inverse 3D → 2D jusqu'au centre optique de la caméra ⑥. Cette étape permet de récupérer les nouvelles coordonnées écran du point 3D de référence ⑦ ;
- Les attributs colorimétriques du pixel obtenu à l'étape précédente sont utilisés comme nouvelle valeur de référence pour actualiser la sélection ⑧ ;
- Lorsque l'utilisateur valide la sélection, nous sauvegardons la liste des points de contours de la région 2D constituée. Pour chaque point de contour, un nouveau lancer de rayon 2D → 3D permet de déterminer son index. Comme dans le cas d'une annotation, nous pouvons alors tracer une prévisualisation de la région. À l'issue de cette étape, nous disposons de la liste ordonnée des indices des points 3D du contour. Nous pouvons donc procéder à la propagation, comme dans le cas d'une annotation 3D classique.

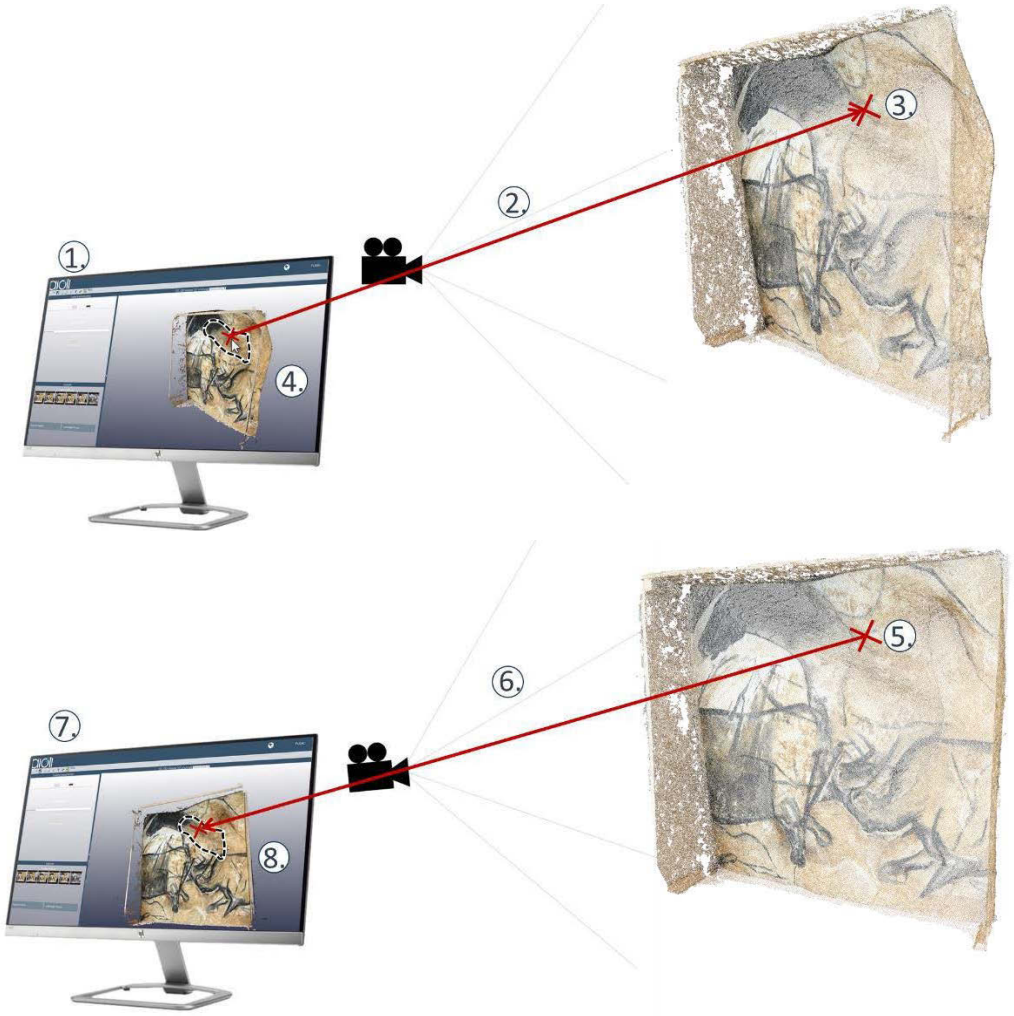


Figure 131 : Schéma du principe de l'outil *baguette magique* 3D. En haut : initialisation. En bas : actualisation de la sélection en cas de déplacement de la vue. Illustration : auteur.

Cette approche présente l'intérêt d'être plus rapide concernant l'affichage, et permet de restreindre très simplement la sélection aux pixels contigus. En contrepartie, le réglage du seuil et des paramètres d'affichage du *viewer* est plus délicat pour éviter de prendre en compte les autres éléments de la scène. Nous pourrions contourner cette limite en testant systématiquement la présence de points 3D correspondant aux contenus de la région sélectionnée, mais cela demanderait un temps de calcul très conséquent qui nuirait considérablement à l'affichage.

Par ailleurs, le recours à de multiples projections 2D → 3D → 2D pour déterminer les coordonnées écran successives du point 3D de référence cause des approximations qui introduisent des erreurs. Ainsi, lors d'un mouvement de la caméra, la sélection peut changer notablement d'aspect (Annexe 14). Enfin, pour que cette méthode constitue véritablement une alternative à la sélection 3D, il faudrait que chaque sélection successive entre deux mouvements de caméra soit sauvegardée, afin que la région finale soit obtenue par union booléenne des sélections intermédiaires.

Cette approche présente des résultats encourageants, en particulier au regard des enjeux liés au relevé patrimonial ou d'art pariétal. Avec de bons paramétrages, un tel outil de sélection permet de tracer une région en un clic, sur la base d'attributs géométriques ou visuels. Ainsi, selon la texture active, nous pouvons sélectionner des régions contenant des points isochromatiques, coplanaires, isohypses, isobathes, etc. L'ajout d'options permettant de réaliser des opérations booléennes sur ces régions permettrait en outre de réaliser des sélections multicritères, c'est-à-dire partageant plusieurs attributs – fonctionnalité naturellement prise en charge dans l'approche précédente.

6.3. Codification graphique des annotations

6.3.1. Paramétrage de matériaux d'annotation

Dans les Chapitres 1et 2, nous avons pu observer que le traitement graphique accordé au relevé constitue un aspect important : c'est l'intelligibilité du document résultant qui est en jeu, le traitement graphique constituant une clé de lecture destinée à accompagner les documents produits. Pour cette raison, il est important d'offrir une certaine liberté dans le paramétrage de l'aspect des différentes annotations.

Comme nous l'avons vu précédemment, notre proposition comprend le chargement des annotations 2D dans le *viewer* comme sous forme de `shapes`, dont la géométrie comprend deux éléments distincts : le tracé du contour, et le remplissage. Le tracé du contour correspond à une polyligne dont l'aspect est régi par différentes propriétés : la couleur, l'épaisseur du tracé, la présence ou non de pointillés, et le cas échéant, la longueur des tirets et l'espacement entre ceux-ci. L'aspect du remplissage de l'annotation dépend essentiellement de sa couleur et son opacité, mais peut également exploiter des textures externes, par exemple pour créer des motifs répétitifs.

Ce principe, que l'on retrouve dans tous les logiciels de traitement de données vectorielles, se prête donc assez bien à l'utilisation de conventions graphiques, tel que dans les approches de Delluc et al. (1986a), et plus généralement dans les relevés analytiques techniques. Dans le cas du relevé analytique plastique, cette approche présente des limites notables, car le principe d'annotation que nous proposons ne permet que de définir les contours de régions aux aspects homogènes, ce qui ne permet pas d'aborder l'expression graphique des annotations de manière nuancée. Notons cependant que cette opération peut

être abordée comme un post-traitement extérieur à la plateforme, par exemple, par le biais d'un export SVG et d'un import dans un logiciel de traitement plus adapté comme Inkscape ou Illustrator.

Figure 132 : Adaptation de la couleur et de l'opacité d'une région 3D pour l'harmonisation des matériaux d'annotation 2D et 3D. La prise en charge des propriétés de contours et des textures de remplissage est plus complexe.

Une autre limite repose sur les différences de natures entre les régions 2D et 3D (Figure 132). En effet, les annotations 3D sont traitées comme des nuages de points, dont l'aspect est principalement régi par le paramétrage de la taille des points, leur couleur et leur opacité. Pour homogénéiser pleinement les représentations 2D et 3D des régions, il faudrait donc tenir compte également des paramètres des contours et des motifs de remplissage. Le premier point impliquerait le chargement, pour chaque région 3D, d'une polyligne suivant les points de contrôle 3D et partageant les propriétés graphiques des `shapes`. Le second point est plus délicat, car il impliquerait de réaliser des projections planaires des textures externes, ce qui implique de déterminer un axe de projection pertinent pour la région observée. Cela n'est pas trivial car cela dépend grandement des morphologies des objets d'étude qui peut être très complexe dans le cas des grottes.

6.3.2. Génération d'une légende graphique

Une légende graphique est une image utilisée pour établir un lien entre les tracés d'une représentation et les informations sous-jacentes qu'ils supportent. Dans le cas d'un relevé, la légende associe donc à un élément graphique l'interprétation principale de la figure, telle qu'elle est établie par l'auteur. Sa communication est essentielle pour limiter les erreurs d'interprétation, particulièrement dans le cas d'études pluridisciplinaires.

Dans notre cas, la codification graphique des annotations est gérée par le paramétrage des matériaux selon l'approche énoncée précédemment. Pour établir une légende graphique à partir d'un projet, nous devons donc dans un premier temps constituer la liste des matériaux utilisés dans le projet, ce qui peut être fait à partir de la liste des régions de chaque calque. Nous devons ensuite proposer une méthode permettant de rendre sur une seule image l'ensemble de ces matériaux, chacun associé à un intitulé (Figure 133).

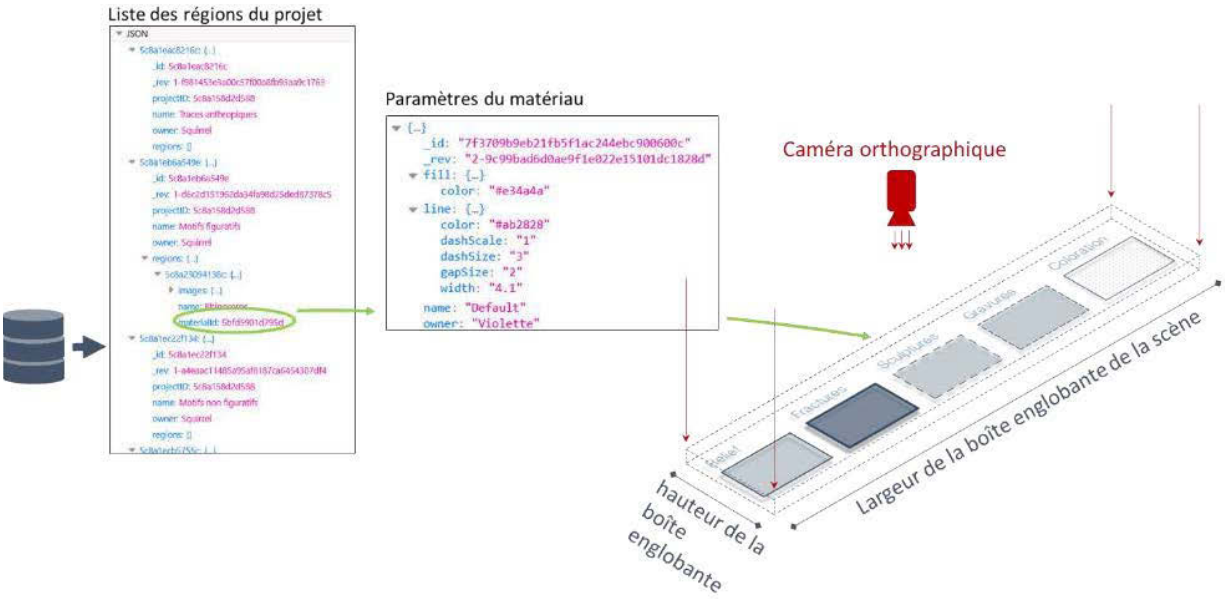


Figure 133 : Schéma du principe du processus de rendu automatique d'une légende graphique à partir des données de projet. Illustration : auteur.

Pour cela, notre approche consiste à créer une scène dédiée, nommée $Scene_{Legendes}$. Munis de la liste des matériaux exploités dans un projet, nous pouvons peupler $Scene_{Legendes}$ d'autant de `shapes` dotés d'une géométrie élémentaire que le projet ne comprend de matériaux différents. À côté de chacun d'entre eux, le nom du matériau correspondant peut être affiché, également sous la forme de `shape` pour éviter le recours à de multiples éléments `HTML`. Enfin, la scène peut être rendue à l'aide d'une caméra orthographique, puis l'image générée être retournée à l'utilisateur pour être téléchargée.

L'unique difficulté (relative) de cette méthode réside dans la disposition des éléments de la scène et le choix des dimensions de l'image à rendre. Cet obstacle peut être levé simplement à partir des boîtes englobantes des contenus de la scène. Notons que ce fonctionnement impose aux utilisateurs de prévoir une dénomination explicite dès la création du matériau. Une amélioration possible serait de remplacer la description automatique du matériau, actuellement basée sur son nom, par une description libre réalisée à l'aide d'un formulaire.

6.4. Bilan

En tenant compte des apports des différents types de ressources manipulées lors d'un relevé et des possibilités d'hybridation des environnements réels et virtuels, nous avons proposé dans ce chapitre des méthodes prospectives d'augmentation d'attributs géométriques et visuels destinés à faciliter les processus d'interprétation et d'annotation. Dans un premier temps nous nous sommes intéressés à la production d'images de caractérisation morphologiques puis à l'élaboration de traitements 2D, 3D, et vidéo temps-réel. Enfin, nous avons proposé différents outils d'annotation multimodaux permettant d'enrichir sémantiquement les données recueillies.

Les méthodes présentées dans ce chapitre reposent sur la spatialisation des ressources impliquées dans un relevé numérique, et plus particulièrement sur la maîtrise des relations projectives entre les différentes représentations de l'objet d'étude. Dans cette optique, il semble que la proposition de relevé

multimodal que nous présentons dans cette thèse constitue une avancée notable quant aux contraintes d'éloignement des parois et aux questions de justesse géométrique qui en découlent.

Du point de vue de la codification graphique, notre approche semble bien se prêter au relevé analytique technique du fait de la proximité des modalités de représentation exploitées et de la structure par calques. Dans le cas du relevé analytique plastique, elle présente des limites concernant l'expression graphique des observations, dont le traitement n'est envisagé qu'en tant qu'opération ultérieure au relevé, nécessairement extérieure à la plateforme. Cela signifie que dans un tel cadre, les trois moments de Saint-Aubin ne sont pas vraiment réunis dans un moment unique.

En reprenant les critères d'évaluation exploités dans les premiers chapitres de ce document, nous pouvons comparer notre approche aux méthodes de relevé précédemment étudiées. D'un point de vue théorique, le relevé multimodal tend logiquement à concentrer les atouts du relevé numérique et du relevé analytique technique, mais présente des lacunes au niveau de la représentation, conformément à ce que nous venons d'énoncer (Figure 134).

En outre, le critère d'intelligibilité du document résultant est un point délicat à évaluer car les modalités de visualisation et d'interaction proposées par notre approche basée sur l'hybridation d'environnements réels et virtuels tendent à requestionner la pertinence des paradigmes de représentation planimétrique. Quoi qu'il en soit, pour vérifier cette évaluation d'un point de vue pratique, il serait absolument nécessaire de tester notre approche sur le terrain en conditions réelles auprès des utilisateurs.

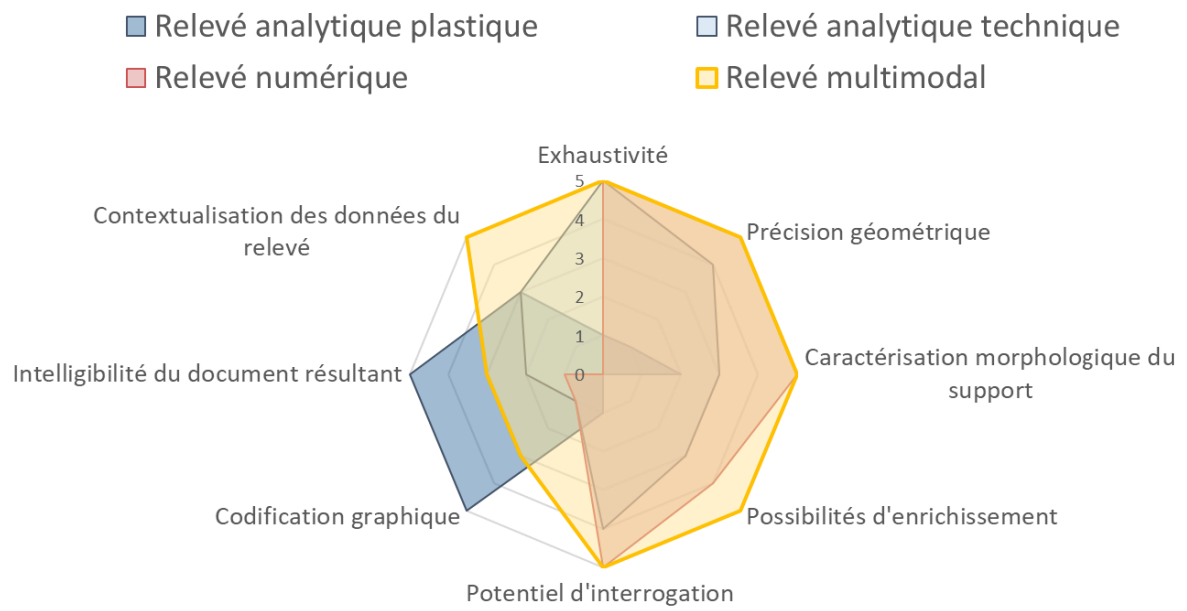


Figure 134 : Comparaison du relevé multimodal et des différentes approches de relevé d'art pariétal. Illustration : auteur.

Cela passe par l'implémentation des méthodes énoncées précédemment, afin d'élaborer une preuve de concept permettant d'évaluer en premier lieu la viabilité technique de notre proposition, puis méthodologique. Nous présenterons dans le chapitre suivant les développements informatiques liés à l'élaboration de cette preuve de concept ainsi que nos résultats.

Chapitre 7. Implémentation informatique de la preuve de concept

7.1. De la réalité augmentée sur le web avec l'API WebXR

Comme nous l'avons énoncé précédemment, pour la partie expérimentale de notre travail notre choix s'est porté sur le WebXR. Il est important de noter que cette API n'est pas encore stable, et que, si une première ébauche a bien été publiée en février 2019 par le W3C, augurant d'un standard à venir, ce n'était pas encore le cas au moment de nos développements. Il ne s'agissait alors que d'expérimentations menées en parallèle par deux équipes, Google-AR d'une part, et Mozilla™ d'autre part (Medley, 2018 ; Bergstrom, 2018). Leurs arguments et les avancées de ces travaux respectifs étaient alors suffisamment convaincants pour présager d'une piste prometteuse pour l'avenir, ce que les publications récentes de l'*Immersive Web Community Group* ont finalement confirmé (W3C, 2018). En effet, aujourd'hui le WebXR est pris en charge de façon (très) expérimentale dans plusieurs navigateurs fonctionnels, en particulier Google Chrome, et Microsoft Edge (Deveria, 2019).

Dans le cas qui nous intéresse, c'est-à-dire prioritairement les mobiles, les navigateurs compatibles sont des navigateurs expérimentaux, au nombre de trois (Figure 135) : deux navigateurs expérimentaux destinés aux mobiles dont le système d'exploitation est iOS et étant compatibles avec ARKit : ***WebARonARKit*** (Google-AR, 2017b)[44] et ***WebXR Viewer*** (Mozilla, 2017), et un navigateur expérimental pour les mobiles sous Android et étant compatibles avec ARCore : ***WebARonARCore*** (Google-AR, 2017a)[42].

Figure 135 : Navigateurs prenant en charge le WebXR, de manière expérimentale.

Il est important de signaler que ces applications ne sont pas des navigateurs web complets et comprennent **un nombre très réduit** de fonctionnalités. Ils se destinent essentiellement aux développeurs souhaitant explorer les capacités de la XR sur le web. Par ailleurs, ces travaux étant des prototypes, ils évoluent rapidement et parfois radicalement[45]. Pour des raisons pratiques, notre choix s'est porté sur le *WebXR Viewer* proposé par Mozilla[46], que nous avons exploité sur des iPad Pro. Ses fonctionnalités ont évolué depuis la fin de nos expérimentations, pour la suite de ce chapitre, sauf mention contraire explicite, nous baserons nos explications sur la version utilisée : version 1.12 du 29 novembre 2018.

[44] Il ne s'agit pas d'un produit officiel de Google ni d'un navigateur complet. Il s'agit uniquement d'une application destinée à permettre aux développeurs d'expérimenter.
[45] WebARonARKit a été mis de côté chemin faisant par l'équipe qui le développait au profit d'autres projets.
[46] En effet, WebARonARCore est construit à partir de la première version de ARCore Developer Preview, ce qui limite considérablement les appareils compatibles. Le choix du WebXR Viewer permettait donc d'utiliser un matériel plus accessible. Par ailleurs, la gamme des appareils compatibles inclut des tablettes dotées de plus grands écrans.

À lui seul, le WebXR ne suffit pas à proposer des contenus en réalité augmentée pour le web. Pour cela, il doit être utilisé conjointement avec le WebGL, une API JavaScript permettant l'affichage d'éléments 2D et 3D dans n'importe quel navigateur web compatible sans nécessiter l'utilisation de modules complémentaires. Dans notre cas, nous avons basé le développement du *viewer* hybride sur les bibliothèques Three.JS et Potree.JS. Le principe de l'application AR peut se schématiser de la façon suivante (Figure 136) :

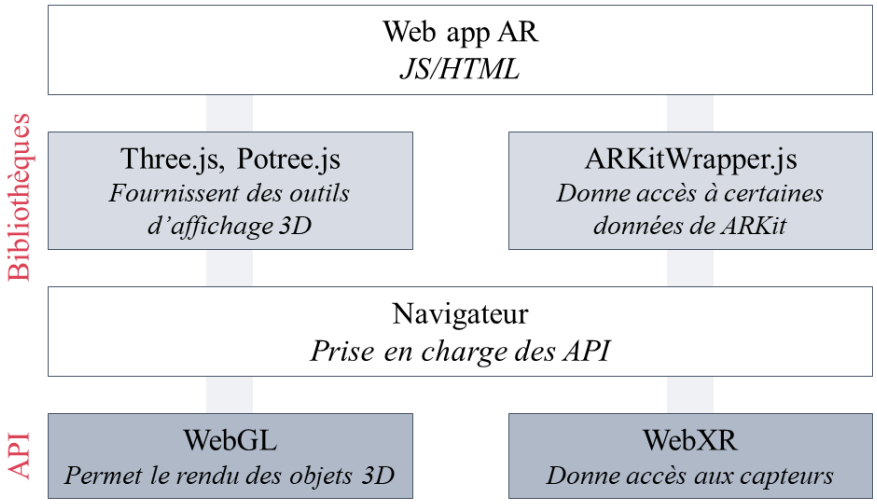


Figure 136 : Représentation schématique du fonctionnement d'une application basée sur les API WebXR et WebGL Illustration : auteur.

7.2. Implémentation du *viewer* hybride

7.2.1. Interface

Avant d'implémenter le *viewer* et les différentes modalités d'interaction selon les approches détaillées dans les chapitres précédents, nous devons préparer l'interface web destinée à les accueillir. Étant donné nos objectifs, il apparaît clairement que cette interface doit être *responsive* afin de s'adapter aux différentes utilisations de la plateforme, qu'elles soient bureautiques ou mobiles. L'interface que nous proposons est une page web HTML5/JavaScript comprenant 6 éléments (Annexe 8) :

1. Le bandeau supérieur, qui comprend les différents onglets menant aux outils de manipulation, annotation, et navigation ;
2. Le panneau latéral des outils, comprenant les éléments GUI permettant d'accéder aux différents outils et les paramétrer ;
3. Le panneau latéral des images, qui comprend les vignettes de toutes les images maîtresses du projet et permet leur sélection ;
4. Le panneau latéral des calques, qui offre une représentation hiérarchique des calques et annotations existantes ;
5. Le panneau latéral des propriétés, qui présente les descripteurs sémantiques associés à l'annotation ou au calque actif ;
6. L'espace de visualisation, au centre, c'est-à-dire le *viewer*.

En outre, nous pouvons distinguer deux cas d'utilisation, selon les modalités d'interactions choisies : *desktop* (ou *hors site* dans le cas d'une utilisation bureautique ou mobile *classique*), et utilisation XR (ou *sur site* dans le cas d'un affichage en réalité augmentée). Dans le cas d'une utilisation *desktop* la page comprend deux couches, le canvas dans lequel est rendue la scène 3D et les éléments de l'interface utilisateur (GUI).

Dans le cas d'une utilisation XR, il faut également rendre les images de l'environnement réel provenant de la caméra. Pour cela, la scène augmentée est donc rendue dans deux canvas superposés : le canvas inférieur pour rendre chaque *frame* du flux vidéo, et le canvas supérieur pour rendre la scène 3D à l'aide des bibliothèques WebGL. Au-dessus de ces éléments se trouve le GUI (Figure 137) :

Figure 137 : Composition de la page HTML selon le mode de visualisation adopté. Illustration : auteur.

7.2.2. Structure du *viewer*

Comme nous l'avons évoqué précédemment, le WebGL est une API graphique issue d'OpenGL ES proposée par le Khronos Group, permettant d'afficher directement des contenus 3D dans une page web via la balise HTML5 `<canvas>`. Cette API est supportée par tous les navigateurs modernes.

Three.JS est une bibliothèque JavaScript *open-source* très populaire et richement documentée, procurant des outils qui simplifient considérablement l'utilisation du WebGL, notamment en ce qui concerne la création de scènes, de caméras virtuelles, de géométries 3D, ou encore de matériaux. Elle permet également l'utilisation de *shaders* en glsl (*Open GL Shading Language*) pour la création d'effets de rendus avancés. Cette fonctionnalité est à la base de la bibliothèque Potree.JS, un complément de Three.JS permettant la visualisation de grands nuages de points. Le fonctionnement global du rendu du *viewer* repose alors sur la structure suivante (Figure 138).

La scène 3D est un contenant comprenant l'ensemble des objets 3D peuplant l'environnement virtuel. Dans $Scene_{Objects}$, dont le rendu repose sur Three.JS, ces objets sont des `Meshes`, c'est-à-dire des maillages constitués d'une géométrie (composée de sommets, d'arrêtes, et de faces), et d'un matériau doté de différentes propriétés qui régissent son rendu (couleur, texture, transparence, etc.).

Dans le cas de $Scene_{Pointclouds}$, dont le rendu repose sur Potree.JS[47], l'affichage des nuages de points ne passe pas par des `Meshes` mais par des `Octrees`, composés de `Nodes` (ou nœuds), eux-mêmes constitués de `Points`. L'aspect visuel des points est géré par un *shader* agissant sur leur position et leur géométrie à l'écran (*vertex shader*) et leurs propriétés colorimétriques (*fragment shader*).

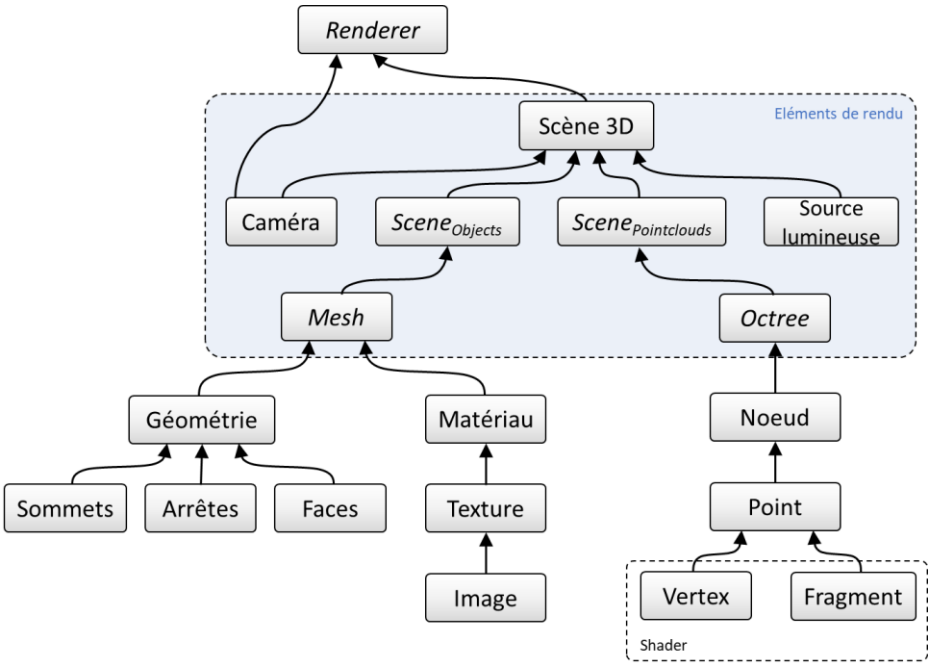


Figure 138 : Fonctionnement global du *viewer*, unifiant les fonctionnements de Three.JS et Potree.JS pour rendre des ressources 3D diverses au sein d'une unique scène 3D. Illustration : auteur.

L'ensemble est rendu par un `Renderer`, c'est-à-dire un moteur de rendu, à l'aide d'une caméra virtuelle et des sources lumineuses. Le modèle de caméra détermine les propriétés projectives utilisées pour générer la vue finale apparaissant à l'écran (orthographique, perspective, projection cylindrique ou sphérique, etc.). Les sources de lumière correspondent quant à elles aux modèles standards utilisés en infographie (lumière ponctuelle, spot, lumière ambiante, etc.).

7.2.3. Rendu temps réel

7.2.3.1. Initialisation

En termes d'implémentation, nous pouvons distinguer deux processus principaux : la phase d'initialisation du *viewer*, et la boucle de rendu (Figure 139).

L'initialisation du *viewer* comprend deux visées : d'une part la création des éléments indispensables à l'élaboration d'un rendu WebGL (canvas, caméra, moteur de rendu, etc.), c'est-à-dire l'initialisation du processus de rendu, et d'autre part la création des différents éléments de la hiérarchie selon la structure présentée en (section 5.2.1), qui devront nous permettre de manipuler simplement les ressources des projets.

[47] Pour ce travail, nous avons préalablement produit une version allégée de la bibliothèque Potree.JS en l'adaptant à nos besoins, notamment en supprimant toutes les fonctionnalités non nécessaires pour nos expérimentations et en modifiant la structure afin de simplifier les contrôles sur les contenus. Sauf mention contraire explicite, c'est à cette version « allégée » que nous ferons référence dans les sections qui suivent.

Les grandes étapes de l'initialisation sont résumées en Annexe 4. Cette phase est générique et indépendante des contenus à afficher, c'est pourquoi elle peut être déclenchée dès l'ouverture du projet, c'est-à-dire dès le début de la session d'utilisation.

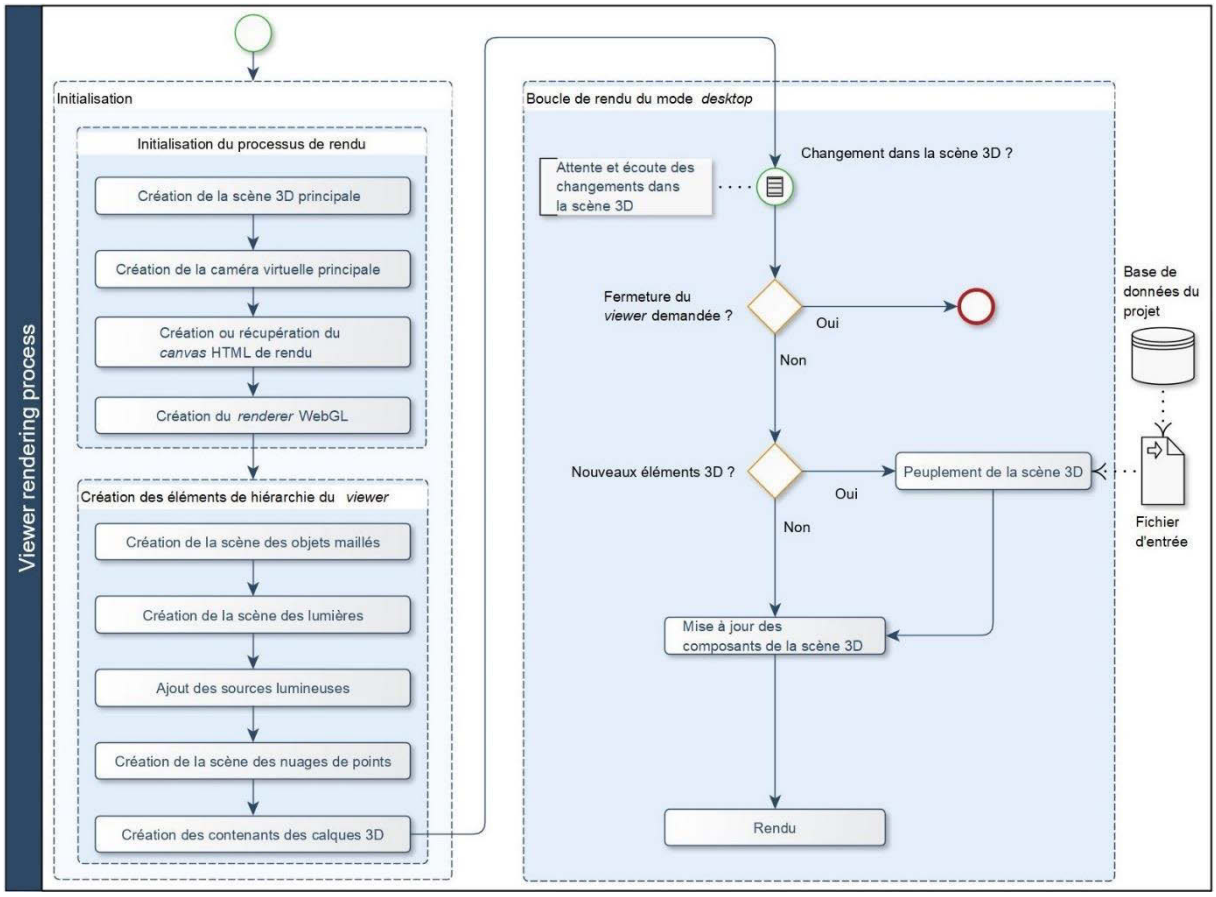


Figure 139 : Processus de rendu de la scène contenant les ressources d'un projet. Illustration : auteur.

Lorsque l'initialisation est achevée, boucle de rendu est déclenchée et doit être exécutée à chaque *frame*. Elle répond à trois objectifs principaux :

- **Écouter les actions de l'utilisateur** sur les éléments du GUI et mettre à jour les composants de la scène (position, aspect, propriétés) en conséquence. Par exemple il peut s'agir de mettre à jour l'opacité du nuage de points principal lorsque l'utilisateur agit sur le *slider* associé ;
- **Peupler la scène** dans le cas où de nouvelles ressources ont été ajoutées. Par exemple il peut s'agir de charger les données du projet lorsqu'elles sont disponibles (nuages de points, et modèles de caméras correspondant aux images maîtresses), ou encore charger une nouvelle annotation à l'issue d'une propagation ;
- **Permettre la fermeture du projet** si l'utilisateur en fait la demande, et le cas échéant, supprimer les éléments concernés de la page HTML et réinitialiser les variables.

7.2.3.2. Peuplement et rendu

Les fonctions permettant de peupler et rendre les éléments de la scène 3D dépendent du type de ressources manipulées. Comme nous avons pu le voir précédemment, nous distinguons en particulier

les ressources de type nuages de points et les ressources assimilables à des maillages (images orientées, mesures, etc.). Les ressources de $scene_{Objects}$ étant toutes spatialisées relativement aux nuages de point globaux des projets, nous détaillons en premier lieu les méthodes impliquées dans l'affichage des contenus de $scene_{Pointclouds}$.

7.2.3.2.1. Données 3D

7.2.3.2.1.1. Structure en octree

Étant donné la quantité d'informations contenues dans ce fichier, qui dépend notamment de l'ampleur du projet (nombre et dimensions des images maîtresses), l'affichage du nuage de points brut dans le *viewer* causerait d'importants problèmes de performance. Pour cette raison, la bibliothèque Potree.JS développée par Markus Schüetz à l'Université d'informatique de Vienne (Autriche) propose une méthode permettant d'adapter continuellement le niveau de détails du rendu du nuage de points, en fonction de la position de l'observateur.

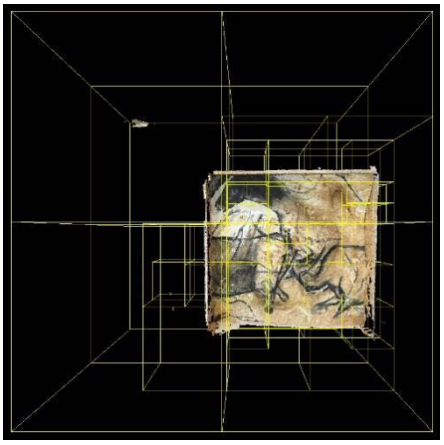


Figure 140 : Répartition des points du nuage du facsimilé dans l'octree.

Cette méthode s'appuie sur un octree. Il s'agit d'une structure arborescente de données dans laquelle l'ensemble est subdivisé de manière récursive en nœuds, qui peuvent compter jusqu'à huit enfants (Figure 140). La position de la caméra virtuelle détermine les nœuds devant être affichés à l'écran, ce qui permet d'alléger considérablement le rendu (Schuetz, 2014). La qualité visuelle du nuage de points, et en particulier de son niveau de détails, dépend donc de la sélection opérée pour répartir les points du nuage initial dans les différents nœuds de l'octree. Une sélection aléatoire de points résulte en un sous-ensemble non uniforme, qui comporte des trous dans certaines zones et des densités trop importantes dans d'autres.

Pour cette raison, lors de la conversion, une sélection uniforme est opérée afin que la distance entre les points affichés soit constante, ce qui améliore nettement la qualité visuelle du rendu. Pour cela, le programme utilise une grille 3D dont la dimension correspond à celle de la boîte englobante du nuage de points, et dans laquelle les dimensions des cellules (largeur, hauteur, profondeur) correspondent à une distance minimale prédéfinie. Suite à cela, le programme calcule pour chaque point l'index de la cellule qui le contient, et vérifie l'état des cellules voisines :

- Si toutes les cellules voisines sont vides, le point est stocké dans le nœud ;
- Si au moins l'une des cellules voisines n'est pas vide, le programme vérifie les distances séparant le point testé des points voisins. Si ces distances sont supérieures à la distance minimale prédéfinie, le point est stocké dans le nœud. Sinon, il est écarté.

En fonction de leur position, les points restants sont divisés en huit nœuds enfants du nœud actif, et la distance minimale est réduite de moitié. Ce processus est répété de manière récursive sur chaque nœud jusqu'à ce que tous les points du nuage de points initial aient été traités.

La conversion du nuage de points brut en octree est réalisée à l'aide d'un programme nommé PotreeConverter, écrit en C++. Ce programme est stocké dans l'un des conteneurs du serveur de traitement, et exécuté à l'issue du processus d'indexation.

7.2.3.2.1.2. Préservation des attributs de caractérisation morphologiques calculés par Aïoli

Pour prendre en charge dans le *viewer* les attributs de caractérisation morphologique calculés lors de l'indexation et stockés dans nos fichiers PLY comme attributs colorimétriques *customs* (voir section 6.1.2.1), nous devons modifier ce programme et assurer que ces derniers soient préservés lors de l'écriture des fichiers de l'octree.

À partir du fichier d'entrée, PotreeConverter procède à la conversion et produit un fichier binaire par nœud, ainsi qu'un fichier de métadonnées générales au format JSON spécifiant les propriétés du nuage, et en particulier ses attributs. Pour la conversion, seules les coordonnées euclidiennes de chaque point sont indispensables. PotreeConverter prend tout de même en charge certains attributs optionnels comme les couleurs, intensités, ou les normales.

Pour réaliser la conversion, dans un premier temps le programme lit le `header` du fichier PLY d'entrée pour déterminer les propriétés de chaque point et leurs ordres d'apparition. Ensuite, pour chaque point du fichier d'entrée, une instance de la classe `Point` est créée et en récupère les coordonnées euclidiennes ainsi que les propriétés prises en charge par Potree.JS. Cela concerne donc les propriétés dont les noms et types sont reconnus par Potree.JS et dont la sauvegarde est explicitement demandée par l'utilisateur. À l'issue du processus de répartition des points en octree, les fichiers définitifs sont écrits (Figure 141).

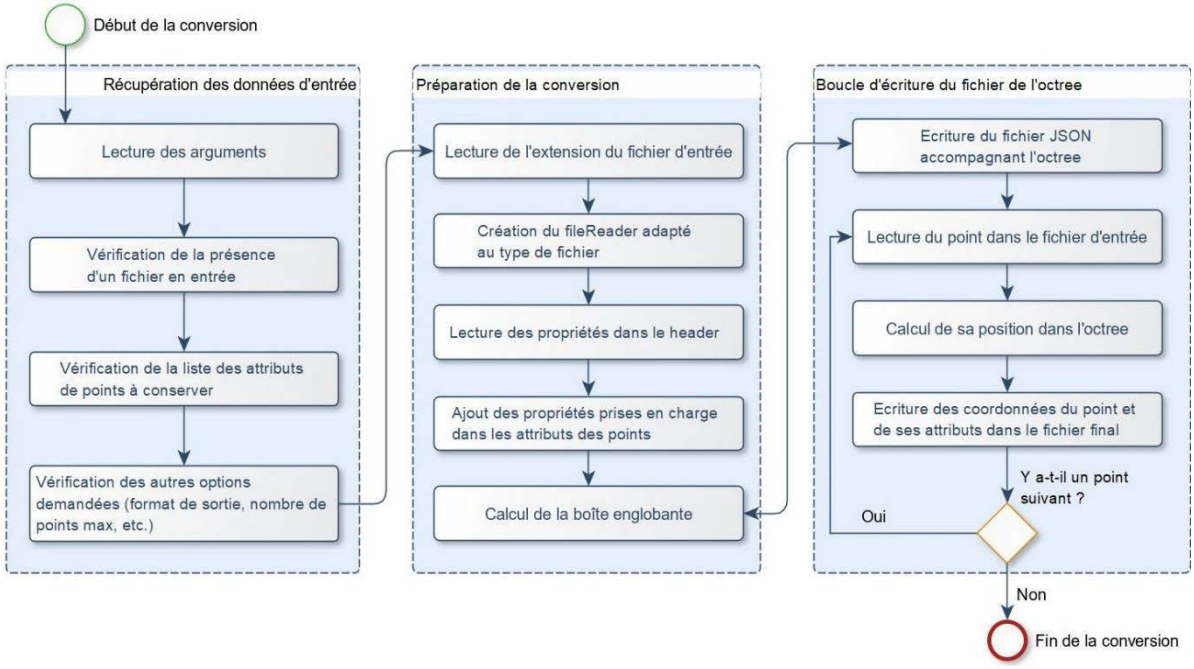


Figure 141 : Déroulement général du processus de conversion d'un nuage de points en octree avec PotreeConverter. Illustration : auteur.

La prise en charge de nos nouvelles propriétés de points est donc relativement simple à implémenter, elle demande principalement les modifications suivantes :

- Pour chaque nouvelle propriété, ajout d'un argument permettant à l'utilisateur de demander sa sauvegarde ;
- Ajout des nouvelles propriétés dans la liste des attributs pris en charge et spécification de leurs noms et types ;
- Ajout des cas correspondant aux nouvelles propriétés dans la fonction de lecture du fichier.

Une fois ces modifications réalisées, nous avons préparé un nouveau conteneur dédié à la conversion en octree, hébergé sur le serveur de traitement.

7.2.3.2.1.3. Chargement sélectif et rendu des nœuds

Comme nous l'avons évoqué, la structure en octree est particulièrement intéressante pour permettre le rendu de grands jeux de données lorsqu'elle est associée à un affichage sélectif des nœuds, basé sur la détection des nœuds situés dans le cône de vision de la caméra. On parle de *view frustum culling* (Figure 142). Pour améliorer la visualisation, Potree.JS propose une phase de vérification visant à rendre conjointement plusieurs nœuds de l'octree qui se chevauchent pour augmenter le niveau de détail, les nœuds les plus proches de la caméra étant favorisés par rapport aux nœuds éloignés. Un budget de points défini par le client limite le nombre de points affichés simultanément à l'écran, afin d'adapter le *viewer* aux capacités des différents matériels utilisés (ordinateurs, tablettes, *smartphones*, etc.).

À l'issue de cette phase, l'application dispose de la liste des nœuds à charger et dont le rendu est prévu. Pour éviter le chargement et le décodage de trop nombreux nœuds, Potree.JS procède en établissant une hiérarchie, qui consiste à charger en priorité les cinq nœuds dont la taille projetée à l'écran est la plus importante.

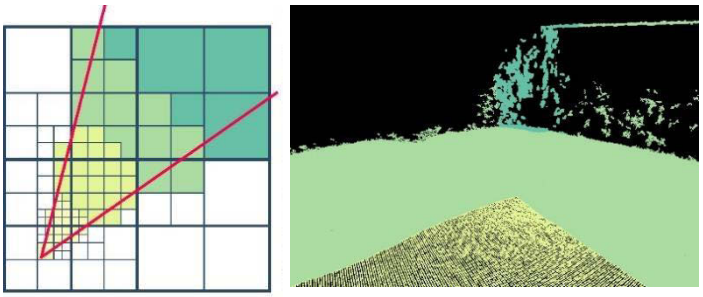


Figure 142 : Affichage sélectif des nœuds de l'octree basé sur le cône de vision de la caméra. Les couleurs correspondent aux différents niveaux de détails liés à la proximité de la caméra. Illustration : auteur, d'après (Schuetz, 2016).

Le rendu des points est réalisé sous forme de carrés ou de cercles (Figure 143). Pour améliorer l'apparence du rendu, Potree.JS se base sur l'algorithme de Botsch et al. (2005) pour fusionner les points se superposant et dont les profondeurs par rapport à la caméra sont semblables. Une phase de normalisation permet d'adoucir les transitions entre les points superposés. Cette méthode permet de gérer les phénomènes d'occlusion, les valeurs de profondeur appliquées dans le *shader* influençant la capacité à voir à travers les surfaces.

Figure 143 : Rendu des points sous forme de carrés ou de cercles par Potree.JS.

La couleur des points affichés à l'écran dépend des attributs stockés dans les fichiers de l'octree. Seuls les attributs pris en charge par le *viewer* sont préservés lors du décodage des fichiers, pour être ensuite communiqués au *fragment shader*, dont le rôle est de déterminer la couleur du rendu de chaque point par le biais de fonctions spécifiques. Dans le cas des nuages de points produits lors de la création d'un projet, l'attribut colorimétrique principal est bien évidemment le triplet RGB, issu des couleurs des pixels provenant des images maîtresses.

Pour prendre un exemple volontairement très simpliste[48], si la variable `inputcolor` correspond à un vecteur contenant ces trois valeurs RGB du point à rendre telles qu'elles sont décodées lors du chargement, et `outputcolor` un vecteur contenant le triplet de valeurs RGB de la couleur à utiliser pour l'affichage de ce point, la fonction suivante (Figure 144) permettrait au *shader* d'assigner au point les attributs de couleurs directement tels qu'ils sont reçus en entrée.

```
FONCTION getRGB (attribution de la couleur du point à afficher)

Entrée : vecteur d'entiers inputcolor contenant les attributs RGB du point


DÉBUT

      Déclarer un vecteur d'entiers outputColor

      outputColor = (inputColor.r, inputColor.g, inputColor.b)

      RETOURNER outputColor

FIN
```

Figure 144 : Tous les attributs de points sont communiqués au *shader*, qui peut les traduire en informations colorimétriques pour le rendu.

L'intérêt majeur de ce fonctionnement est qu'il autorise une grande flexibilité dans la manipulation des attributs des points, notamment pour ce qui concerne la traduction d'un attribut exprimé sous la forme d'une valeur mesurée en une valeur colorimétrique permettant son affichage.

En reprenant notre exemple, si au lieu de colorer le nuage de points à l'aide des couleurs originales nous souhaitions le colorer en niveaux de gris correspondant à la luminance de chaque point, nous pourrions proposer une variante de la fonction précédente. Il s'agirait alors simplement de lire les valeurs RGB du point telles qu'écrites dans le fichier, les multiplier par des constantes selon le modèle CIE RGB 1931 (CIE, 1931) et les sommer (Figure 145).

Une fois le nuage chargé et les attributs communiqués aux *shader*, des éléments du GUI doivent permettre à l'utilisateur de modifier la valeur d'une variable dédiée pour spécifier le mode de coloration désiré pour le rendu des points. Ce choix est communiqué au *shader* qui appelle la ou les fonction(s) correspondante(s) en fonction de la requête.

[48] En pratique, cette fonction ne serait pas très utile car il suffirait de retourner directement inputColor. Il faut voir cet exemple comme un préambule à l'exemple présenté en figure 145.

```
FONCTION getLuminance (coloration du point selon sa luminance)

Entrée : vecteur d'entiers inputcolor contenant les attributs RGB du point


DÉBUT

	Déclarer un vecteur d'entiers outputColor

	Déclarer un réel luminance

	Déclarer un entier valeurGris

	luminance = (0.17697 * inputcolor.r) + (0.81240 * inputcolor.g) +
			(0.01063 * inputcolor.b)

	valeurGris ← interpolation linéaire de luminance entre 0 et 255

	outputColor = (valeurGris, valeurGris, valeurGris)


	RETOURNER outputColor

FIN
```

Figure 145 : Ces attributs peuvent être facilement manipulés et exploités pour offrir de nouveaux modes de visualisations.

En ce qui concerne les attributs morphologiques issus de la phase d'indexation, nous avons pris le parti de les encoder comme des couleurs de points pour homogénéiser les représentations. Ainsi, au niveau du *viewer*, outre l'ajout des cas correspondant aux nouveaux attributs dans la fonction de décodage du fichier, l'affichage de ces attributs n'implique que la création de nouvelles fonctions de coloration des points dans le *shader*, sur le modèle des fonctions déjà existantes. Nous pouvons alors proposer à l'utilisateur différents modes de texturage du nuage de point : couleurs des points, normales, courbure, rugosité, occlusion ambiante, et profondeur.

7.2.3.2.1.4. Insertion des nuages de points dans la scène 3D

Potree.JS intègre nativement toutes les fonctions nécessaires pour le chargement et le décodage des fichiers binaires des nuages de points, ainsi que leur affichage. Un aspect problématique est que dans la version originale de la bibliothèque, ces fonctions ne laissent pas le choix concernant la destination de l'objet chargé. Il est obligatoirement inséré dans une scène nommée `ScenePointCloud`, elle-même enfant d'une entité de type `EventDispatcher` **improprement nommée *Scene***. Comme nous l'avons évoqué précédemment, cette structure complexe manque de flexibilité quant aux contrôles exercés sur les poses des différents éléments de la hiérarchie, principalement car l'élément parent ne se comporte pas comme un objet 3D manipulable.

Pour adapter la bibliothèque Potree.JS à nos besoins et contraintes, nous avons notamment proposé une modification de la fonction de chargement de nuages de points distinguant le processus qui relève du chargement et décodage des données, de celui qui concerne l'insertion de l'octree dans l'un des sous-groupes de la scène parente. Ce fonctionnement permet d'ores et déjà de charger et afficher tous les nuages de points impliqués dans un projet Aïoli dans le *viewer* en respectant notre structure hiérarchique, le tout étant compris dans le client web de la plateforme (Figure 146).

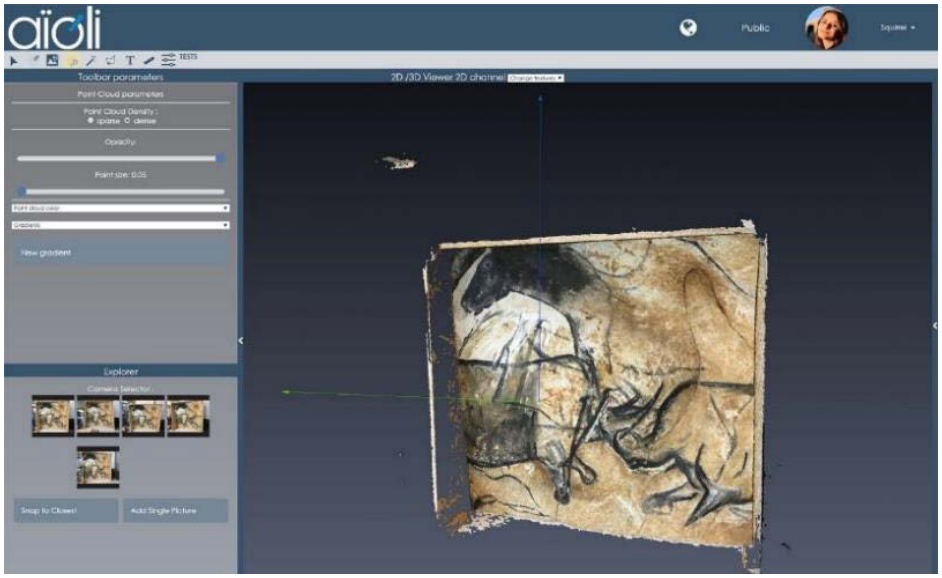


Figure 146 : Chargement du nuage de point d'un projet dans le *viewer*, en respectant la structure hiérarchique préétablie.

7.2.3.2.2. Données 2D

La majorité des images 2D impliquées dans un projet peut être traitée selon le processus présenté en (section 5.2) qui vise à créer des pyramides de vue selon le modèle sténopé à partir des données issues de calculs photogrammétriques. Celles-ci sont créées et rendues à l'aide de Three.JS comme des objets de type `mesh` (Annexe 5). Cependant, la prise en charge de certains types de données 2D demande un traitement spécifique.

7.2.3.2.2.1. Prise en charge des images RTI

Contrairement aux autres images 2D statiques, les images RTI doivent permettre aux utilisateurs de faire varier la position de la source de lumière. Nous ne pouvons donc pas les afficher comme des textures classiques, c'est-à-dire à l'aide de la classe `THREE.Texture`.

Le laboratoire *Visual-Computing Lab* du CNR-ISTI propose une bibliothèque nommée Relight.JS sous licence GNUv3 permettant la création et visualisation d'images RTI (Ponchio, Potenziani, et al., 2019). Parmi les outils proposés, la bibliothèque propose des outils permettant le chargement et l'affichage de ces images dans un canvas via l'API WebGL. En l'état, cette méthode constitue un *viewer* 2D à part entière que nous ne pouvons utiliser comme tel étant donné nos objectifs.

Cependant, nous pouvons exploiter certaines méthodes, en particulier celles dédiées au chargement et décodage des données, ainsi que les fonctions d'actualisation qui prennent en paramètre la direction du faisceau lumineux. Ainsi, nous pouvons charger et afficher l'image RTI dans un canvas. Celui-ci n'est pas affiché dans la page HTML, mais directement assigné à un matériau par le biais de la classe `THREE.CanvasTexture`, une variante de la classe de texture lui permettant en outre d'être actualisée à chaque *frame* (Figure 147).

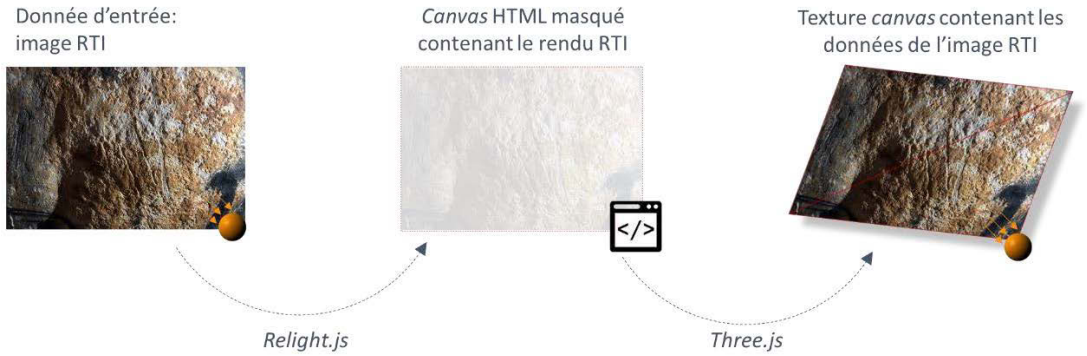


Figure 147 : Principe proposé pour l'utilisation d'une image RTI comme texture dynamique d'un plan 3D. Illustration : auteur.

La prise en charge des images RTI impose, outre ce travail sur les textures, l'ajout d'un élément de GUI dédié aux variations lumineuses. Dans les logiciels de RTI les plus répandus, il se présente sous la forme d'une sphère colorée que l'utilisateur peut faire tourner sur elle-même. En l'absence d'un *input* HTML adapté, nous avons créé une classe permettant de générer ce type d'élément de GUI par le biais d'une scène 3D indépendante comprenant une `mesh` sphérique et des contrôles orbitaux appliqués à l'objet, le tout rendu dans un canvas. La classe est dotée d'un gestionnaire d'évènements permettant de renvoyer l'orientation de l'axe X du repère local de la sphère à chaque mouvement. Nous utilisons ce même gestionnaire pour déclencher l'actualisation de la texture RTI.

Des expérimentations sur un jeu de données du facsimilé nous ont permis de valider cette méthode. Notre approche nous permet bien d'afficher dans le *viewer* des images RTI, préalablement orientées par photogrammétrie sur le serveur de traitement, puis de les utiliser comme support d'annotation tout en conservant le contrôle de la source lumineuse (Figure 148). Sur le cas d'étude testé, ce fonctionnement a facilité l'observation de fissures et le tracé des régions correspondantes, en nous permettant de modifier en cours d'annotation les paramètres d'éclairage en lumière rasante (Annexe 10).

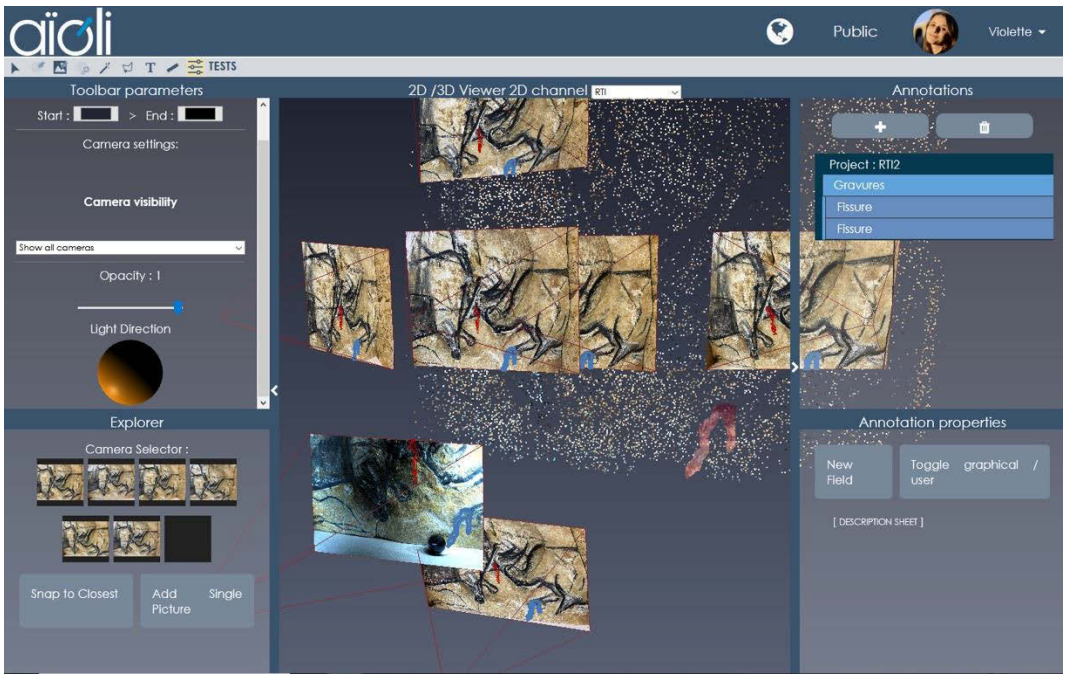


Figure 148 : Capture d'écran d'une image RTI spatialisée et annotée dans Aïoli. Illustration : auteur.

7.2.3.2.2.2. Implémentation de l'algorithme de *Decorrelation stretching* pour les images maîtresses

Pour implémenter l'algorithme de *decorrelation stretching* d'images, nous avons opté pour le développement d'une application en C++, compilée et conteneurisable sur le serveur de traitement, à l'aide de la bibliothèque OpenCV (Bradski, 2019).

La fonction de décorrélation suit la méthode détaillée en (section 6.1). Elle prend en entrée une image RGB ainsi que deux attributs spécifiant la moyenne et l'écart-type attendus en sortie pour chaque canal. Elle retourne en sortie une image décorrélée (Figure 149). En outre, la bibliothèque OpenCV propose différentes fonctions permettant de convertir une image d'un espace colorimétrique dans un autre espace colorimétrique, par exemple pour passer de l'espace RGB à l'espace LAB. Ainsi, une variante simple de cette implémentation de l'algorithme pourrait être envisagée, en ajoutant comme paramètre d'entrée l'espace colorimétrique souhaité.

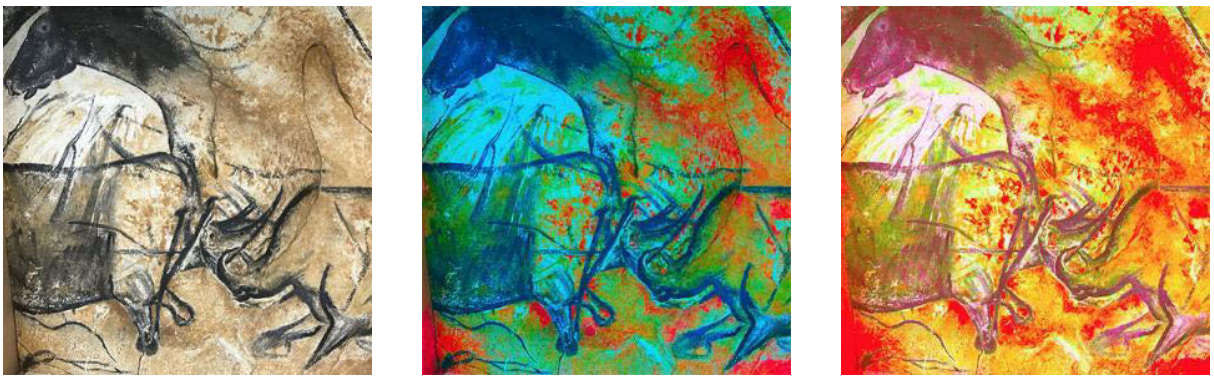


Figure 149 : Résultats de l'implémentation de l'algorithme de *decorrelation stretching* en variant les paramètres. À gauche : image originale. Au centre : image décorrélée, avec m_{target}=120. À droite : image décorrélée avec m_{target}=180.

Cette application peut également être compilée et conteneurisée sur le serveur de traitement pour permettre son utilisation directement dans Aïoli. Dans ce cas, il faut proposer la sauvegarde dans la base de données des informations concernant l'image décorrélée produite, pour tenir compte de sa présence comme nouvelle texture auxiliaire de la vue maîtresse correspondante. Pour l'instant, nous n'avons exploité cette application qu'en local à des fins d'expérimentation.

7.2.3.2.3. Spatialisation de données multimédias complémentaires

Certaines ressources manipulées lors d'un relevé ne sont pas aisément spatialisables. C'est par exemple le cas de documents textes, de vidéos, de dessins, d'images d'archives difficile à corréler, etc. Elles ne sont pourtant pas à négliger, car elles peuvent s'avérer des sources précieuses d'informations concernant l'objet étudié. Pour faire face à ces cas, nous proposons de prendre en charge dans le *viewer* les descriptions sémantiques de type *relations* (voir section 2.4.1.2). Ces descriptions permettent aux utilisateurs de spécifier des liens entre une annotation et des données complémentaires, en joignant aux descripteurs sémantiques (de types *attributs*) des entrées de type *URI* ou des fichiers multimédias téléversés sur le serveur web. Par exemple, un utilisateur peut joindre à une région des diagrammes illustrant des mesures physiques réalisées *in situ* sur cette zone.

Lorsque ce cas se présente, c'est-à-dire lorsque la fiche descriptive d'une région comprend au moins un attribut de type URI ou une donnée multimédia, nous insérons dans la scène une fenêtre contenant ces

informations (lien cliquable dans le cas d'un URI, miniature d'une image ou du lecteur vidéo/audio dans le cas de données multimédias). Cette fenêtre est une balise `<div>` positionnée dans la page HTML sur la projection en coordonnées écran du barycentre de la région 3D (Figure 150).

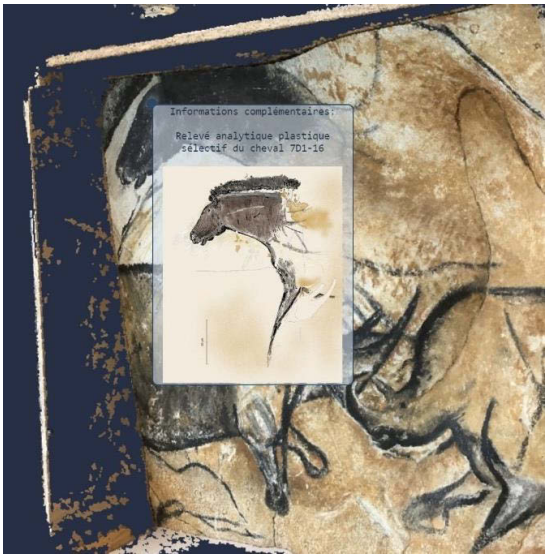


Figure 150 : Capture d'écran. Visualisation d'une ressource complémentaire non spatialisable : ancrage d'un DIV à un point 3D.

Bien entendu, l'utilisateur peut choisir de masquer ces fenêtres, notamment lorsque la scène est encombrée de trop nombreux éléments, ou de les afficher pour mieux visualiser les ressources relatives à son projet.

Ces ressources complémentaires peuvent donc être visualisées de manière spatialisée dans la scène au même titre que les autres ressources du projet. Cependant, ne bénéficiant pas de la relation projective, elles ne pourront être annotées.

7.2.4. Implémentation des outils d'annotation, mesure, et mise à l'échelle

Les outils d'annotation 2D ont été implémentés à partir de l'outil *raycast* de Three.JS, qui permet de récupérer les coordonnées UV du point du maillage impacté par un rayon. Lors de la récupération des coordonnées des points de contrôle, nous avons simplement dû prendre garde aux différents repères, car le repère local de la géométrie plane utilisée comme plan image, le repère des coordonnées UV, et le repère utilisé par le module de propagation sont tous différents (Figure 151).

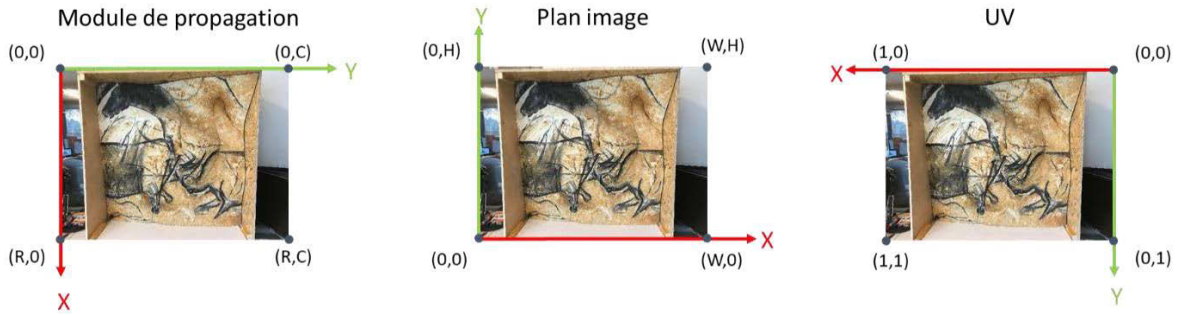


Figure 151 : Les différents repères impliqués dans le processus d'annotation d'une image. Illustration : auteur.

Les autres méthodes d'interaction 3D basées sur la sélection de points 3D présentées en Chapitre 6 (annotation 3D, prévisualisation d'annotations 2D, baguette magique, mesure d'angles, de distances, de valeurs colorimétriques, etc.) ont été implémentées à partir de la fonction de *point-picking*[49] de

[49] Le fonctionnement de cet outil est détaillé par Schuetz (2016). L'idée est de réaliser un premier lancer de rayon pour détecter les nœuds de l'octree intersectés. Ces nœuds sont rendus dans le *buffer*, mais à la place des couleurs ce sont les indices des points qui sont utilisés. Une petite fenêtre carrée de 17 pixels de côté est extraite pour récupérer l'index du point le plus proche de la souris (c'est-à-dire le point le plus proche du centre et dont l'index est non nul). Cet index est ensuite utilisé pour récupérer

Potree.JS. Cette fonction vise à permettre la sélection d'un point 3D à partir d'un clic de souris[50] et fournit en sortie la liste des attributs des points impactés : indices, coordonnées, et attributs calculés après l'indexation.

L'implémentation d'outils de mesure a été plus complexe car l'extraction de distances, surfaces et volumes nécessite de connaître le rapport d'échelle du nuage de points. C'est aussi le cas de l'approche présentée pour le recalage (section 5.3), qui suppose que les données affichées dans le *viewer* soient déjà à l'échelle.

En effet, l'une des différences fondamentales entre la lasergrammétrie et la photogrammétrie est qu'en photogrammétrie, les résultats sont exprimés à un facteur d'échelle près. Autrement dit, le positionnement relatif des éléments entre eux est cohérent, mais dans un système de coordonnées arbitraire. Dans le cas du relevé, cela constitue une limite notable étant donnée l'importance de la mesure dans les opérations d'analyse et d'interprétation. Face à ce constat, Aïoli propose bien un outil de mise à l'échelle automatique, utilisant un marqueur de type QRCode que l'utilisateur doit placer dans la scène. Lors de la création du projet, ce marqueur est détecté sur les photographies maîtresses et le projet est automatiquement mis à l'échelle. Cette technique permet également de définir le plan horizontal de référence. Cependant, dans le cas des grottes ornées cette opération est délicate, car nous devons minimiser les modifications de l'environnement et si l'insertion de cible n'est pas considérée comme viable pour le suivi AR, elle ne l'est pas non plus pour la mise à l'échelle (voir section 3.3.7).

Pour cette raison, nous avons implémenté un outil simple de mise à l'échelle qui s'exécute lors de la première ouverture du projet. Il s'agit pour l'utilisateur de sélectionner deux points du nuage, et de spécifier manuellement la distance qui les sépare. Pour cela, nous nous basons encore sur le *point-picking* : lorsque l'utilisateur clique sur le nuage de points, nous récupérons les attributs de position du point sélectionné. Lorsque deux points ont été sélectionnés, nous calculons la distance qui les sépare dans le système de coordonnées du projet. L'utilisateur spécifie manuellement la distance réelle attendue, en mètres, via un formulaire HTML, et nous pouvons alors calculer le ratio de mise à l'échelle. Ce dernier est sauvé dans la base de données, et appliqué à la scène 3D parente à chaque ouverture du projet. Si nécessaire, l'utilisateur peut réitérer cette opération, notamment en cas d'erreur ou d'imprécision.

Le ratio ainsi obtenu permet la mise à l'échelle de l'ensemble des ressources du projet, en appliquant une homothétie directement au plus grand groupe de la hiérarchie du *viewer* : la scène 3D. Il nous a également permis de modifier les fonctions de la classe `MeasuringTool` de Potree.JS, afin que les outils de mesure qu'elle contient puissent exprimer des valeurs métriques.

La valeur numérique d'une mesure extraite est affichée à l'écran dans un élément `DIV` de la page HTML. Ainsi, elle ne souffre pas des phénomènes éventuels d'occlusion lorsque la scène comprend différentes ressources. La position de cet élément est calculée à partir de la projection inverse du point 3D par rapport à la caméra virtuelle. Elle est recalculée à chaque mise à jour de la scène, que ce soit en cas de déplacement de l'utilisateur ou en cas d'édition des propriétés des ressources affichées. Le tracé des

l'ensemble des attributs du point sélectionné, en particulier ses coordonnées. L'avantage de cette méthode est qu'elle permet de tenir compte des paramètres de taille d'affichage des points et d'occlusions impliqués dans le rendu choisi par l'utilisateur.
[50] Ou un évènement équivalent, par exemple un toucher sur un écran tactile.

éléments sélectionnés (par exemple un segment et les points à ses extrémités lors d'une mesure de distance) est quant à lui chargé dans un groupe dédié de $scene_{Objects}$.

7.2.5. Export de vues 2D du projet depuis la scène 3D

Nous avons vu précédemment qu'une finalité du relevé était la production d'un document synthétisant les observations réalisées sur l'objet d'étude, afin de faciliter la communication des connaissances dans des contextes souvent pluridisciplinaires. Étant données les perspectives apportées par les outils informatiques, il apparaîtrait légitime de requestionner la pertinence des paradigmes de représentation par projections planes, hérités des pratiques architecturales de la Renaissance et toujours exploités de nos jours pour communiquer des informations complexes. Cependant, dans l'hypothèse où une visualisation 3D interactive ou AR ne serait pas appropriée au contexte d'utilisation, notre prototype nous permet tout à fait de proposer des modalités de communication plus proches des usages actuels.

Nous avons ainsi présenté dans le (Chapitre 6) une méthode permettant de générer automatiquement une légende graphique. Pour être intéressante, cette méthode doit être associée à l'export de vues 2D du projet. Pour cela, nous proposons différentes modalités d'extraction de vues 2D à partir de la scène 3D manipulée (Figure 152). Ces outils sont basés sur la méthode `HTMLCanvasElement.toDataURL()`, qui renvoie un URI de données contenant une représentation de l'image du canvas. L'idée générale est de récupérer la *frame* active du canvas depuis le buffer avant son actualisation, et la sauvegarder dans un format image. Selon les cas, l'utilisateur peut ne désirer sauver qu'une partie des éléments de la scène (annotations spécifiques, mesures, etc.), c'est pourquoi nous proposons également des outils permettant de paramétrer manuellement l'apparence des différents objets du projet (affichage, masquage, modification de l'opacité et selon les objets, de la texture).

Figure 152 : Exports 2D issus d'une scène 3D. À gauche : projection perspective. Au centre : projection orthographique. À droite : extrait d'image panoramique équirectangulaire.

Concernant l'export, le cas le plus simple est celui de la projection perspective, car il s'agit du mode actuellement privilégié pour la navigation. Dans ce cas, l'image à sauvegarder est simplement celle produite par la caméra principale du *viewer*. Dans le cas d'une image orthographique, il suffit de remplacer temporairement le modèle de caméra et laisser l'utilisateur la positionner comme il le souhaite selon un plan de référence choisi librement, avant de procéder à la sauvegarde. Enfin, dans le cas où l'utilisateur souhaite générer une image panoramique de la scène 3D, la caméra virtuelle principale doit être remplacée par une caméra virtuelle de la classe `THREE.CubeCamera`. Il s'agit d'un objet composé de six caméras perspectives disposées pour rendre la scène sous la forme de faces de cube. Pour sauvegarder l'image produite par cet objet, (Sanchez, 2019) propose un script permettant de convertir

les faces de cubes en une image équirectangulaire à l'aide d'un *shader*. Ensuite, l'image cette dernière peut être sauvée selon la même méthode que pour les deux cas précédents.

7.3. Implémentation de la méthode de recalage avec l'API WebXR et ARKit

Nous avons proposé dans la section précédente une implémentation informatique visant à mettre en relation les différentes ressources impliquées dans un relevé d'art pariétal au sein d'un *viewer* unique compris dans une plateforme web. Pour cela, nous nous sommes basés sur l'API WebGL à partir de bibliothèques JavaScript. À ce stade, nous disposons donc d'un *viewer* hybride 2D/3D permettant de charger et manipuler les nuages de points issus de photogrammétrie, les photographies orientées accompagnées d'images auxiliaires, ainsi que les annotations 2D et 3D. Nous l'avons également doté d'outils de mesure et d'annotation 2D et 3D manuels ou semi-automatiques. À partir de cette preuve de concept, nous pouvons maintenant nous intéresser à l'implémentation de la méthode de recalage (section 5.3), afin de réconcilier les environnements réels et virtuels par le biais de modalités d'interactions en réalité augmentée.

7.3.1. Preuve de concept

L'objectif de cette implémentation est de valider l'approche propose en section 5.3 pour le recalage semi-automatique de la scène d'un projet sur la réalité au sein d'un environnement web.

7.3.1.1. Déroulement d'une session XR

Le déroulement de l'affichage d'une page web en XR comprend nécessairement les étapes suivantes :

1. Vérification de la prise en charge du mode XR (compatibilité du navigateur et du dispositif d'affichage) ;
2. Si le support est compatible et disponible, la page annonce la disponibilité du mode XR à l'utilisateur ;
3. Si l'utilisateur active le mode XR, une session est initialisée ;
4. La session est utilisée pour exécuter une boucle de rendu qui produit en temps réel les images à afficher sur le dispositif XR ;
5. La production d'images de rendu se poursuit jusqu'à ce que l'utilisateur quitte le mode XR ;
6. La session est close.

Le processus global présenté précédemment pour le rendu du *viewer* en mode *desktop* comprend deux phases principales : l'initialisation, et la boucle de rendu. Globalement, l'implémentation d'un mode XR n'influence pas la phase d'initialisation car celle-ci concerne essentiellement la préparation des éléments minimaux du *viewer*. En revanche, d'après les étapes énoncées ci-dessus, elle impose l'intégration de deux phases intermédiaires de vérification de compatibilité et d'initialisation XR, ainsi que des modifications dans la boucle de rendu (Figure 153).

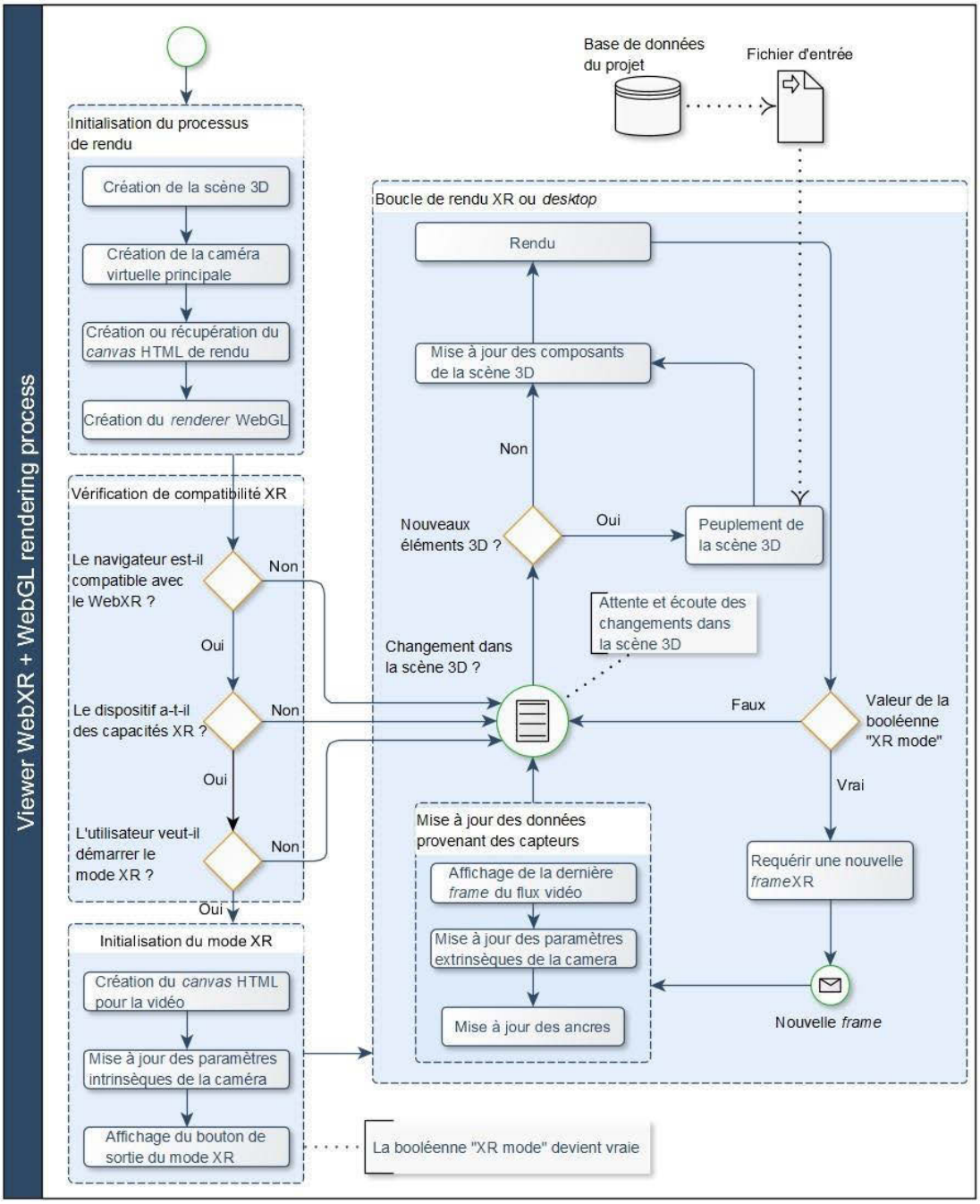


Figure 153 : Diagramme du processus de rendu permettant plusieurs modalités de visualisation (*desktop* ou XR). Illustration : auteur.

➢ Vérification de la compatibilité et initialisation XR

Pour étendre les capacités de notre *viewer* en lui fournissant un mode XR, nous devons dans un premier temps ajouter les étapes évoquées précédemment, destinées à vérifier la compatibilité du navigateur et du dispositif d'affichage. À l'issue de la phase d'initialisation, nous réalisons donc trois vérifications successives :

1. Le navigateur est-il compatible WebXR ?
2. Le mode XR du dispositif d'affichage est-il pris en charge ?
3. Le mode XR est-il demandé par l'utilisateur ?

Si la réponse à l'une de ces trois questions est négative, la boucle de rendu est inchangée, et l'utilisateur bénéficie d'un affichage *desktop* classique. Dans le cas contraire, la session XR est initialisée. Il est alors nécessaire de modifier la structure de la page pour l'adapter à la visualisation en réalité augmentée, en créant le canvas qui recevra les images provenant du flux vidéo du mobile, mais aussi en générant les éléments de GUI spécifiques, en particulier le bouton de sortie de mode XR. Il faut également mettre à jour les paramètres intrinsèques de la caméra virtuelle du viewer afin qu'ils correspondent à ceux de la caméra réelle. Sans cela, les images virtuelles issues du *viewer* et celles issues de la caméra mobile ne seront pas superposables.

- **Boucle de rendu**

Le mode XR étant initialisé, nous pouvons entrer dans la boucle de rendu. Lorsque le mode XR est actif, la boucle a deux objectifs principaux : **assurer la synchronicité** des mouvements de l'utilisateur dans son environnement réel et ceux de la caméra virtuelle de la scène afin que les contenus restent systématiquement superposables, et **actualiser les contenus 3D** du projet à afficher dans la scène en fonction des actions de l'utilisateur, comme dans la boucle de rendu *desktop* classique.

À chaque *frame*, les données du suivi fournies par le SDK XR utilisé sont fournies au client par le navigateur sous forme de chaînes de caractères encodées en base 64. Ce fonctionnement présente l'avantage de permettre le transfert de données complexes comme des images à travers des médias ne supportant que des données textuelles. Outre la nécessité de décoder à chaque *frame* les données qui nous intéressent, un inconvénient de cette approche est que les ressources encodées par ce biais pèsent 10% à 30% plus lourd que les données originales, selon la compression utilisée (Boudrias, 2012). Ainsi, pour des raisons d'optimisation les images du flux vidéo ne peuvent être communiquées en pleine résolution sans impacter sérieusement la fluidité de l'application.

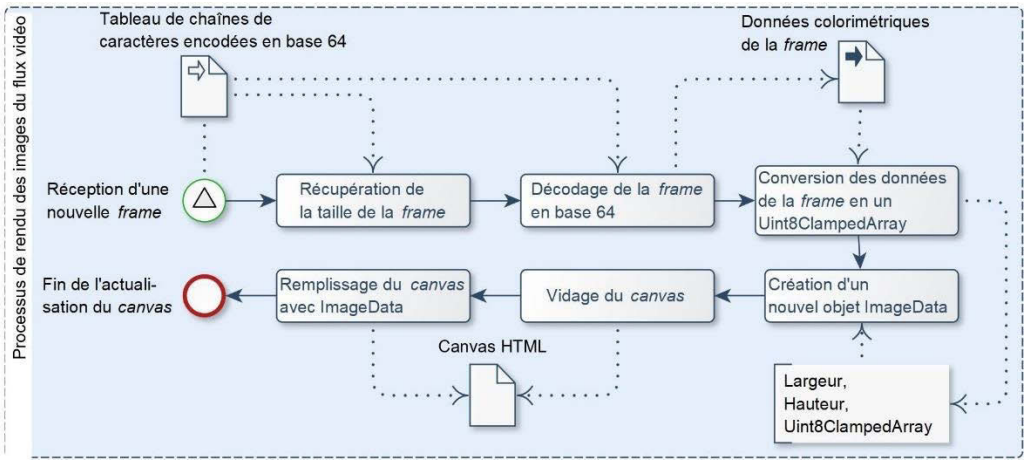


Figure 154 : Diagramme du processus de rendu du flux vidéo dans le canvas. Illustration : auteur.

Nous devons dans un premier temps mettre à jour l'image du flux vidéo dans la page HTML (Figure 154). Pour cela, les données de la *frame* sont décodées et converties en un tableau typé `Uint8ClampedArray`, c'est-à-dire un tableau d'entiers non signés de 8 bits, dont les valeurs sont comprises entre 0 et 255. Ce tableau est utilisé pour générer une image en attribuant pour chaque pixel des valeurs RGB. Enfin, le contenu du canvas est supprimé (*frame* précédente), afin d'y placer la nouvelle image. Dans un second temps, nous devons mettre à jour les paramètres extrinsèques de la caméra virtuelle en fonction des poses calculées lors du suivi, afin que ses déplacements soient

synchrones avec ceux du dispositif d'affichage. Pour cela, nous appliquons directement la matrice calculée par ARKit à la caméra virtuelle.

Enfin, nous devons mettre à jour les ancres de l'environnement réel. Il s'agit d'actualiser les positions des ancres existantes, constamment affinées par la procédure de suivi, mais aussi d'ajouter les nouvelles ancres, créées par la prise en compte d'images-clés. Lorsque cette étape est achevée, nous poursuivons la boucle avec les processus présents dans le rendu *desktop*, c'est-à-dire la prise en compte des actions de l'utilisateur et des modifications opérées dans la scène.

Cette boucle est répétée à chaque *frame*, jusqu'à ce que l'utilisateur manifeste l'envie de quitter le mode XR. Dans ce cas, après confirmation de l'utilisateur, la session XR est close, les éléments HTML dédiés sont supprimés de la page (canvas de rendu vidéo et éventuellement éléments du GUI spécifiques à l'usage XR), et la boucle de rendu repasse en mode *desktop* (Figure 155).

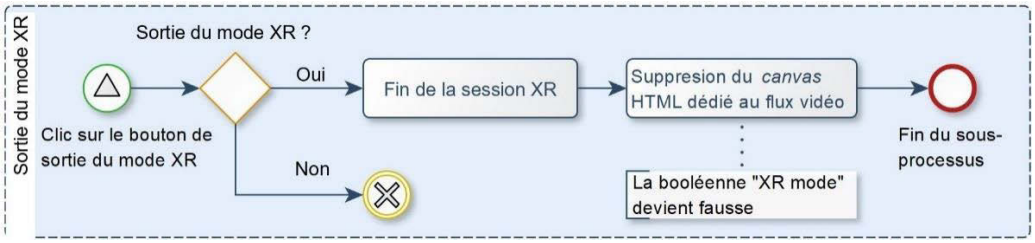


Figure 155 : Diagramme de processus de sortie du mode XR. Illustration : auteur.

Ce processus, résumé sur le schéma (Figure 153) permet d'obtenir une visualisation en réalité augmentée stable. La scène virtuelle est placée dans l'environnement de l'utilisateur et reste en place, comme si elle faisait partie du décor. L'utilisateur peut évoluer avec sa tablette dans le monde réel et visualiser son environnement augmenté. **Cependant à ce stade, rien n'est fait pour que la scène soit positionnée de manière cohérente par rapport à l'objet physique étudié** (Figure 156).

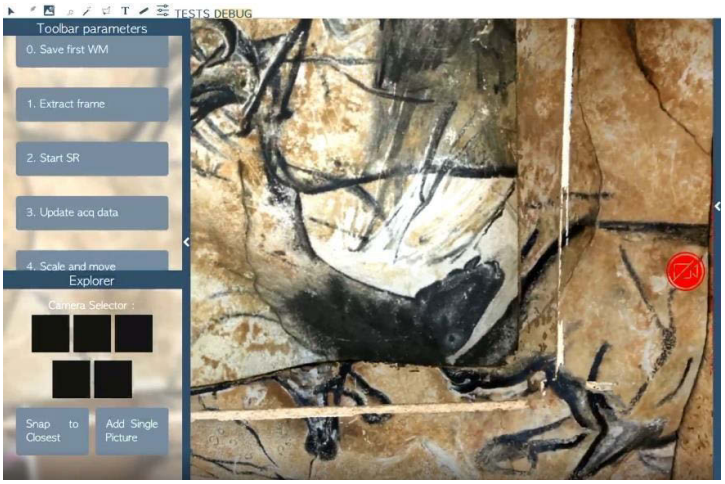


Figure 156 : Capture d'écran de l'application en mode XR. À ce stade, la scène apparaît de manière stable d'en l'environnement réel, mais n'est pas positionnée de manière cohérente.

7.3.1.2. Spatialisation de la *frame* référence

Maintenant que nous avons assuré le rendu de la scène 3D en mode XR, il reste à implémenter la méthode de recalage afin de calculer la position réelle des objets affichés à l'écran. Nous avons proposé une méthode de recalage basée sur la spatialisation d'une *frame* de référence. Nous avions alors identifié deux méthodes possibles (section 5.3.3).

Généralement, la méthode par ajustement de faisceaux est considérée comme plus précise, du fait de l'utilisation des paramètres intrinsèques et de l'optimisation itérative. Cependant, elle est aussi plus lente ce qui peut s'avérer très problématique dans notre cas, en partie pour l'utilisateur qui doit attendre la fin du processus de recalage pour visualiser son projet en AR, en partie parce que le SLAM peut provoquer une dérive progressive, ce qui nous conduit plutôt à limiter le temps de calcul pour favoriser le temps d'utilisation effective. Par ailleurs, nous avons vu qu'en l'état actuel le flux de la caméra est communiqué *frame* par *frame* au client sous forme d'une chaîne de caractères, encodée en base64 et rendu en arrière-plan. Les images ne sont pas fournies avec des métadonnées EXIF, bien que nous puissions tout à fait en ajouter à la volée, connaissant la calibration interne de la caméra utilisée.

La méthode par résection spatiale est quant à elle plus polyvalente puisqu'elle ne nécessite pas de métadonnées, mais aussi plus rapide. Elle est globalement moins précise : si cette méthode de calcul fournit une orientation globalement cohérente, le manque de métadonnées concernant la distance focale peut parfois causer une incohérence concernant la profondeur par rapport à la position réelle.

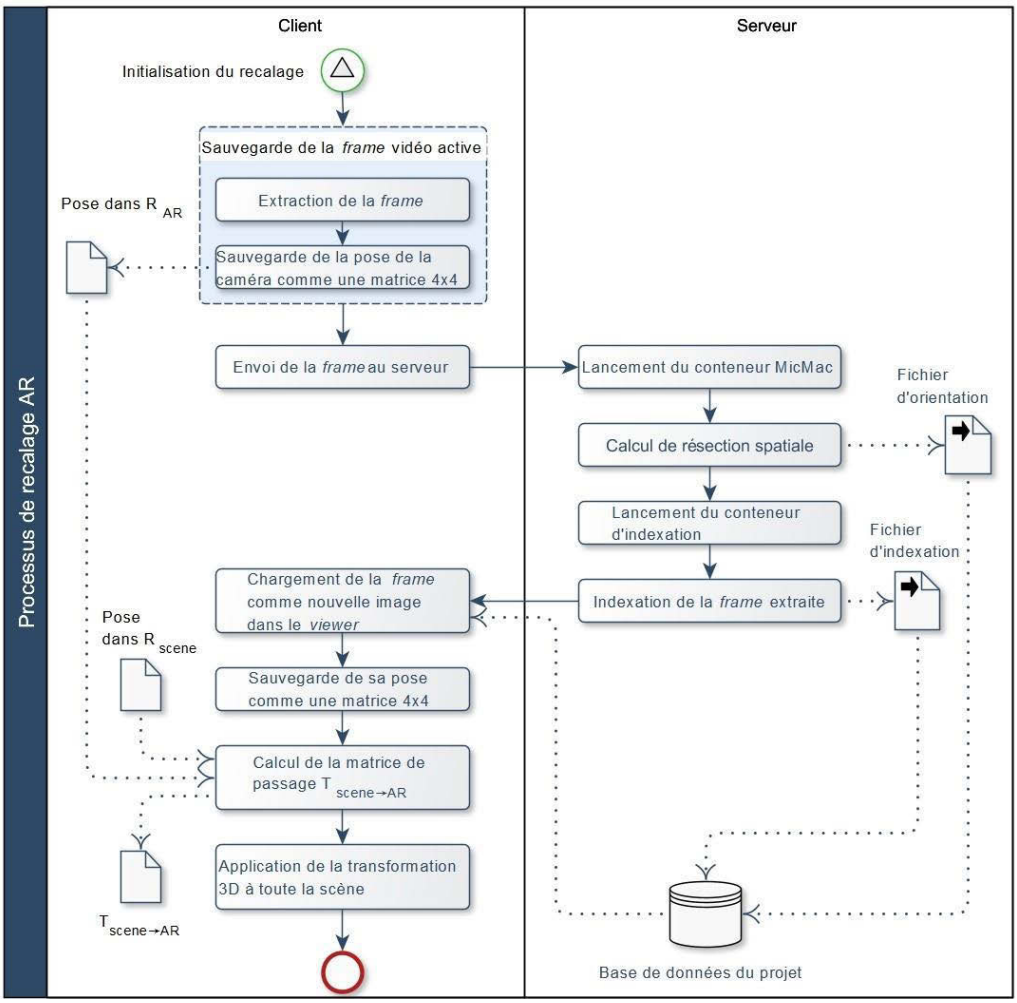


Figure 157 : Diagramme du processus de recalage basé sur un calcul de résection spatiale. Illustration : auteur.

Après évaluation des atouts et limites des deux méthodes, nous avons opté pour la méthode basée sur la résection spatiale en raison de sa flexibilité. Une fois que la *frame* est orientée par rapport à l'acquisition maîtresse, elle est indexée dans un conteneur dédié pour pouvoir bénéficier des annotations préexistantes et devenir à son tour support d'annotation. Enfin, les données de la nouvelle image sont stockées dans la base de données.

Lorsque l'application reçoit le signal de la fin du processus de calcul, nous pouvons charger la *frame* dans la scène en tant que nouvelle image, et calculer $T_{scene \rightarrow AR,}$ la matrice de passage entre la pose exprimée dans R_{scene} et la pose dans R_{AR} sauvée lors de l'extraction. Enfin, cette matrice de passage est appliquée à la scène 3D, qui contient toutes les données du projet, et le recalage est terminé (Figure 157). En cas de besoin, des éléments de GUI permettent à l'utilisateur d'ajuster manuellement la position et l'orientation de la scène selon les trois axes.

7.3.2. Résultats préliminaires

Les tests ont été menés sur des projets réalisés à partir de six prises de vues convergentes du facsimilé du panneau des chevaux de la grotte Chauvet-Pont d'Arc. Des projets ont été créés à partir de ces photographies, et des annotations ont été réalisées en amont dans la version *desktop* de la plateforme. Une fois le projet créé et annoté, nous passons à la version mobile pour activer le mode XR. Tout le contenu de la scène 3D du projet est rendu sur le flux vidéo de la caméra, et les contrôles utilisateur sont initialisés pour synchroniser les mouvements de la caméra virtuelle avec ceux de l'appareil. À ce stade, la position de la scène n'est pas cohérente avec la réalité. Lorsque nous démarrons la procédure de recalage, une *frame* est extraite, et envoyée au serveur de traitement pour les calculs nécessaires à l'alignement.

Entre l'envoi de l'image au serveur et l'alignement subséquent, le temps de calcul varie entre 2 à 5 minutes environ selon le contexte. Plusieurs facteurs peuvent influencer cette durée, principalement la définition de l'image candidate, sa proximité avec celles de l'acquisition maîtresse, et les variations environnementales notamment lumineuses. Évidemment, ce temps de calcul est trop conséquent pour envisager une utilisation confortable, mais néanmoins raisonnable pour une preuve de concept. Par ailleurs, nous verrons par la suite que des solutions existent pour le réduire considérablement. La précision du calcul de résection spatiale dépend beaucoup de l'image de référence, qui doit, dans la mesure du possible, partager de nombreux points homologues avec les images de l'acquisition maîtresse. Ainsi, la qualité de la procédure d'alignement n'est pas constante d'un recalage à l'autre.

Figure 158 : Résultat. Capture d'écran d'un projet visualisé en réalité augmentée après un recalage réussi.

Visuellement, les résultats du recalage semblent satisfaisants, les nuages de points du projet et des annotations sont positionnés de manière cohérente par rapport à la réalité (Figure 158). Nous pouvons observer par transparence la proximité entre les figures pariétales du facsimilé et celles du nuage de points (Annexe 15). Nous pouvons constater un léger décalage, qui peut être lié à la précision du calcul de résection spatiale ou à celle de la mise à l'échelle du projet.

Une dérive peut toutefois se produire lors de l'utilisation de l'application, notamment en cas de mouvements brusques de l'utilisateur ou de mauvaises conditions d'éclairage. Généralement, cela se caractérise par un déplacement progressif de la scène de sa position initiale, à un degré variable. Par exemple, lors d'essais en conditions de très faible luminosité, le phénomène s'est produit à plusieurs reprises, provoquant une dérive de 5 à 10 centimètres après plusieurs minutes d'utilisation. Ce phénomène n'est pas systématique, et peut être imputé à l'algorithme de suivi utilisé.

En effet, la stabilité de la scène repose exclusivement sur la qualité du suivi (estimation des poses des caméras et des positions des ancres). Or, le WebXR n'en est qu'à ses balbutiements et les SDK sur lesquels l'API repose sont encore voués à évoluer et s'améliorer dans le futur. Ainsi, si la dérive peut causer à court terme un certain inconfort dans l'utilisation de l'application (par la nécessité d'ajuster le calage manuellement lorsque cela se produit), elle ne remet pas fondamentalement en question notre méthode de recalage dans la mesure où les choix applicatifs que nous avons réalisés rendent la problématique du suivi indépendante.

7.3.3. Évaluation des résultats

7.3.3.1. Définition d'un protocole d'évaluation

L'appréciation visuelle de l'alignement n'est pas suffisante pour évaluer la précision de notre méthode de recalage ni sa robustesse aux diverses variations environnementales qui pourraient affecter les sites d'art pariétal. Nous devons donc définir un protocole permettant l'évaluation de notre approche.

Idéalement, l'évaluation devrait être basée sur la comparaison des nuages de points issus d'une part de la cartographie 3D établie lors du suivi, et issu d'autre part du calcul photogrammétrique. La distance moyenne entre les deux nuages et l'écart-type fourniraient alors une bonne indication de la qualité de l'alignement : plus ces valeurs seraient élevées, plus le nuage de points du projet serait éloigné de sa position théorique. De plus, l'utilisation de l'algorithme ICP (Besl, McKay, 1992) pour obtenir la meilleure correspondance entre les deux nuages de points pourrait aussi nous renseigner sur la proximité de notre résultat à un cas optimal. Enfin, différentes approches décrites par Bogoslavskyi et Stachniss (2017) pourraient être utilisées pour analyser la qualité des appariements des nuages de points 3D.

Néanmoins, plusieurs difficultés s'opposent à cette perspective. La première est la nature des points de caractéristiques extraits par ARKit que nous avons utilisé à travers le navigateur expérimental WebXR Viewer. Ces points représentent des caractéristiques notables détectées dans l'image de la caméra pendant la session. Comme l'indique Apple dans la documentation, leurs positions 3D sont extrapolées et ne correspondent qu'approximativement aux contours des objets, sans aucune garantie que leur nombre et leur disposition restent stables entre deux *frames* de la même session (Apple, 2019). De fait, les points caractéristiques ne constituent pas une représentation intelligible de l'environnement, et ne constituent pas une donnée suffisamment stable pour servir de référence lors de la comparaison des nuages de points : un tel protocole d'évaluation ne serait pas reproductible. En outre, Voloshin (2018) indique que très peu de points 3D sont produits au cours d'une session. Ainsi, dans une scène très détaillée, il arrive que moins de 500 points soient extraits, en incluant des doublons.

Deuxièmement, en raison de l'aspect expérimental de l'API, l'utilisation de WebXR réduit notre capacité à accéder aux données fournies par le SDK ARKit. En particulier, le navigateur utilisé pour notre expérience ne permet pas de collecter les points caractéristiques 3D côté client, pour des raisons de latence. En outre, le système de coordonnées peut changer avec le temps au cours du suivi, ce qui signifie que dans l'hypothèse où nous pourrions collecter ces points il n'y aurait aucune garantie que les coordonnées (*x, y, z*) obtenues soient comparables d'une *frame* à l'autre.

Dans notre cas, la qualité de du recalage dépend essentiellement de deux facteurs : la précision du calcul de la résection spatiale d'une part, et la précision des poses de la caméra obtenue par SLAM d'autre part. Une autre approche consisterait donc à évaluer ces deux aspects indépendamment, afin de déduire un indicateur de qualité pour la méthode globale. Toutefois, cette approche ne fournirait que des informations indirectes et incomplètes : elle ne permettrait d'évaluer que la qualité de nos données d'entrées, et non de statuer sur la précision du recalage réalisé à l'issue des calculs.

Étant donné que la définition d'un protocole d'évaluation dans notre contexte n'est pas une question triviale, nous proposons de simplifier le problème en le réduisant à un protocole d'évaluation 2D. Notre approche est d'obtenir un indicateur de qualité basé sur la comparaison de couples d'images. Chaque couple est constitué d'une image de référence et d'une image AR, de poses identiques. L'image de référence est une image brute extraite du flux vidéo, qui ne représente donc que l'environnement réel. L'image AR représente la même image, augmentée par le nuage de points recalé du projet (Figure 159).

Ces deux images sont comparées à l'aide d'un algorithme de comparaison d'images, incluant l'extraction et la description des points d'intérêt par SURF (Bay et al., 2008) et leur appariement avec FLANN (Muja, Lowe, 2009). La distance entre chaque paire de points homologues peut alors être calculée, ainsi que la distance moyenne (Figure 160).

Figure 159 : Exemple de couple d'images extrait pour l'évaluation. À gauche : image de référence. À droite : image AR.

Dans le cas idéal où la scène 3D est parfaitement recalée, le résultat attendu est une distance moyenne proche de zéro, car les deux images sont alors quasiment identiques. De même, plus l'erreur de recalage de la scène 3D est grande, plus la distance moyenne attendue est élevée.

Évidemment, dans la pratique, l'alignement n'est jamais parfait. Pour cette raison nous devons d'abord obtenir une note de référence à partir de l'évaluation d'une scène 3D recalée manuellement sur sa

position réelle. Étant donné que cette étape repose essentiellement sur une appréciation visuelle, les données de référence ne seront pas idéales, et les distances moyennes, quoique faibles, ne seront pas nulles. L'évaluation du recalage semi-automatique est ensuite réalisée à partir de ces données de référence. Par ailleurs, pour évaluer le recalage de manière globale, nous devons appliquer cette méthode à plusieurs couples d'images représentant des points de vue variés de l'environnement augmenté, afin de disposer d'un panel représentatif de données.

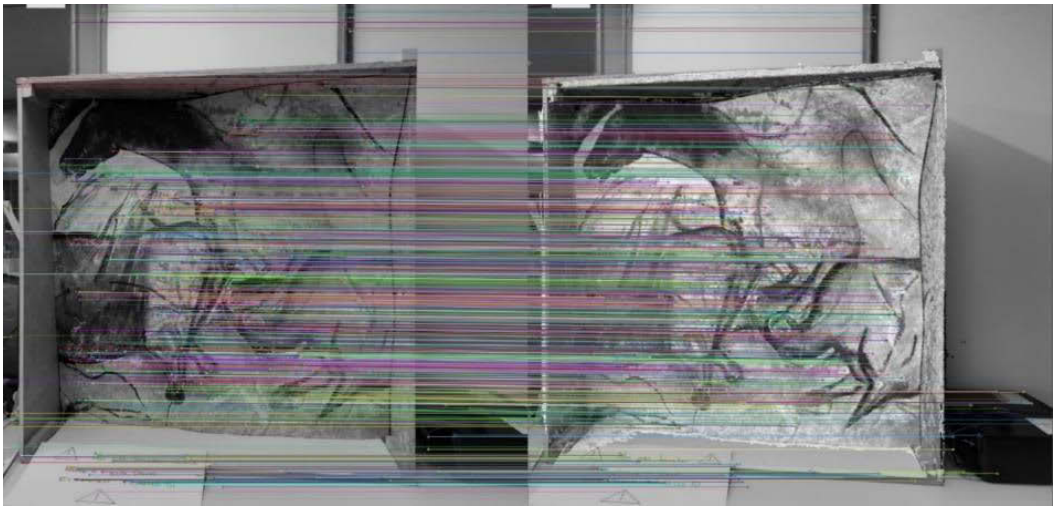


Figure 160 : Comparaison des deux images par SURF et FLANN. La distance entre chaque paire de points homologues donne une indication concernant l'écart du nuage de points à sa position réelle.

Ce protocole d'évaluation, liée au contexte particulier de l'utilisation d'une API émergente non standard, n'est pas suffisamment générique pour nous permettre de nous situer par rapport aux méthodes existantes de recalage existant dans la littérature ni d'évaluer notre approche sur des jeux de données de référence comme KITTI, TUM ou EuroC (Geiger et al., 2012 ; Sturm et al., 2012 ; Burri et al., 2016). Cependant, il nous permet tout de même de comparer de manière relative nos recalages successifs, par exemple pour estimer l'influence de certains éléments contextuels tels que la luminosité ou la complexité de la scène sur la précision de l'alignement.

7.3.3.2. Interprétation et limites de la méthode d'évaluation

Nous avons mis à l'épreuve ce protocole pour évaluer le recalage de l'un des projets réalisés avec les données du facsimilé. Nous avons extrait des couples d'images de points de vue différents conformément à l'approche détaillée précédemment (Figure 161), ainsi qu'un couple de référence obtenu après recalage manuel du projet.

Considérons dans un premier temps le cas des vues 0 à 5. Sur les images extraites, nous pouvons constater visuellement un décalage plus ou moins marqué selon les points de vue. Celui-ci est très visible sur les vues 3 et 4 prises après une discontinuité dans le suivi AR, et plutôt léger sur les autres. Les résultats mesurés à l'aide de notre protocole d'évaluation confirment cette première appréciation, puisque la distance moyenne entre les points homologues est nettement supérieure pour les vues 3 et 4 que sur les vues 0, 1, 2 et 5 (Tableau 1).

Figure 161 : Vues AR 0 à 5 extraites pour l'évaluation d'un recalage.

	Vue 0	Vue 1	Vue 2	Vue 3	Vue 4	Vue 5	Vue 6	RÉFÉRENCES
Nombre de points	989	969	1435	1152	1118	1102	686	**1032**
Distance moyenne	4,8994	5,0584	5,6514	8,9046	9,2761	5,7950	40,2568	**3,1490**
Distance minimale	0,0004	0,0009	0,0007	0,0006	0,0003	0,0002	2,0396	**0,0001**
Distance maximale	22,1833	16,8489	22,4616	214,0980	30,2436	19,4689	101,8540	**19,5773**
Valeur médiane	3,2999	4,0434	0,0828	0,5301	13,7920	6,9709	37,2951	**1,6276**
Écart-type	5,3911	3,5590	6,2354	10,8788	9,2467	5,8277	21,9548	**3,5803**

Tableau 1 : Données quantitatives obtenues à l'aide de notre protocole d'évaluation.

Par ailleurs, chacune de ces vues laisse apparaître une portion importante de l'environnement de l'objet d'étude, qui n'a pas fait l'objet de numérisation 3D. Ainsi, une partie des points homologues détectés sur les images concerne cet environnement plutôt que l'objet d'étude. Cet environnement étant stable, il est logique que les distances minimales obtenues lors de l'évaluation soient très faibles.

Figure 162 : Points homologues détectés sur les deux images du couple 3. Nous pouvons très nettement identifier un faux positif, caractérisé par le segment oblique vert.

En revanche, les distances maximales obtenues peuvent être attribuées à d'importants décalages entre les images – donc entre les nuages de points recalés et leur position théorique – mais aussi à des erreurs d'appariement liées à notre méthode d'évaluation. Sur la vue 3 par exemple, nous pouvons observer un faux positif (ligne oblique sur la Figure 162), qui explique la distance mesurée de plus de 214 pixels.

Pour mieux tenir compte de ces erreurs indépendantes de la méthode de recalage, il apparaît donc intéressant de regarder conjointement l'écart-type des distances obtenues, qui procure une indication sur l'homogénéité des mesures. Ainsi, un écart-type faible indique que les mesures sont globalement proches de la moyenne, quand un écart-type élevé nous montre que les mesures présentent d'importantes disparités suggérant des erreurs d'appariement.

Figure 163 : Couple d'images n°6.

Intéressons-nous maintenant au cas de la vue n°6 qui souligne bien les limites de cette méthode d'évaluation (Figure 163). Le couple d'images n°6 représente une vue rapprochée de l'objet d'étude. Elle a été extraite rapidement après le couple 5 et aucune discontinuité dans le suivi AR n'est survenue entre les deux extractions, c'est-à-dire que le recalage du nuage de points sur la réalité est globalement semblable sur les deux couples.

La différence majeure entre ces deux couples repose donc sur le point de vue adopté, qui se caractérise ici par l'absence de l'environnement réel. Pourtant, les mesures obtenues pour cette vue sont près de dix fois supérieures, donc très mauvaises si l'on ne tient compte que du tableau. Il apparaît donc clairement que l'influence des points homologues détectés dans l'environnement est beaucoup trop importante dans cette méthode d'évaluation. D'ailleurs, l'une de nos expérimentations, menée sur un objet différent, a fourni des résultats quantitatifs extrêmement positifs alors que le recalage était mauvais, car les points homologues évalués étaient majoritairement extraits en dehors de l'objet d'étude.

Ainsi, lorsque l'on adopte un point de vue approché, nous faisons face à deux phénomènes. D'une part l'absence de points homologues *stables* baisse la moyenne par rapport aux autres vues, d'autre part, le nuage de points que nous recalons est lui-même basé sur la réalité et peut présenter des erreurs de reprojection, qui influencent d'autant plus l'évaluation que le nuage de points est visible sur l'image.

Cette forte influence du cadrage de l'image extraite sur l'évaluation soulève alors plusieurs problèmes. Pour garantir la répétabilité du protocole d'évaluation pour évaluer plusieurs recalages d'un même projet, il convient donc d'extraire pour chaque évaluation des séries de vues présentant exactement les mêmes poses, ce qui pourrait passer par l'automatisation de la trajectoire de test, par exemple à l'aide

d'un dispositif robotisé. D'autre part, pour pouvoir comparer les résultats quantitatifs obtenus pour le recalage de plusieurs projets différents, il faut s'assurer que l'influence de l'environnement y est semblable, ce qui s'avère beaucoup plus délicat. Comme nous l'avons énoncé, un cadrage serré sur l'objet réduit cette influence, mais augmente celle des erreurs liées aux calculs photogrammétriques. Ainsi, il pourrait être intéressant d'intégrer dans l'évaluation du recalage des descripteurs caractérisant la qualité de la numérisation.

Par la suite, il conviendra d'améliorer le protocole d'évaluation à la lumière des limites énoncées, afin de mener de nouveaux tests sur des projets plus variés et selon différentes conditions environnementales. Il s'agira d'évaluer l'influence du contenu des projets sur le recalage, mais aussi celle des conditions d'extraction de la *frame* de référence sur la qualité du suivi et du calcul de résection spatiale.

7.3.4. Limites et perspectives de la méthode de recalage

Comme nous l'avons vu dans les sections précédentes, l'approche que nous proposons pour l'alignement réel/virtuel repose sur le suivi AR ainsi que sur l'extraction d'une image du flux vidéo et son orientation par rapport aux images maîtresses du projet. Ainsi, la qualité du recalage est sujette à plusieurs limites inhérentes au suivi ou à la photogrammétrie.

Concernant les limites liées au suivi, notre approche présente l'avantage d'être indépendante de l'algorithme utilisé. En effet, étant donné que nous utilisons l'API WebXR, c'est le navigateur qui joue le rôle d'intermédiaire entre les SDK de réalité augmentée et le client. Ainsi, à chaque mise à jour des SDK et du navigateur, notre application devrait bénéficier des améliorations logicielles sans nécessiter d'opération de maintenance particulière. Nous pouvons donc supposer que les estimations de poses sur lesquels notre méthode repose seront plus précises dans le futur.

En ce qui concerne les limites liées aux calculs d'orientation, la proximité de l'image extraite avec les images maîtresses s'avère un facteur décisif, tant d'un point de vue spatial qu'environnemental. Ainsi, si l'image extraite est trop éloignée des images maîtresses, ou si elle partage trop peu de caractéristiques communes avec ces dernières (par exemple en cas de variations lumineuses), le calcul risque de fournir des résultats erronés voire de ne pas aboutir. Pour limiter les possibilités d'erreurs, nous pourrions envisager d'extraire une série d'images plutôt qu'une seule et ne sélectionner que celles qui procurent les meilleurs résultats en termes d'appariement, mais cela risque d'allonger le temps de calcul.

Comme nous l'avons déjà énoncé, le temps de calcul constitue une seconde limite de notre approche. Pour limiter ce problème, nous avons proposé de limiter notre méthode de recalage à une phase de précalibration, et de réaliser tous les recalages ultérieurs par **persistance AR** (voir section 5.3.4). L'implémentation de cette méthode n'a pas présenté d'obstacle particulier, et nous avons pu vérifier qu'elle permet en effet de limiter considérablement le temps de calcul – qui passe de quelques minutes à quelques secondes – tout en assurant un alignement cohérent[51] (Annexe 16) En revanche, cette méthode dépend beaucoup plus du suivi que la précédente, car elle se base sur le chargement des ancres sauvegardées. Ainsi, il est probable que certains cadres d'utilisation se prêtent moins à cette approche,

[51] Des vidéos de la preuve de concept présentant les résultats du recalage sont disponibles à l'adresse suivante : *www.aioli.cloud/experiments/*

il conviendra d'évaluer cette possibilité en testant des objets d'études et des environnements plus nombreux et plus variés.

7.4. Bilan

Dans ce chapitre, nous avons présenté une implémentation informatique des approches exposées dans les chapitres précédents. Celle-ci a donné lieu à un prototype de système d'information multimodal dédié au relevé d'art pariétal. Il comporte un *viewer* hybride 2D/3D fonctionnel, associé à des modalités d'interactions *desktop* et AR, et doté d'outils visant à faciliter l'interprétation des ressources numériques (accentuation d'attributs perceptifs et annotation multimodale).

Pour cela, nous avons implémenté une méthode de recalage générique permettant de visualiser les différents projets en réalité augmentée dans leur contexte réel, et ce, indépendamment de leurs contenus. La preuve de concept, réalisée à l'aide d'outils émergents tels que l'API WebXR, les navigateurs expérimentaux compatibles, ou encore les SDK de réalité augmentée sur lesquelles ils reposent, nous a permis d'obtenir des résultats prometteurs. Elle semble ouvrir à de nombreuses perspectives, notamment en matière de collecte de données numériques, de travail collaboratif, de suivi multitemporel, mais aussi – et surtout – de traçabilité.

Cependant, si visuellement le recalage que nous obtenons est satisfaisant, il reste difficile à évaluer en raison de l'aspect expérimental des outils mobilisés. Pour cette raison, nous avons proposé un protocole d'évaluation permettant dans quantifier l'écart entre le positionnement des projets recalés et leurs positions théoriques. Nous avons néanmoins observé que ce protocole présente des limites non négligeables qu'il conviendra de dépasser à l'avenir avant de tester la viabilité du prototype face à différents contextes d'utilisations.

Par ailleurs, au-delà des limites techniques soulevées par les différentes expérimentations présentées précédemment, il demeure des questions importantes à traiter, d'ordres ergonomiques et cognitives. En effet, en cherchant à proposer des méthodes de visualisation et d'interaction adaptées aux enjeux du relevé d'art pariétal, nous considérons le système comme une entité dévolue à un usage spécifique, ce qui fait naturellement intervenir le contexte d'utilisation mais aussi d'éventuels verrous *humains*. Ce contexte introduit donc la nécessité de procéder à une évaluation des outils proposés auprès des futurs utilisateurs qui dépasse les considérations uniquement fonctionnelles, mais cette perspective est pour l'heure délicate à envisager du fait des caractères exploratoires et expérimentaux des outils présentés.

Conclusion

Ce travail de recherche s'est intéressé à l'élaboration d'un système d'information spatialisé multimodal dédié au relevé patrimonial, et plus spécifiquement au relevé d'art pariétal. Les enjeux abordés dans cette thèse ont concerné la mise en relation de données spatialisées au sein d'un *viewer* 2D/3D, la proposition d'une méthode générique pour l'alignement réel/virtuel semi-automatique de scènes basées sur la réalité, et le développement de méthodes d'augmentation de l'environnement destinées à faciliter le processus d'interprétation. Ces recherches ont été menées sur la base de travaux antérieurs du laboratoire MAP concernant l'annotation sémantique de ressources numériques.

Problématique

Dans la première partie de ce travail, nous nous sommes attachés à expliciter la notion de relevé afin d'énoncer les enjeux inhérents à cet exercice (0). Les grottes ornées sont des objets d'étude d'une grande complexité, tant du point de vue morphologique que du point de vue de la diversité des informations qu'elles supportent. Le but du relevé est de rendre intelligible cette complexité, tout en respectant certaines exigences. Celles-ci comprennent notamment la fiabilité géométrique, la possibilité d'un traitement analytique permettant d'isoler les représentations selon des critères morphologiques ou sémantiques, ou encore la possibilité de mettre en relation la morphologie du support et la disposition des figures. En outre, la littérature souligne que l'étude des grottes ornées ne saurait faire l'économie du processus de découverte apporté par le travail *in situ*.

L'étude des différents outils et méthodes d'analyse (Chapitre 2) nous a révélé que le relevé contemporain ne se limite plus à un acte individuel ni à un document unique. Le relevé fait aujourd'hui appel à des disciplines et des techniques complémentaires, et est constitué de ressources extrêmement variées, tant 2D (photographies numériques, multispectrales, macro, RTI, etc.) que 3D (lasergrammétrie, photogrammétrie, maillages, etc.). Ce contexte cause un effet de dispersion des données qui nuit considérablement à la possibilité de croiser les regards au sein d'équipes pluridisciplinaires. En outre, malgré l'apport considérable des technologies d'acquisition 3D modernes en termes de caractérisation géométrique, leur démocratisation a eu pour effet d'écarter progressivement la dimension cognitive de la pratique du relevé. Autrement dit, la sauvegarde massive des attributs géométriques de l'objet d'étude décale le moment où s'opère la sélection d'éléments signifiants. Dans le cadre d'un relevé numérique, la majorité du travail de sélection et d'interprétation intervient *ex situ* auprès d'une copie numérique. Ce changement de paradigme marque une rupture significative entre les activités de relevé sur site et hors site.

Face à ce contexte, nous avons souligné la nécessité de proposer un nouveau paradigme pour le relevé d'art pariétal. Plus précisément, il s'agit d'envisager le relevé comme un cycle non linéaire, un processus dynamique et itératif comprenant des phases *in situ* et *ex situ*, conciliant les apports du relevé numérique en termes de caractérisation géométrique, et ceux du relevé *classique* en termes d'interprétation et d'enrichissement sémantique.

Bilan des contributions

Comme nous nous sommes attachés à le montrer en Chapitre 3, cette problématique articule notre recherche à la croisée de deux domaines : les systèmes d'information patrimoniaux, et la réalité mixte. Les systèmes d'information patrimoniaux constituent d'une part un levier vers la définition d'une approche intégrée susceptible de fédérer à la fois les données et les acteurs. D'autre part, par son lien profond à la question de la mesure de l'existant, la réalité mixte constitue l'aboutissement logique d'un processus qui démarre de la lecture du réel, passe par l'analyse et l'interprétation de sa complexité, et qui retourne au réel pour l'enrichir des informations produites. Il s'agit à ce titre d'un second levier fondamental.

L'état de l'art concernant ces deux domaines nous a permis de préciser en Chapitre 4 notre approche et le cadre expérimental. Nous avons décidé de nous inscrire dans la continuité du travail mené par le laboratoire MAP dans le contexte du développement de la plateforme Aïoli. L'approche proposée consiste à étendre à l'objet réel la relation projective établie entre les différentes ressources 2D et 3D relatives à un objet d'étude. Concrètement, il s'agit de faire converger ce qui relève de la saisie de données spatiales avec ce qui relève de l'interprétation de l'objet dans un environnement unique. Pour cela, nous avons proposé d'utiliser le nuage de points 3D à la fois pour fédérer les ressources impliquées dans le relevé et comme intermédiaire entre l'objet réel et les données numériques. Concernant le cadre expérimental de ce travail, nous avons visé l'élaboration d'une preuve de concept se présentant sous la forme d'une application web. Celle-ci tient Aïoli comme base de départ (architecture logicielle, structuration des données, création de projets, processus de propagation, etc.), étendue de trois fonctionnalités fondamentales où se situent nos contributions :

- Mise en relation de données spatialisées 2D et 3D dans un *web-viewer* unique temps réel (section 5.2) ;

- Recalage semi-automatique des projets sur la réalité pour offrir un mode de visualisation et d'interactions en réalité augmentée (section 5.3) ;

- Augmentation géométrico-visuelle de l'environnement pour faciliter le processus d'analyse et annotation multimodale (Chapitre 6).

Au regard des enjeux énoncés, nous pensons que l'approche proposée répond à un certain nombre d'exigences du domaine du relevé d'art pariétal. Afin de permettre **l'unification des trois moments énoncés par Saint Aubin en un moment unique face à l'objet d'étude**, nous avons proposé une méthode permettant de spatialiser les ressources impliquées dans le relevé au sein d'un espace de représentation unique cohérent. La diversité des représentations réunies dans ce *viewer* permet d'une part de limiter l'effet de dispersion des données, mais surtout de faciliter la compréhension de l'objet et rendre compte de sa complexité, tant géométrique que visuelle. Le système d'annotations sémantiques spatialisées structurées en calques correspondants chacun à un niveau d'observation offre bien la possibilité d'un traitement analytique guidé par des critères sémantiques, ce qui constituait l'une des exigences énoncées par Jean Airvaux (Chapitre 1). À partir de cette première preuve de concept, nous avons proposé une méthode de recalage dans l'optique de concilier les apports du relevé *in situ* et *ex situ*.

Face aux précautions imposées par le caractère sensible des grottes ornées, nous avons proposé une méthode d'alignement réel/virtuel générique et sans marqueurs. En outre, nous nous sommes attachés à privilégier un environnement web en raison du potentiel collaboratif, ce qui nous a conduits à utiliser une ébauche d'API, le WebXR. En dépit des difficultés liées au caractère expérimental des outils exploités, les résultats obtenus sont très encourageants et permettent bien de visualiser et d'interagir avec les contenus du *viewer* en réalité augmentée. Cela constitue une avancée importante, car en garantissant la continuité du travail réalisé sur site et hors site, nous offrons de concilier à la fois les capacités perceptives de l'observateur et les capacités calculatoires de l'informatique. À ce titre, nous considérons que cette pratique, que nous désignons comme du *relevé multimodal* constitue une réponse au paradoxe de Moravec : loin de s'opposer, les capacités humaines sur les tâches haut niveau et informatiques sur les tâches bas niveau, se complètent.

Enfin, en tenant compte des apports des différents types de ressources manipulées lors d'un relevé, nous avons développé des méthodes d'augmentation d'attributs géométriques et visuels destinés à faciliter les processus d'interprétation et d'annotation. Pour cela, nous nous sommes notamment intéressés à la production d'images et de textures de caractérisation morphologiques, à l'élaboration de traitements temps-réel de flux vidéo, ainsi qu'à l'élaboration d'outils d'interaction capables d'en tirer parti. La complexité morphologique des sites d'art pariétal et l'éloignement des parois justifié par les enjeux de conservation causent de nombreuses préoccupations concernant la justesse géométrique des observations. Or, la spatialisation, et plus spécifiquement la maîtrise des relations projectives entre les différentes représentations 2D, 3D et l'environnement réel, constitue justement le fondement de notre approche et de ses modalités d'interaction. Dans cette optique, il semble que l'approche présentée dans ce travail pourrait constituer une avancée notable pour le relevé d'art pariétal. Comme nous le verrons dans la section suivante, elle présente cependant des faiblesses et des limites qui doivent être analysées de manière approfondie.

Limites techniques et méthodologiques

Spatialisation des ressources dans le *viewer*

Que ce soit pour l'affichage ou les interactions, notre méthode repose sur le référencement spatial des ressources autour du nuage de points par photogrammétrie. Notre approche est donc sujette aux limites propres à la photogrammétrie, en particulier en ce qui concerne les photographies dont l'acquisition doit respecter un protocole rigoureux. En cas d'erreurs ou d'imprécisions dans le calcul de photogrammétrie, c'est tout le *viewer* qui est impacté. Au niveau de l'affichage, cela peut se manifester par des incohérences dans la géométrie des caméras, dans l'alignement des images, ou encore par des erreurs au niveau du nuage de points. Par exemple, lors de nos expérimentations, nous avons rencontré le cas d'une caméra aberrante dont le point principal était extérieur au plan image. En ce qui concerne les interactions, les outils d'annotation reposent sur la maîtrise des relations projectives entre les ressources. Si celles-ci sont faussées, alors les annotations le seront également.

Prise en charge de ressources 2D visuellement éloignées du jeu d'images maîtresses

Concernant la prise en charge des ressources impliquées dans un relevé d'art pariétal, nous avons vu que notre approche permet d'afficher puis utiliser comme support d'annotation n'importe quelle ressource 2D pour peu qu'elle puisse être orientée, que ce soit de manière directe ou indirecte. Ce fonctionnement interroge donc sur la robustesse et la flexibilité des processus de calcul photogrammétrique, car cela impose soit d'être capable de détecter un nombre suffisant de points homologues sur l'image à orienter, soit de connaître sa position relative par rapport à une image intermédiaire supportant le calcul. Dans le cas d'images visuellement très éloignées du jeu d'images maîtresses, par exemple des dessins, des images d'états antérieurs avant dégradation, ou même des images de facsimilé du panneau étudié, ce fonctionnement peut s'avérer inadapté. Face au caractère décisif du moteur de corrélation dans notre approche, il serait intéressant de s'intéresser à des stratégies de fusion de données multimodales en identifiant des protocoles d'acquisition et de traitement interopérables, reproductibles, et adaptés aux spécificités des différentes ressources.

Fusion de données 3D issues de lasergrammétrie et photogrammétrie

Nous avons vu en (Chapitre 5) qu'il était possible de prendre en compte des données issues de lasergrammétrie (nuages de points des stations et images panoramiques associées). La fusion des données 3D issues de lasergrammétrie et de photogrammétrie soulève cependant plusieurs limites.

En premier lieu, cela implique d'établir une stratégie pour autoriser le recalage des données. Une première approche serait de passer par la géolocalisation des données issues de photogrammétrie afin que l'ensemble des ressources soient exprimées dans un repère global connu. Cependant, comme nous l'avons vu, les données géolocalisées sont délicates à traiter dans le *viewer* du fait du nombre de décimales, qui causent des imprécisions dans le positionnement des ressources. Cette approche nécessite donc de repenser notre protocole de chargement des ressources qui n'est pas encore adapté à ce cas de figure. Une autre approche serait d'agir spécifiquement au niveau du module de photogrammétrie afin de contraindre les images autour du nuage de points issu de lasergrammétrie. Nous en revenons alors aux limites évoquées précédemment concernant l'élaboration de stratégies de spatialisation adaptées aux protocoles d'acquisitions multimodales. Enfin, d'autres approches consisteraient à s'intéresser au calage *a posteriori* des deux familles de ressources, soit à partir de méthodes automatiques (par exemple ICP) ou manuelles (sélection de points de correspondance). Dans cette situation, les modifications concernent le développement d'outils d'interaction au niveau du *viewer*, et de calcul au niveau du serveur de traitement.

Une seconde limite liée à la prise en charge de données issues de lasergrammétrie concerne l'interopérabilité. En effet, la généralisation de la prise en charge de ces ressources impliquerait une étude plus vaste concernant les données que l'on peut exporter via les différents logiciels propriétaires exploités pour le traitement des données laser. Or, s'il existe des formats de fichier interopérables concernant les nuages de points eux-mêmes, l'absence de standard concernant les métadonnées associées aux différentes stations est quant à elle bien plus problématique. Nous disposons alors de deux possibilités. La première consiste à édicter nous-mêmes les règles concernant le téléversement des fichiers sur la plateforme : nous pouvons par exemple proposer un formulaire permettant à l'utilisateur

d'entrer manuellement les informations de positionnement des stations. Cette solution nous dédouane du problème d'harmonisation, mais peut s'avérer très contraignante pour l'utilisateur dans la mesure où les objets d'étude complexes imposent souvent de grandes quantités de stations. Une seconde approche serait de s'intéresser aux particularités liées aux différents matériels afin de proposer des protocoles de chargement distincts. Cette solution présente des limites évidentes en termes de maintenabilité, car elle exige une veille technologique continue.

Méthode de recalage

La méthode de recalage présentée en (section 5.3) présente plusieurs limites. Comme nous l'avons expliqué, celle-ci repose sur l'extraction d'une image du flux vidéo et son orientation par rapport aux images maîtresses du projet. Ainsi, la qualité de l'alignement réel/virtuel est elle aussi sujette aux limites inhérentes à la photogrammétrie. Si l'image extraite est trop éloignée des images maîtresses, ou plus généralement, si elle partage peu de caractéristiques communes avec ces dernières (par exemple pour des raisons environnementales), le calcul peut fournir des résultats erronés, voire ne pas aboutir. L'extraction de l'image du flux vidéo est donc une étape décisive concernant l'alignement réel/virtuel. Pour limiter les possibilités d'erreurs, nous pourrions envisager d'extraire une série d'images plutôt qu'une seule, et de sélectionner celle qui procure les meilleurs résultats en termes d'appariement. Cependant, cette approche risquerait d'augmenter le temps de calcul déjà conséquent.

Le temps de calcul constitue une seconde limite de la méthode proposée. En réponse à ce problème, nous avons proposé de nous baser sur la notion de persistance pour limiter le temps de calcul à la seule étape de précalibration. Ainsi, toutes les sessions AR ultérieures peuvent bénéficier d'un alignement quasiment instantané. Cependant, il demeurerait intéressant de rechercher une méthode d'optimisation du processus de calcul de résection spatiale afin de réduire le temps de calcul.

Dérive du suivi AR

En ce qui concerne la modalité de visualisation en réalité augmentée, nous avons concentré nos efforts sur la seconde couche des *building blocks*, c'est-à-dire la couche logicielle s'intéressant aux contenus, à la façon de les présenter, et à la façon dont les utilisateurs interagissent avec (Figure 82, section 3.3.8). En plus des limites propres à notre approche, notre méthode présente donc aussi des limites liées aux éléments de la couche inférieure, c'est-à-dire aux éléments fondamentaux du système de réalité augmentée.

Cela concerne d'une part les éléments matériels, pour lesquels les limites sont principalement liées à la capacité de calcul, la résolution d'affichage, et la précision des capteurs. D'autre part, cela concerne les éléments logiciels, en particulier les capacités du contrôleur de rendu et la méthode de suivi. Les limites liées au suivi sont sans doute les plus visibles lors de l'utilisation de notre preuve de concept. En effet, une dérive peut survenir au cours du temps, le suivi des ancres peut manifester une certaine instabilité, et l'application peut adopter un comportement variable selon les environnements dans lequel évolue l'utilisateur (difficultés dans les environnements sombres ou monochromes notamment).

En principe, le choix de l'API WebXR nous rend indépendants de ces limites, dans la mesure où les navigateurs seront amenés à étendre d'eux-mêmes la liste des matériels pris en charge et à tenir compte

des améliorations logicielles à travers les mises à jour des différents SDK sous-jacents. En harmonisant les méthodes liées aux systèmes de réalité augmentée, ils simplifient considérablement le développement d'applications web et leur maintenabilité. Ainsi, nous devrions théoriquement bénéficier des évolutions des méthodes de suivi de manière transparente, sans qu'elles ne remettent en question notre approche. En pratique, cela n'est pas certain, car l'API n'est encore qu'une ébauche et n'est implémentée que de manière très expérimentale dans un nombre restreint de navigateurs. En outre, la compatibilité avec d'autres matériels (par exemple, les HMD amenés à être pris en charge) implique un important travail de réflexion sur l'expérience utilisateur et les techniques d'interaction qui dépasse la simple question de la faisabilité.

La représentation des observations

Les limites présentées précédemment sont principalement d'ordre technique. Néanmoins, l'approche présentée dans ce travail de recherche soulève aussi des interrogations d'ordre méthodologique.

En ce qui concerne la représentation des observations, nous avons souligné les disparités existant entre les annotations 2D et les annotations 3D. Ces différences sont liées à la nature des objets utilisés pour l'affichage : géométries solides dans le premier cas, nuages de points dans le second. Pour harmoniser les représentations, nous pouvons envisager de charger les contours des régions 3D sous forme de *splines* en plus des nuages de points. Le paramétrage des matériaux appliqués à ces courbes et l'utilisation de *shaders* adaptés pour les nuages de points permettraient d'appliquer plus rigoureusement la charte graphique définie par l'utilisateur. Cette approche permet de résoudre le problème de la cohérence de la représentation des différentes ressources. Cependant, une autre question se pose, celle de la dimension graphique du relevé, c'est-à-dire de la production du relevé en tant que document synthétique, qui relève du troisième moment énoncé par Saint-Aubin.

Si l'approche présentée dans ce travail semble plutôt bien se prêter au relevé analytique technique, dont les codes graphiques actuellement exploités sont relativement proches de ce que nous proposons, le cas du relevé analytique plastique est plus complexe. Nous avons vu en Chapitre 1 que dans le cas du relevé analytique plastique, le traitement graphique avait pour vocation d'accentuer certains caractères de l'objet étudié. Il ne s'agit pas nécessairement de produire une représentation au plus proche de la réalité, mais d'exploiter le traitement graphique pour mieux communiquer un propos, ou *faire comprendre*. Dans l'immédiat, nous considérons ce travail de traitement graphique comme une opération extérieure de post-traitement, rendue possible par les modalités d'export proposées. Ce faisant, nous occultons néanmoins la nature spéculative du dessin dont l'utilisation est avant tout un acte de pensée :

« L'utilisation du dessin va au-delà du simple travail de transcription passive d'une réalité donnée car le dessin […] est avant tout un outil d'analyse, de conceptualisation et finalement de projection. » (Estevez, 2001)

Si la recherche d'esthétisme dans le relevé intervient dès le tracé des régions, c'est-à-dire si celle-ci est incluse dans l'acte de dessiner, alors notre approche présente une limite importante. Car ce point de vue implique que le geste du dessin – au-delà du tracé délimitant les observations – est une partie prenante du processus d'analyse, ce qui signifie que les limites pratiques de notre preuve de concept (par exemple

des outils d'annotation et de charte graphique) soulèvent aussi des questions relevant des sciences cognitives et de l'ergonomie.

Enfin, l'approche que nous proposons dans ce travail de recherche consiste à offrir aux acteurs du relevé un accès simplifié et immédiat à l'état des connaissances concernant leur objet d'étude. Nous considérons que ce fonctionnement favorise des scénarios dans lesquels il deviendra possible d'identifier de potentielles corrélations entre les différentes observations, entre les différentes disciplines, voire de révéler des éléments sous-étudiés qui bénéficieraient d'une analyse plus approfondie. Cependant, nous pouvons nous demander si, à l'inverse, cette proposition ne va pas à l'encontre de l'idée du processus de découverte tel qu'évoqué par Paillet (2015) : l'accès aux observations et aux assignations préétablies est-il susceptible d'influencer l'observateur, c'est-à-dire d'orienter son analyse au point de causer des biais dans la lecture de l'objet ?

Perspectives de recherche

Outre les questions soulevées par les limites énoncées précédemment, l'approche présentée dans ce travail ouvre des perspectives de recherche.

Déploiement et évaluation du système

Comme nous l'avons expliqué en (Chapitre 3), la constitution générale d'un système de réalité augmentée comporte quatre niveaux distincts. Dans la partie expérimentale de ce travail, nous nous sommes attachés à réaliser une preuve de concept, pour laquelle nous nous sommes principalement focalisés sur le second niveau. Celui-ci concerne la couche logicielle, s'intéressant principalement aux contenus et outils d'interactions. Nous nous sommes également intéressés au troisième niveau, c'est-à-dire le système en lui-même en tant qu'entité tournée vers un usage spécifique, dans la mesure où c'est bien l'étude du relevé d'art pariétal qui nous a menés à proposer un système de réalité augmentée et a guidé nos choix expérimentaux.

Étant donné qu'un système AR se destine nécessairement à un utilisateur final, il existe un quatrième niveau, qui s'intéresse à l'acceptabilité soulevée par le verrou *humain*. Celui-ci dépend de l'utilisateur, notamment de ses habitudes de travail, de son expérience, de ses compétences techniques, etc. Il s'agit de la couche de plus haut niveau, dont l'étude fait intervenir différentes disciplines, notamment les sciences cognitives et l'ergonomie. Ainsi, la preuve de concept ayant démontré la faisabilité de notre approche et sa validité selon des critères techniques, il devient maintenant important de s'intéresser à l'évaluation de notre proposition du point de vue de l'usage. Cette évaluation de l'expérience utilisateur (UX) concerne deux aspects : d'une part l'*utilisabilité* de l'interface, et d'autre part, l'*efficacité* du système pour l'usage envisagé (dans notre cas, le relevé pariétal) et les processus cognitifs impliqués.

L'objectif de cette phase indispensable est de permettre l'amélioration de l'ergonomie de l'application pour faciliter son utilisation, la rendre plus intuitive. Cela passe par le recueil de retours d'expériences et de données objectives concernant l'intérêt, l'efficacité, l'utilisabilité, et l'acceptabilité du système. Il s'agit donc d'une évaluation à mener sur le terrain auprès des futurs utilisateurs.

Un apport continu en données : vers un suivi multitemporel

Une autre perspective réside dans l'opportunité de collecter en continu des données concernant l'objet d'étude lors des sessions de relevé en réalité augmentée. En effet, lors d'un relevé en réalité augmentée, l'utilisateur est amené à évoluer autour de sa zone d'intérêt. L'extraction en arrière-plan des images clés du suivi pourrait alors permettre d'établir des jeux de données exploitables par photogrammétrie pour créer des sauvegardes sous la forme de nuages de points 3D, témoignant de différents états temporels.

Si nous parvenons à proposer un traitement garantissant que ces nuages de points soient exprimés dans un même repère, nous pouvons aisément envisager de propager les annotations au sein des jeux de données correspondant aux différents états temporels. Comme nous l'avons présenté précédemment, les annotations réalisées sur la plateforme font à la fois l'objet de caractérisations géométrico-visuelles et sémantiques. La propagation de ces annotations à différents états temporels permettrait alors d'apprécier quantitativement et qualitativement l'évolution d'altérations au cours du temps. Ce fonctionnement pourrait s'avérer précieux pour la conservation, et s'inscrit tout à fait dans la logique d'une vision du relevé en tant que démarche cyclique itérative, telle qu'envisagée en (Chapitre 4).

Cette perspective implique de résoudre différentes difficultés. Dans l'immédiat, les données exposées avec le WebXR n'incluent pas d'informations concernant les images clés du suivi, et les préoccupations de latence excluent la sauvegarde de l'ensemble du flux vidéo. Il s'agit donc d'établir une stratégie concernant les images à extraire. La sélection automatique temps réel des images provenant d'un flux vidéo en vue de reconstructions par photogrammétrie n'est pas une question triviale, en témoignent par exemple les travaux de Pavoni et al. (2016) et Saleri (2018).

Une seconde difficulté réside dans la volonté de conserver un repère commun aux différents nuages de points. Une solution est proposée par Peteler et al. (2015) mais ne fonctionne qu'à la condition que les images des différents jeux de données offrent suffisamment de points communs, et donc, que les différents états temporels ne présentent pas de changements significatifs. Or, dans le cas que nous envisageons ici, les images extraites dépendant des zones d'intérêt des utilisateurs puisqu'il s'agit d'une fonction réalisée en arrière-plan. Nous ne pouvons donc pas garantir qu'elles répondent à ces contraintes. De plus, cette perspective implique de proposer une méthode permettant de détecter au préalable à quelle zone de l'objet d'intérêt se réfère le jeu de données extrait. Enfin, la propagation d'annotation entre différents états suppose le choix d'un temps de référence, et nécessite une stratégie efficace pour la mise en correspondance des indexations des différents jeux de données représentant le même objet.

Traçabilité : un pas vers la compréhension de la relation observateur/objet

L'aspect collaboratif de la plateforme et les contextes d'utilisation auxquels elle se destine nécessitent d'accorder un soin particulier aux questions liées à la traçabilité, qu'il s'agisse des données ou des observations. La traçabilité renvoie à des enjeux sous-jacents tels que la reproductibilité des résultats, des aspects juridiques concernant la propriété des contenus, la description des contenus sous forme de métadonnées en vue d'une diffusion, ou simplement la capacité à préserver l'historique d'un élément - par exemple une annotation - pour suivre son évolution (création, édition, assignations, etc.). La convergence des différents moments du relevé dans un environnement unique permet aujourd'hui de mener une réflexion globale sur cette question de traçabilité.

Nous pouvons sauvegarder dans la base de données autant d'informations que nécessaire concernant les processus de calculs réalisés ou les actions des utilisateurs sur les données (auteur d'une annotation, auteurs des éléments de sa description, tracé original, support et outil utilisé, etc.). Il ne s'agit alors plus uniquement de garantir la description des contenus et des actions réalisées, mais bien d'envisager la traçabilité comme levier vers la compréhension de la relation qui lie un observateur à son objet d'étude.

Par exemple, en mémorisant et analysant les supports prioritairement exploités pour les annotations supportant une sémantique donnée, nous pouvons espérer comprendre quels attributs géométriques ou visuels représentent le mieux un concept. Plus encore, pourrait-on établir un lien entre les disciplines impliquées dans un relevé interdisciplinaire et les supports privilégiés pour annoter ? Autrement dit, quels rapprochements ou quelles différences peut-on noter dans le regard porté par des acteurs de différentes disciplines sur un même objet d'étude ?

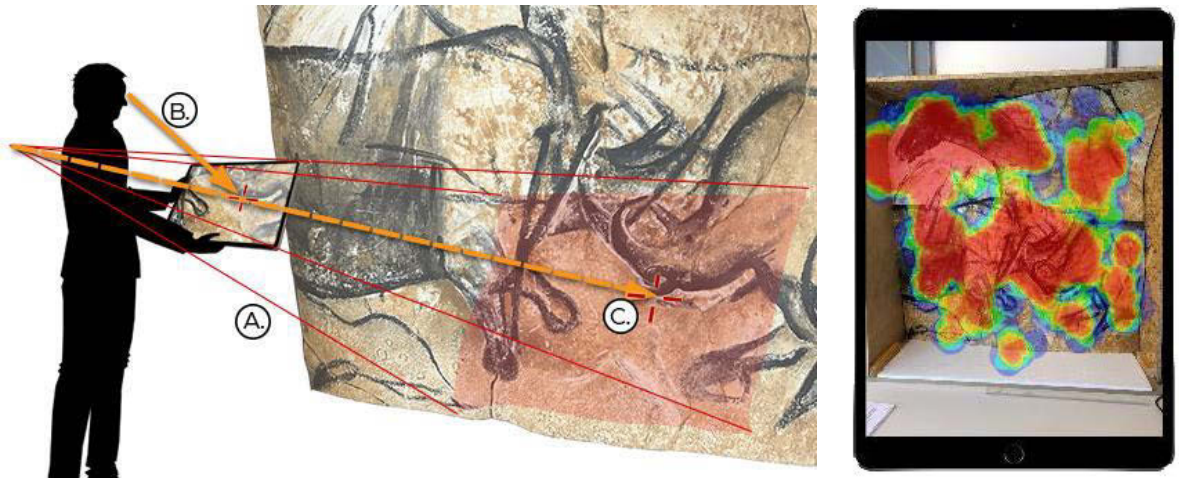


Figure 164 : Les modalités d'interaction en réalité augmentée ouvrent de nombreuses perspectives en termes de traçabilité. A. Les poses du dispositif d'affichage donnent des indications sur les zones d'intérêt de l'observateur. B. L'utilisation de la caméra frontale permet de déterminer le point d'intérêt de l'utilisateur sur l'écran. C. Un lancer de rayon basé sur les informations A. et B. permet de connaître précisément à chaque instant le point 3D observé. À droite : schéma d'une carte de chaleur synthétisant ces informations. Illustration : auteur.

Dans cette optique, la modalité de relevé en réalité augmenté élargit considérablement les perspectives. Les poses successives de la caméra arrière du dispositif nous informent non seulement sur les trajectoires empruntées par l'utilisateur lors de sa visite, mais aussi sur les directions successives de son attention (Figure 164 A). Pour obtenir des informations plus complètes, l'utilisation simultanée de la caméra frontale du mobile nous permettrait de réaliser une fonction d'oculométrie visant à examiner à chaque instant quel point de l'image restituée est observé par l'utilisateur (Figure 164 B). Si nous appliquons en plus une fonction de lancer de rayon à partir de ce point (à partir des mêmes approches que celles détaillées en Chapitre 6 pour l'annotation), nous pouvons connaître la position 3D du point observé (Figure 164 C).

Ces informations pourraient s'avérer précieuses à plusieurs titres. Quel cheminement a mené un utilisateur à son observation ? Autrement dit, comment le contexte de la visite influe-t-il sur le regard porté à l'objet ? D'autre part, la récolte de données concernant les zones d'intérêt des différents acteurs permet d'envisager des scénarios d'analyse s'intéressant à la recherche de zones délaissées ou de recoupements. Par exemple, dans le cas où différents acteurs manifestent un intérêt envers la même zone de l'objet, ces recoupements se traduisent-ils au niveau des annotations sémantiques par une

augmentation de la corrélation entre les régions des différents calques ? Et dans l'hypothèse où un acteur tiers aurait délaissé cette même zone, serait-il intéressant de lui suggérer de l'observer ?

Cette perspective reste à explorer et nécessite d'accumuler et croiser une quantité significative de données quantitatives et qualitatives. Elle fait intervenir le domaine des sciences cognitives (Fontoura et al., 2019).

Potentiels d'utilisation au-delà du domaine du relevé d'art pariétal

Pour ce travail, nous nous sommes intéressés à un cas d'étude spécifique, celui des grottes ornées. Cependant, l'approche présentée est générique, dans la mesure où elle est indépendante de l'environnement de l'objet d'étude et du contenu des projets. Dans cette optique, notre approche pourrait s'appliquer à un domaine plus vaste, celui du patrimoine au sens large, que l'on parle du patrimoine bâti ou matériel. Notre preuve de concept pourrait ainsi être testée sur des objets plus variés, par exemple des éléments architecturaux, des vestiges archéologiques, des peintures, des sculptures, etc. En outre, nous avons beaucoup insisté sur l'interprétation, la connaissance de l'objet d'étude. Cependant, la méthode proposée pourrait trouver des applications dans les domaines de la conservation ou de la gestion d'opérations de restauration.

Enfin, la convergence des acteurs et des données au sein d'un environnement unique pourrait tout à fait intéresser des disciplines extérieures au domaine du patrimoine, par exemple dans des domaines industriels pour lesquels les contraintes opérationnelles sont similaires à celles qui ont guidé nos décisions, notamment pour la modalité en réalité augmentée (transparence du système concernant l'environnement de l'objet étudié, et indépendance vis-à-vis du contenu).

Bibliographie

ABDEL-AZIZ, Y.I. et KARARA, H.M., 2015. **Direct Linear Transformation from Comparator Coordinates into Object Space Coordinates in Close-Range Photogrammetry**. In : *Photogrammetric Engineering & Remote Sensing*. Vol. 81, n° 2, p. 103-107. DOI 10.14358/PERS.81.2.103.

ABERGEL, V., 2019. [en ligne] **Aïoli experiments**. [Consulté le 11 juin 2019] Disponible à l'adresse : http://www.aioli.cloud/?page_id=2239.

ABERGEL, V., PINÇON, G., KONIK, S. et JACQUOT, K., 2019. **Harmonisation et diffusion des ressources numériques 3D des grottes ornées**. In : *In Situ*, n° 39. DOI 10.4000/insitu.21550.

AIRVAUX, J., FOUCHER, P., MAGNANT, D., MARANDET, F. et BERNARD, J., 1991. **Les techniques informatiques du traitement de l'image appliquées à l'étude des gravures paléolithiques**. In : *Paléo*. Vol. 3, n° 1, p. 139-147. DOI 10.3406/pal.1991.1044.

ALLEY, R.,1999. **Algorithm Theoretical basis document for: decorrelation stretch**. Jet Propulsion Laboratory, Version 2.2, Pasadena, CA, USA.

AMMANN, L., BARLA, P., GUENNEBAUD, G., GRANIER, X. et REUTER, P., 2012. **Surface relief analysis for illustrative shading**. In : *Computer Graphics Forum*. Wiley Online Library. p. 1481-1490.

ANDREWS, P., ZAIHRAYEU, I. et PANE, J., 2012. **A classification of semantic annotation systems**. In : *Semantic Web*. n° 3, p. 223–248. DOI 10.3233/SW-2011-0056.

APOLLONIO, F., BASILISSI, V., CALLIERI, M., DELLEPIANE, M., GAIANI, M., PONCHIO, F., RIZZO, F., RUBINO, A., SCOPIGNO, R. et SOBRA', G., 2018. **A 3D-centered information system for the documentation of a complex restoration intervention**. In : *Journal of Cultural Heritage*. Vol. 29, p. 89-99. DOI 10.1016/j.culher.2017.07.010.

APPLE, Inc. ©., 2019. [en ligne] **rawFeaturePoints - ARFrame | Apple Developer Documentation**. [Consulté le 28 août 2019]. Disponible à l'adresse : https://developer.apple.com/documentation/arkit/arframe/2887449-rawfeaturepoints?language=objc.

ARCHAMBEAU, M., 1982. **Un repère orthonormé pour relever les gravures et les peintures pariétales**. In : *Bulletin de la Société préhistorique française*. Vol. 79, n° 7, p. 217-220. DOI 10.3406/bspf.1982.5388.

ARGAN, G. et WITTKOWER, R., 1990. **Perspective et Histoire au Quattrocento**. Les Éditions de la Passion. Paris, France ISBN 2-906229-10-5.

ARNALDI, B., GUITTON, P. et MOREAU, G. (éd.), 2018. **Virtual reality and augmented reality: Myths and Realities**. ISTE. Hoboken, ISTE Ltd/John Wiley and Sons Inc. Computer Engineering Serie. ISBN 978-1-78630-105-5.

AUJOULAT, N., 1987. **Le relevé des œuvres pariétales paléolithiques: enregistrement et traitement des données**. Éditions de la Maison des sciences de l'homme. Paris, France. Documents d'archéologie française, n° 9. ISBN 978-2-7351-0208-2.

AUJOULAT, N., 1993a. **Les projections**. In : *L'art pariétal paléolithique: techniques et méthodes d'étude*. Édition du Comité des travaux historiques et scientifiques. Paris, France. Ministère de l'Enseignement supérieur et de la Recherche, Documents préhistoriques, 5. p. 339-342. ISBN 978-2-7355-0286-8.

AUJOULAT, N., 1993b. **L'évolution des techniques**. In : *L'art pariétal paléolithique: techniques et méthodes d'étude*. Paris, France. Ministère de l'Enseignement supérieur et de la Recherche, Comité des travaux historiques et scientifiques. Documents préhistoriques, 5. p. 317-327. ISBN 978-2-7355-0286-8.

AUJOULAT, N., 1993c. **L'outil photographique**. In : *L'art pariétal paléolithique: techniques et méthodes d'étude*. Paris, France. Ministère de l'Enseignement supérieur et de la Recherche, Comité des travaux historiques et scientifiques. Documents préhistoriques, 5. p. 347-353. ISBN 978-2-7355-0286-8.

AUKSTAKALNIS, S., 2016. **Practical augmented reality: a guide to the technologies, applications and human factors for ar and vr**. Addison Wesley, Boston, Massachusetts, USA, Pearson Education, Inc. ISBN 978-0-13-409423-6.

AURAT, D., LEROY, L., HUGUES, O. et FUCHS, P., 2013. **Un flou adaptatif en fonction du point de focalisation pour réduire la fatigue visuelle en vision stéréoscopique**. In : *8èmes journées de l'Association Française de Réalité Virtuelle, Augmentée, Mixte et d'Intéraction 3D (AFRV 2013)*. Laval, France.

AZÉMA, M., GÉLY, B., BOURRILLON, R. et LHOMME, D., 2012. **La grotte ornée paléolithique de Baume Latrone (France, Gard): la 3D remonte le temps**. In : *L'art pléistocène dans le monde - Actes du Congrès IFRAO.* Tarascon-sur-Ariège : Société préhistorique Ariège-Pyrénées.

AZUMA, R. T., 1997. **A Survey of Augmented Reality**. In : *Presence: Teleoperators and Virtual Environments*. Vol. 6, n° 4, p. 355-385. DOI 10.1162/pres.1997.6.4.355.

BABELON, J-P., 1989. **Châteaux de France au siècle de la Renaissance**. Paris : Flammarion/Picard, 840 p. ISBN 978-2-08-012062-5.

BAHN, P. et VERTUT, J., 1997. **Journey through the ice age**. University of California Press. Berkeley, CA, USA. ISBN 978-0-520-21306-7.

BARBER, D., MILLS, J. et BRYAN, P., 2001. **Laser scanning and photogrammetry: 21st century metrology**. In : *Proceedings of the 18th CIPA Symposium*. Postdam, Germany.

BARBIER, J., KENNY, P., YOUNG, J., NORMAND, J-M., KEANE, M. T., O'SULLIVAN, M. et VENTRESQUE, A., 2017. **MAAP Annotate: When archaeology meets augmented reality for annotation of megalithic art**. In : *Proceedings of the 23rd International Conference on Virtual System & Multimedia (VSMM)*. Dublin : IEEE, p. 1-8.

BAY, H., ESS, A., TUYTELAARS, T. et VAN GOOL, L., 2008. **Speeded-Up Robust Features (SURF)**. In : *Computer Vision and Image Understanding*. Vol. 110, n° 3, p. 346-359. DOI 10.1016/j.cviu.2007.09.014.

BECK, L., LEBON, M., LAHLIL, S., GRÉGOIRE, S., ODIN, G., P., ROUSSELIÈRE, H., CASTAING, J., DURAN, A., VIGNAUD, C., REICHE, I., LAMBERT, E., SALOMON, H., GENTY, D., CHIOTTI, L., NESPOULET, R., PLASSARD, F. et MENU, M., 2014. **Analyse non destructive des pigments préhistoriques: de la grotte au laboratoire**. In : *Paléo. 2014. Vol. Actes du colloque "Micro-analyses et datations de l'art préhistorique dans son contexte archéologique - MADAPCA-Paris 2011"*, numéro spécial, p. 63-74.

BEKELE, M., PIERDICCA, R., FRONTONI, E., MALINVERNI, E. et GAIN, J., 2018. **A Survey of Augmented, Virtual, and Mixed Reality for Cultural Heritage**. In : *Journal on Computing and Cultural Heritage*. Vol. 11, n° 2, p. 1-36. DOI 10.1145/3145534.

BERGSTROM, L., 2018. [en ligne] **New API to Bring Augmented Reality to the Web – Mozilla Hacks - the Web developer blog**. In : *Mozilla Hacks – the Web developer blog*. [Consulté le 10 décembre 2018]. Disponible à l'adresse : https://hacks.mozilla.org/2018/09/webxr.

BERGSTROM, L., 2019. [en ligne] **Bringing Firefox Reality to HoloLens 2**. In : *Mozilla Mixed Reality Blog*. [Consulté le 31 octobre 2019]. Disponible à l'adresse : https://blog.mozvr.com/bringing-firefox-reality-to-hololens-2/.

BESL, P. et MCKAY, N., 1992. **A method for registration of 3-D shapes**. In : *IEEE Transactions on Pattern Analysis and Machine Intelligence*. Vol. 14, n° 2, p. 239-256. DOI 10.1109/34.121791.

BIMBER, O. et RASKAR, R., 2005. **Spatial augmented reality: merging real and virtual worlds**. Wellesley, Mass : Peters. ISBN 978-1-56881-230-4.

BIMBER, O. et RASKAR, R., 2006. **Modern approaches to augmented reality**. In : *ACM SIGGRAPH 2006 Courses on - SIGGRAPH '06*. ACM Press, Boston, Massachusetts, USA.

BLANCO-PONS, S., CARRIÓN-RUIZ, B., LUIS LERMA, J. et VILLAVERDE, V., 2019. **Design and implementation of an augmented reality application for rock art visualization in Cova dels Cavalls (Spain)**. In : *Journal of Cultural Heritage*. DOI 10.1016/j.culher.2019.03.014.

BOEHLER, W., HEINZ, G. et MARBS, A., 2001. **The potential of Non-Contact Close Range Laser Scanners for Cultural Heritage Recording**. In : *Proceedings of the 18th CIPA Symposium*. Postdam, Germany.

BOGOSLAVSKYI, I. et STACHNISS, C., 2017. **Analyzing the quality of matched 3D point clouds of objects**. In : *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. Vancouver, BC. IEEE, p. 6685-6690.

BONORA, V., TUCCI, G. et VACCARO, V., 2005. **3D data fusion and multi-resolution approach for a new survey aimed to a complete model of Rucellai's chapel by Leon Battista Alberti in Florence**. In : *Proc. of CIPA XX Int. Symposium, Turin, Italy*.

BORDIN, P., 2002. **SIG: concepts, outils et données**. Hermès Science Publications, Lavoisier. Paris, France. ISBN 978-2-7462-0554-3.

BORST, P., 1997. **Construction of engineering ontologies for knowledge sharing and reuse**. Thèse de doctorat. Enschede, Centre for Telematics and Information Technology (CTIT).

BOTSCH, M., HORNUNG, A., ZWICKER, M. et KOBBELT, L., 2005. **High-Quality Surface Splatting on Today's GPUs**. In : *SPBG'05, Proceedings of the 2nd Eurographics - IEEE VGTC conference on Point-Based Graphics,* p. 17-24. Eurographics Association, New-York, USA.

BOTTECCHIA, S., 2010. **Réalité augmentée et NTIC au service des opérateurs et des experts dans le cadre d'une tâche de maintenance industrielle supervisée**. Thèse de doctorat. Toulouse, France : Université Toulouse III - Paul Sabatier, Ed. MITT.

BOUCHENY, C., ALEJ et ON, R., 2011. [en ligne] **Eye-Dome Lighting: a non-photorealistic shading technique**. [Consulté le 3 juillet 2019]. Disponible à l'adresse : https://blog.kitware.com/eye-dome-lighting-a-non-photorealistic-shading-technique/.

BOUCHER, A. et LE, T-L., 2005. **Comment extraire la sémantique d'une image?** In : *Conference Internationale Sciences Électroniques, Technologies de l'Information et des Télécommunications (SETIT'05)*. Sousse, Tunisia, p. 295–306, IEEE. ISBN: 9973-51-546-3.

BOUDON, P., 1987. **Une architecture mesurée**. In : *L'objet architecture*. 1987. Vol. XLIII, n° 476-477, p. 121-133.

BOUDON, P., GUILLERME, J. et TABOURET, R., 1974. **Figuration graphique en architecture**. *2, Le Théâtre de la figuration*. D.G.R.S.T ; A.R.E.A.

BOUDRIAS, S., 2012. [en ligne] **Le schéma Data-URI**. [Consulté le 12 novembre 2019]. Disponible à l'adresse : https://www.alsacreations.com/article/lire/1439-data-uri-schema.html.

BOURDIER, C., ABGRALL, A., HUARD, O., BRUN, E. Le, PEYROUX, M. et PINÇON, G., 2011. **Histoires de bisons et de chevaux : regard sur l'évolution de la frise pariétale de Cap-Blanc (Marquay, Dordogne) à travers l'analyse du panneau de l'alcôve**. In : *PALEO. Revue d'archéologie préhistorique*, n° 21, p. 17-38.

BOURDIER, C., FUENTES, O. et PINÇON, G., 2015. **Contribution of 3D technologies to the analysis of form in late Paleolithic rock carvings: The case of the Roc-aux-Sorciers rock-shelter (Angles-sur-l'Anglin, France)**. In : *Digital Applications in Archaeology and Cultural Heritage*. Vol. 2, n° 2-3, p. 140-154. DOI 10.1016/j.daach.2015.05.001.

BRADSKI, G., 2019. [en ligne] **opencv/opencv**, Disponible à l'adresse : https://github.com/opencv/opencv.

BROT, J., 2005. [en ligne] **L'utilisation des reliefs naturels dans l'art gravé et sculpté pariétal du paléolithique supérieur français**. Thèse de Master, Paris : École Pratique des Hautes Études.

BUNDGAARD, J. A., 1957. **Mnesicles: A Greek Architect At Work**. Gyldendal. Scandinavian University Books.

BURRI, M., NIKOLIC, J., GOHL, P., SCHNEIDER, T., REHDER, J., OMARI, S., ACHTELIK, M. W et SIEGWART, R., 2016. **The EuRoC micro aerial vehicle datasets**. In : *The International Journal of Robotics Research*. DOI 10.1177/0278364915620033.

BUSARAYAT, C., 2010. **La maquette numérique comme support pour la recherche visuelle d'informations patrimoniales: définition d'une approche pour la sémantisation de sources iconographiques par référencement spatial**. Thèse de doctorat. Aix-en-Provence : Arts et Métiers ParisTech.

CAMPANARO, D., LANDESCHI, G., DELL'UNTO, N. et LEANDER TOUATI, A-M., 2016. **3D GIS for cultural heritage restoration: A 'white box' workflow**. In : *Journal of Cultural Heritage*. Vol. 18, p. 321-332. DOI 10.1016/j.culher.2015.09.006.

CARRERO-PAZOS, M., VILAS-ESTÉVEZ, B. et VÁZQUEZ-MARTÍNEZ, A., 2018. **Digital imaging techniques for recording and analysing prehistoric rock art panels in Galicia (NW Iberia)**. In : *Digital Applications in Archaeology and Cultural Heritage*. Vol. 8, p. 35-45. DOI 10.1016/j.daach.2017.11.003.

CARTAILHAC, E., 1902. **Les cavernes ornées de dessins. La grotte d'Altamira, Espagne. « Mea culpa » d'un sceptique**. MASSON ET CIE (éd.). Vol. 13, p. 348-354.

CASEY, Y., 2016. [en ligne] I**ntroducing the WebVR 1.0 API Proposal.** [Consulté le 31 octobre 2019]. Disponible à l'adresse : https://hacks.mozilla.org/2016/03/introducing-the-webvr-1-0-api-proposal.

CASSEN, S., ROBIN, G., 2010. **Recording art on Neolithic stelae and passage tombs from digital photographs**. Journal of Archaeological Mehod and Theory, vol.17 (1), p1-14. DOI 10.1007/s10816-009-9075-y.

CATALDI, G., ROSSI, M. et MANDELLI, E., 2000. **Luigi Vagnetti: architetto (Roma, 1915- 1980) : disegni, progetti, opere : catalogo della Mostra**. Firenze, Italy, Alinea. Studi e documenti di architettura, 21. ISBN 978-88-8125-376-0.

CAUDELL, T.P. et MIZELL, D.W., 1992. **Augmented reality: an application of heads-up display technology to manual manufacturing processes**. IEEE. p. 659-669 vol.2.

CCIG, 2012. [en ligne] **GéoInformations - Espace interministériel de l'information géographique**. [Consulté le 29 novembre 2019].Disponible à l'adresse :http://www.geoinformations.developpement-durable.gouv.fr.

CERRILLO-CUENCA, E., ORTIZ-CODER, P. et MARTÍNEZ-DEL-POZO, J-Á., 2014. **Computer vision methods and rock art: towards a digital detection of pigments**. In : *Archaeological and Anthropological Sciences*. Vol. 6, n° 3, p. 227-239. DOI 10.1007/s12520-013-0147-2.

CHABREDIER, L., 1966. **Étude méthodologique des relevés d'art pariétal préhistorique**. In : *Bulletin de la Société préhistorique française. Études et travaux*. Vol. 63, n° 3, p. 501-512. DOI 10.3406/bspf.1966.4082.

CHEW, L., 1993. **Guaranteed-quality mesh generation for curved surfaces**. In : *Proceedings of the ninth annual symposium on Computational geometry - SCG '93*. San Diego, California, USA. ACM Press. p. 274-280.

CHOEY, F. et PAOLI, M., 2005. **La lettre à Léon X**. Paris : Éd. de l'Imprimeur. Collection tranches de villes. ISBN 978-2-910735-93-7.

CHOUDARY, O., CHARVILLAT, V., GRIGORAS, R. et GURDJOS, P., 2009. **MARCH: mobile augmented reality for cultural heritage**. In : *Proceedings of the 17th ACM international conference on Multimedia*. ACM. p. 1023–1024.

CIE, 1931. [en ligne] **ISO 9241-302:2008 - CIE RGB 1931** [Consulté le 3 août 2019] Disponible à l'adresse : https://www.iso.org/obp/ui#iso:std:iso:9241:-302:ed-1:v1:en:term:3.2.1.

CNP, 2015. **Commission Nationale des Monuments Historiques - 6ème section**. Ministère de la Culture et de la Communication, Paris, France.

CNRTL, 2012. [en ligne] **SÉMANTIQUE : Définition de SÉMANTIQUE**. In : *Centre National de Ressources Textuelles et Lexicales*. [Consulté le 19 juillet 2019]. Disponible à l'adresse: https://www.cnrtl.fr/definition/Semantique

CNRTL, 2012. [en ligne] **MODALITÉ : Définition de MODALITÉ** . In : *Centre National de Ressources Textuelles et Lexicales*. [Consulté le 19 juillet 2019]. Disponible à l'adresse: https://www.cnrtl.fr/definition/Modalite

COMAR, P., 1992. **La perspective en jeu**. Gallimard. France. Découvertes. ISBN 2-07-053185-6.

COSENTINO, A., 2015. **Practical notes on ultraviolet technical photography for art examination**. In : *Conservar Património*. Vol. 21, p. 53-62. DOI 10.14568/cp2015006.

COULTON, J. J., 1984. **Incomplete preliminary planning in greek architecture: some new evidence**. In : *Actes du colloque de Strasbourg*. Strasbourg, p. 103-112.

CREAP, 2018. [en ligne] **Méthodes appliquées aux relevés graphiques**. In : *Centre de Recherche et d'Études pour l'Art Préhistorique*. [Consulté le 23 octobre 2019]. Disponible à l'adresse : http://www.creap.fr/methodes_releves.htm.

DANGAMPOLA, D., 2015. [en ligne] **The Decorrelation Stretching algorithm**. [Consulté le 26 septembre 2019]. Disponible à l'adresse : http://dhanushkadangampola.blogspot.com /2015/02/decorrelation-stretching.html.

DE LUCA, L., 2009. **La photomodélisation architecturale, Relevé, modélisation et représentation d'édifices à partir de photographies**. Eyrolles. France. ISBN 978-2-212-12524-5.

DE LUCA, L., BUSARAYAT, C., STEFANI, C., RENAUDIN, N., FLORENZANO, M. et VÉRON, P., 2010. **An Iconography-Based Modeling Approach for the Spatio-Temporal Analysis of Architectural Heritage**. In : *2010 Shape Modeling International Conference*. Aix-en-Provence, France : IEEE, p. 78-89.

DE LUCA, L., PIERROT-DESSEILLIGNY, M., MANUEL, A., CHEVRIER, C., LOILLIER, B., BENISTANT, P., PAMART, A., PETELER, F., ABERGEL, V. et ALAOUI, A., 2017. [en ligne] **Aïoli – A reality-based 3D annotation cloud platform for the collaborative documentation of cultural heritage artefacts**. Marseille, France. Disponible à l'adresse : http://www.aioli.cloud/.

DE NUNZIO, G., 2018. **A software tool for the semi-automatic segmentation of architectural 3D models with semantic annotation and Web fruition**. In : *ACTA IMEKO*. Vol. 7, n° 3, p. 64. DOI 10.21014/acta_imeko.v7i3.592.

DELLUC, B., DELLUC, G., LEJEUNE, M. et SOLEILHAVOUP, F., 1986a. **Conventions de relevé analytique des surfaces rocheuses décorées par l'homme préhistorique**. In : *Bulletin de la Société Historique et Archéologique du Périgord Périgueux*. Vol. 113, n° 1, p. 17-22.

DELLUC, B., DELLUC, G., LEJEUNE, M. et SOLEILHAVOUP, F., 1986b. **Lecture analytique des supports rocheux et relevé synthétique (supplément)**. In : *L'Anthropologie (Paris)*. Vol. 90, n° 4, p. 801-804.

DESEILLIGNY, M. et CLÉRY, I., 2011. **Apero, an open source bundle adjusment software for automatic calibration and orientation of set of images**. In : *Proceedings of the ISPRS Symposium, 3DARCH11*, Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XXXVIII-5/W16, p. 269-276. DOI 10.5194/isprsarchives-XXXVIII-5-W16-269-2011.

DEVERIA, A., 2019. [en ligne] **Can I use # WebXR API**. [Consulté le 21 août 2019]. Disponible à l'adresse : https://caniuse.com/#search=xr.

DEVLIN, K., CHALMERS, A. et BROWN, D., 2003. **Predictive lighting and perception in archaeological representations**. In : *UNESCO World Heritage in the Digital Age 30th Anniversary Digital Congress*.

DÍAZ-GUARDAMINO, M., GARCÍA SANJUÁN, L., WHEATLEY, D. et RODRÍGUEZ ZAMORA, V., 2015. **RTI and the study of engraved rock art: A re-examination of the Iberian south-western stelae of Setefilla and Almadén de la Plata 2 (Seville, Spain)**. In : *Digital Applications in Archaeology and Cultural Heritage*. Vol. 2, n° 2-3, p. 41-54. DOI 10.1016/j.daach.2015.07.002.

DIMITRIOU, N., DROSOU, A. et TZOVARAS, D., 2016. **Scan4Reco: Towards the Digitized Conservation of Cultural Heritage Assets via Spatiotemporal (4D) Reconstruction and 3D Printing**. CATALANO, C. et DE LUCA, L. (éd.), *Eurographics Workshop on Graphics and Cultural Heritage - GCH 2016 Proceedings*. Genova, Italy. p. 53-56.

DROST, B., ULRICH, M., NAVAB, N. et ILIC, S., 2010. **Model globally, match locally: Efficient and robust 3D object recognition**. In : *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. San Francisco, CA, USA : IEEE, p. 998-1005.

DUFAYET, C., 2017. [en ligne] **Méthodes et techniques de relevé de l'art préhistorique**. [Consulté le 22 octobre 2019]. Disponible à l'adresse : https://www.hominides.com/html/art/method e-releve-art-prehistorique-avant-1966.php.

DUFFY, S., 2010. **Polynomial texture mapping at Roughting Linn rock art site**. In : *Proceedings of the ISPRS Commission V Mid-Term Symposium: Close Range Image Measurement Techniques*. ISPRS Archives. 38. p. 213–217.

DYER, J., VERRI, G. et CUPITT, J., 2013. **Multispectral imaging in reflectance and photo-induced luminescence modes: a user manual**. London, UK. CHARISMA Project, The British Museum.

EL-HAKIM, S.F., BERALDIN, J.-A., PICARD, M., et GODIN, G., 2004. **Detailed 3D reconstruction of large-scale heritage sites with integrated techniques**. In : *IEEE Computer Graphics and Applications*. Vol. 24, n° 3, p. 21-29. DOI 10.1109/MCG.2004.1318815.

ENGEL, J., KOLTUN, V. et CREMERS, D., 2017. **Direct Sparse Odometry**. In : *IEEE Transactions on Pattern Analysis and Machine Intelligence*. Vol. 40, n° 3, p. 611-625. DOI 10.1109/TPAMI.2017.2658577.

ENGEL, J., SCHÖPS, T. et CREMERS, D., 2014. **LSD-SLAM: Large-Scale Direct Monocular SLAM**. In : FLEET, D., PAJDLA, T., SCHIELE, B. et TUYTELAARS, T. (éd.), *Computer Vision – ECCV 2014*. Springer International Publishing. p. 834-849. ISBN 978-3-319-10604-5.

ESTEVEZ, D., 2001. **Dessin d'architecture et infographie: l'évolution contemporaine des pratiques graphiques**. Paris, France. CNRS Éditions. ISBN 978-2-271-05799-0.

FERUGLIO, V., DUTAILLY, B., BALLADE, M., BOURDIER, C., FERRIER, C., KONIK, S., LACANETTE, D., MORA, P., VERGNIEUX, R. et JAUBERT, J., 2013. **Un outil de relevés 3D partagé en ligne : premières applications pour l'art et la taphonomie des parois ornées de la grotte de Cussac (ArTaPOC / programme LaScArBx)**. VERGNIEUX, R. et DELEVOIE, C. (éd.), Virtual retrospect 2013. Pessac, France : Ausonius. p. 49-54.

FISCHLER, M., et BOLLES, R., 1981. **Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography**. In : *Communications of the ACM*. Vol. 24, n° 6, p. 381-395. DOI 10.1145/358669.358692.

FONTOURA, P., SCHAEFFER, J-M et MENU, M., 2019. **The vision and interpretation of paintings: bottom-up visual processes, top-down culturally informed attention, and aesthetic experience**. In : *Proceedings of the 11th ACM Symposium on Eye Tracking Research & Applications - ETRA '19*. Denver, Colorado, USA. ACM Press. p. 1-3.

FORTEA PÉREZ, F., 1989. **Cien años después de Sautuola: estudios en homenaje a Marcelino Sanz de Sautuola en el centenario de su muerte**. Santander : Diputación Regional de Cantabria, Consejería de Cultura, Educación y Deporte. Estudios de Cantabria, 1. ISBN 978-84-85349-86-9.

FRAENKEL, B., 2007. **L'invention de l'art pariétal préhistorique: Histoire d'une expérience visuelle**. In : *Gradhiva*, n° 6, p. 18-31. DOI 10.4000/gradhiva.984.

FRITZ, C. et TOSELLO, G., 2007. **The Hidden Meaning of Forms: Methods of Recording Paleolithic Parietal Art**. In : *Journal of Archaeological Method and Theory*. Vol. 14, n° 1, p. 48-80. DOI 10.1007/s10816-007-9027-3.

FRITZ, C., TOSELLO, G., AZÉMA, M., MOREAU, O., PERAZIO, G. et PÉRAL, J., 2010. **Restauration virtuelle de l'art pariétal paléolithique : le cas de la grotte de Marsoulas**. In : *In Situ,* n° 13. DOI 10.4000/insitu.6774.

FRITZ, C., WILLIS, M., et TOSELLO, G., 2016. **Reconstructing Paleolithic cave art: The example of Marsoulas Cave (France)**. In : *Journal of Archaeological Science: Reports*. Vol. 10, p. 910-916. DOI 10.1016/j.jasrep.2016.05.012.

FUCHS, P., HUGUES, O., et NANNIPIERI, O., 2010. **Proposition d'une Taxonomie Fonctionnelle des Environnements de Réalité Augmentée**. In : *AFRV2010. Cinquièmes Journées de l'Association Française de Réalité Virtuelle et de l'Interaction 3D.* Orsay, France, p. 8.

FUENTES, O., 2017. **La 3D et l'étude de l'art pariétal, ses apports et ses limites**. Centre National de Préhistoire. Périgueux, France. 308 p.

FUENTES, O., LEPELÉ, J. et PINÇON, G., 2019. **Transferts méthodologiques 3D appliqués à l'étude de l'art paléolithique : une nouvelle dimension pour les relevés d'art préhistorique**. In : *In Situ*, n° 39. DOI 10.4000/insitu.21510.

FUENTES-PACHECO, J., ASCENCIO, J. et RENDON-MANCHA, J, 2015. **Visual Simultaneous Localization and Mapping: A Survey**. In : *Artificial Intelligence Review*. Vol. 43, n° 1, p. 55-81. DOI 10.1007/s10462-012-9365-8.

GARRIDO-JURADO, S., MUÑOZ-SALINAS, R., MADRID-CUEVAS, F.J. et MEDINA-CARNICER, R., 2016. **Generation of fiducial marker dictionaries using Mixed Integer Linear Programming**. In : *Pattern Recognition*. Vol. 51, p. 481-491. DOI 10.1016/j.patcog.2015.09.023.

GEIGER, A., LENZ, P. et URTASUN, R., 2012. **Are we ready for Autonomous Driving? The KITTI Vision Benchmark Suite**. In : *Conference on Computer Vision and Pattern Recognition (CVPR)*. Providence, RI, pp. 3354-3361. DOI 10.1109/CVPR.2012.6248074

GENESTE, J-M, 2006. **Les grottes ornées et les sites d'art pariétal paléolithique**. In : *Les Grottes ornées*. Éditions du Patrimoine. Monumental, Revue scientifique et technique des monuments historiques, 2. p. 14-18. ISBN 978-2-85822-907-9.

GERMANN, G., 1991. **Vitruve et le vitruvianisme, introduction à l'histoire de la théorie architecturale**. Presses Polytechniques et Universitaires Romandes. Lausanne, Suisse. Architecture. ISBN 2-88074-210-2.

GHOUAIEL, N., 2014. **Ingénierie de la conception de systèmes de réalité augmentée mobile : applications de la réalité augmentée mobile au tourisme culturel**. Thèse de doctorat. Toulouse, France. Université Toulouse 3.

GONZALEZ, R., et WOODS, R., 2018. **Digital image processing**. Pearson, New York, USA. 976 p. ISBN 978-0-13-335672-4.

GOOGLE-AR, 2017a. [en ligne] **WebARonARCore**. Disponible à l'adresse : https://github.com/google-ar/WebARonARCore.

GOOGLE-AR, 2017b. [en ligne] **WebARonARKit**. Disponible à l'adresse : https://github.com/google-ar/WebARonARKit.

GRAPP (éd.), 1993. **L'art pariétal paléolithique: techniques et méthodes d'étude**. Paris : Ministère de l'Enseignement supérieur et de la Recherche, Comité des travaux historiques et scientifiques. Documents préhistoriques, 5. ISBN 978-2-7355-0286-8.

GRILLI, E., DININNO, D., PETRUCCI, G. et REMONDINO, F., 2018. **From 2D to 3D supervised segmentation and classification for cultural heritage applications**. In : *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. Vol. XLII-2, p. 399-406. DOI 10.5194/isprs-archives-XLII-2-399-2018.

GRIMAUD, V., 2015. **Réexplorer et valoriser les architectures funéraires monumentales du Néolithique de l'Ouest de la France, étude du cairn de Gavrini*s***. Thèse de doctorat. Nantes, France. Université de Nantes, 2 vol.

GROENEN, M., 1999. **Du relevé de la représentation à la lecture du dispositif pariétal**. Vol. Anthropologie et Préhistoire, n° 110, p. 3-23.

GROENEN, M., 2018. [en ligne] **Reproduire l'art des grottes ornées paléolithiques : du relevé au facsimilé**. In : *Revue et encyclopédie multimédia des arts, sous l'égide de l'académie royale de Belgique*. [Consulté le 29 mai 2019]. Disponible à l'adresse : http://koregos.org/fr/marc-groenen-reproduire-art-grottes-ornees-paleolithiques-du-releve-fac-simile/13362/.

GRUBB, H., 1901. **A new collimating-telescope gunsight for large and small ordnance**. Scientific transactions of the Royal Dublin Society. Dublin, Irland.

GRUEN, A., REMONDINO, F. et ZHANG, L., 2005. **The Bamiyan project: multi-resolution image-based modeling**. In : *Recording, Modeling and Visualization of Cultural Heritage*. Vol. 415, n° 39208, p. 45–54.

GUIDI, G., RUSSO, M., ERCOLI, S., REMONDINO, F., RIZZI, A. et MENNA, F., 2009. **A Multi-Resolution Methodology for the 3D Modeling of Large and Complex Archeological Areas**. In : *International Journal of Architectural Computing*. Vol. 7, n° 1, p. 39-55. DOI 10.1260/147807709788549439.

HARRIS, C. et STEPHENS, M., 1988. **A combined corner and edge detector**. In : *Proceedings of the 4th Alvey vision conference*. p. 147-151. DOI 10.5244/C.2.23.

HARTLEY, R. et ZISSERMAN, A., 2003. **Multiple view geometry in computer vision**. 2nd ed. Cambridge, UK-New York. Cambridge University Press. ISBN 978-0-521-54051-3.

HAVARD, V., 2018. **Développement de méthodes et outils basés sur la réalité augmentée et virtuelle pour l'assistance ou l'apprentissage d'opérations dans un contexte industriel**. Thèse de doctorat. Rouen, France : Université de Rouen.

HAVEMANN, S., SETTGAST, V., BERNDT, R., EIDE, Ø. et FELLNER, D., 2009. **The arrigo showcase reloaded—towards a sustainable link between 3D and semantics**. In : *Journal on Computing and Cultural Heritage*. Vol. 2, n° 1, p. 1-13. DOI 10.1145/1551676.1551680.

HÉNO, R. et CHANDELIER, L., 2014. **Numérisation 3D de bâtiments : cas des édifices remarquables**. ISTE, Collection SIG et Intelligence territoriale, 174 p. ISBN 978-1-78405-034-4.

HERRERA, D., KIHWAN K., KANNALA, J., PULLI, K. et HEIKKILA, J., 2014. **DT-SLAM: Deferred Triangulation for Robust SLAM**. In : *2014 2nd International Conference on 3D Vision*. Tokyo. IEEE, p. 609-616. DOI 10.1109/3DV.2014.49.

HINTERSTOISSER, S., LEPETIT, V., ILIC, S., HOLZER, S., BRADSKI, G., KONOLIGE, K. et NAVAB, N., 2013. **Model Based Training, Detection and Pose Estimation of Texture-Less 3D Objects in Heavily Cluttered Scenes**. In : LEE, K., MATSUSHITA, Y., REHG, J. et HU, Z. (éd.), *Computer Vision – ACCV 2012*. Berlin, Heidelberg. Springer, p. 548-562. ISBN 978-3-642-37330-5.

HOFFMEISTER, D., 2017. **Simulation of tallow lamp light within the 3D model of the Ardales Cave, Spain**. In : *Quaternary International*. Vol. 430, p. 22-29. DOI 10.1016/j.quaint.2016.05.010.

HOOTSUITE MEDIA INC, 2019. [en ligne] **Digital in 2019 - Social Media Marketing & Management Dashboard**. [Consulté le 13 août 2019]. Disponible à l'adresse : https://hootsuite.com/fr/pages/digital-in-2019.

HOUGH, P., 1962. **Method and means for recognizing complex patterns**. Patent US3069654A.

HUREL, A., 2013. [en ligne] **Les peintures préhistoriques de la grotte d'Altamira à Santillane (Espagne)**. Bibnum, Sciences Humaines et Sociales. Disponible à l'adresse: https://journals.openedition.org/bibnum/709

JACOBSEN, K., 1998. **Block adjustment**. In : *University of Hannover. Viitattu*. Vol. 27, p. 2018.

JACQUOT, K., 2014. **Numérisation et restitution virtuelle des maquettes de la collection de Louis XIV, le cas des fortifications bastionnées**. Thèse de doctorat. Nancy, France : Université de Lorraine.

KALMAN, R., 1960. **A New Approach to Linear Filtering and Prediction Problems**. In : *Journal of Basic Engineering*. Vol. 82, n° 1, p. 35-45. DOI 10.1115/1.3662552.

KATO, H., 2017. [en ligne] ***ARToolKit***. Washington, USA. HIT Lab. Disponible à l'adresse : https://github.com/artoolkit/artoolkit5.

KAZHDAN, M. et HOPPE, H., 2013. **Screened poisson surface reconstruction**. In : *ACM Transactions on Graphics*. Vol. 32, n° 3, p. 1-13. DOI 10.1145/2487228.2487237.

KHRONOS GROUP, 2019. [en ligne] **WebGL 2.0 Specification**. [Consulté le 4 novembre 2019]. Disponible à l'adresse : https://www.khronos.org/registry/webgl/specs/latest/2.0/.

KIRYAKOV, A., POPOV, B., TERZIEV, I., MANOV, D. et OGNYANOFF, D., 2004. **Semantic annotation, indexing, and retrieval**. In : *Journal of Web Semantics*. Vol. 2, n° 1, p. 49-79. DOI 10.1016/j.websem.2004.07.005.

KLEIN, G. et MURRAY, D., 2007. **Parallel Tracking and Mapping for Small AR Workspaces**. In : *2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality*. Nara, Japan : IEEE, p. 1-10.

KOTOULA, E., ROBINSON, D., et BEDFORD, C., 2018. **Interactive relighting, digital image enhancement and inclusive diagrammatic representations for the analysis of rock art superimposition: The main Pleito cave (CA, USA)**. In : *Journal of Archaeological Science*. Vol. 93, p. 26-41. DOI 10.1016/j.jas.2018.02.012.

KRULL, A., MICHEL, F., BRACHMANN, E., GUMHOLD, S., IHRKE, S. et ROTHER, C., 2015. **6-DOF Model Based Tracking via Object Coordinate Regression**. In : CREMERS, D., REID, I., SAITO, H. et YANG, M-H (éd.), *Computer Vision -- ACCV 2014*. Springer International Publishing. p. 384-399. ISBN 978-3-319-16816-6.

KUNTZ, S., KULPA, R. et ROYAN, J., 2018. **The Democratization of VR-AR**. In : ARNALDI, B., GUITTON, P. et MOREAU, G. (éd.), *Virtual reality and augmented reality: Myths and Realities.* ISTE Ltd/John Wiley and Sons Inc. Hoboken. Computer Engineering Serie. p. 73-122. ISBN 978-1-78630-105-5.

LACANETTE, D., MINDEGUIA, J-C., BRODARD, A., FERRIER, C., GUIBERT, P., LEBLANC, J-C, MALAURENT, P. et SIRIEIX, C., 2017. **Simulation of an experimental fire in an underground limestone quarry for the study of Paleolithic fires**. In : *International Journal of Thermal Sciences*. Vol. 120, p. 1-18. DOI 10.1016/j.ijthermalsci.2017.05.021.

LACANETTE, D., VINCENT, S., SARTHOU, A., MALAURENT, P. et CALTAGIRONE, J-P., 2009. **An Eulerian/Lagrangian method for the numerical simulation of incompressible convection flows interacting with complex obstacles: Application to the natural convection in the Lascaux cave**. In : *International Journal of Heat and Mass Transfer*. Vol. 52, n° 11-12, p. 2528-2542. DOI 10.1016/j.ijheatmasstransfer.2008.12.028.

LAMING-EMPERAIRE, A., 1962. **La signification de l'art rupestre paléolithique : méthodes et applications**. Édition commerciale : Thèse de doctorat : Lettres. Paris. A. et J. Picard.

LANDES, T. et GRUSSENMEYER, P., 2011. **Les principes fondamentaux de la lasergrammétrie terrestre: systèmes et caractéristiques**. In : *Revue XYZ*. n° 128, p. 37-49.

LARNAOUT, D., 2014. **Localisation d'un véhicule à l'aide d'un SLAM visuel contraint**. Thèse de doctorat. Clermont-Ferrand: Université Blaise Pascal - Clermont-Ferrand II.

LE QUELLEC, J-L., HARMAN, J., DEFRASNE, C. et DUQUESNOY, F., 2013. **DStretch® et l'amélioration des images numériques: applications à l'archéologie des images rupestres**. In : *Les Cahiers de l'AARS*. Vol. 16, p. 177–198.

LEROI-GOURHAN, A., 1965. **Préhistoire de l'art occidental**. Mazenod, Collection l'art et les grandes civilisations. Paris, France. 624 p.

LESCOP, L., 2015. **Le Soleil et la pierre**. In : *Semaine du* Golfe [en ligne], CG Morbihan, Vannes, France. Disponible à l'adresse : https://hal.archives-ouvertes.fr/hal-01508973.

LIMA, J., SIMÕES, F., FIGUEIREDO, L. et KELNER, J., 2010. **Model Based Markerless 3D Tracking applied to Augmented Reality**. In : *SBC Journal on 3D Interactive Systems*. Vol. 1.

LIU, H., ZHANG, G. et BAO, H., 2016. **Robust Keyframe-based Monocular SLAM for Augmented Reality**. In : *2016 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. Merida, Mexico. IEEE. p. 1-10.

LO BUGLIO, D., 2016. **Caractérisation de formes architecturales, Une approche expérimentale intégrant complexité et intelligibilité des représentations numériques**. Thèse de doctorat. Bruxelles, Belgique : Université Libre de Bruxelles.

LORBLANCHET, M., 1993a. **Finalités du relevé**. In : *L'art pariétal paléolithique: techniques et méthodes d'étude*. Paris, France : Ministère de l'Enseignement supérieur et de la Recherche, Comité des travaux historiques et scientifiques. Documents préhistoriques, 5. p. 329-337. ISBN 978-2-7355-0286-8.

LORBLANCHET, M., 1993b. **Le support**. In : *L'art pariétal paléolithique: techniques et méthodes d'étude*. Paris, France : Ministère de l'Enseignement supérieur et de la Recherche, Comité des travaux historiques et scientifiques. Documents préhistoriques, 5. p. 69-80. ISBN 978-2-7355-0286-8.

LORBLANCHET, M., 1993c. **Les conventions graphiques**. In : *L'art pariétal paléolithique: techniques et méthodes d'étude*. Paris, France : Ministère de l'Enseignement supérieur et de la Recherche, Comité des travaux historiques et scientifiques. Documents préhistoriques, 5. p. 369-374. ISBN 978-2-7355-0286-8.

LORIOT, B., 2009. **Automatisation de l'acquisition et des post-traitements en numérisation 3D**. Thèse de doctorat. Dijon, France : Université de Bourgogne.

LUCAS, L., LOSCOS, C. et REMION, Y., 2013. **Vidéo 3D: capture, traitement et diffusion**. Hermès Science : Lavoisier. Collection IC2 Traitement du signal et de l'image. ISBN 978-2746245457.

MAERZ, M., et STOPPA, M., 2018. **Understanding ARKit Tracking and Detection - WWDC 2018**. In : *2017 Apple Worldwide Developers Conference*. Developers Conference, session 610. San Jose, CA, USA.

MAIMONE, A., GEORGIOU, A. et KOLLIN, J.., 2017. **Holographic near-eye displays for virtual and augmented reality**. In : *ACM Transactions on Graphics*. Vol. 36, n° 4, p. 1-16. DOI 10.1145/3072959.3073624.

MAIMONE, A., LANMAN, D., RATHINAVEL, K., KELLER, K., LUEBKE, D. et FUCHS, H., 2014. **Pinlight displays: wide field of view augmented reality eyeglasses using defocused point light sources**. In : *ACM SIGGRAPH 2014 Emerging Technologies on - SIGGRAPH '14*. Vancouver, Canada : ACM Press. p. 1-1.

MALZBENDER, T., GELB, D. et WOLTERS, H., 2001. **Polynomial texture maps**. In : *Proceedings of the 28th annual conference on Computer graphics and interactive techniques - SIGGRAPH '01*. ACM Press. p. 519-528.

MANFERDINI, A. et REMONDINO, F., 2010. **Reality-Based 3D Modeling, Segmentation and Web-Based Visualization**. In : IOANNIDES, M., FELLNER, D., GEORGOPOULOS, A. et HADJIMITSIS, D. (éd.), *Digital Heritage*. Berlin, Heidelberg : Springer. p. 110-124. ISBN 978-3-642-16872-7.

MANUEL, A., 2016. **Annotation sémantique 2D/3D d'images spatialisées pour la documentation et l'analyse d'objets patrimoniaux**. Thèse de doctorat. Aix-en-Provence, France : Arts et Métiers ParisTech.

MANUEL, A., DE LUCA, L. et VÉRON, P., 2014. **A Hybrid Approach for the Semantic Annotation of Spatially Oriented Images**. In : *International Journal of Heritage in the Digital Era*. Vol. 3, n° 2, p. 305-320. DOI 10.1260/2047-4970.3.2.305.

MANUEL, A., M'DARHRI, A., ABERGEL, V., ROZAR, F. et DE LUCA, L., 2018. **A semi-automatic 2D/3D annotation framework for the geometric analysis of heritage artefacts**. In : *Proceedings of the 3rd Digital Heritage International Congress (DigitalHERITAGE)*. San Francisco, CA, USA. IEEE, p. 1-7.

MARSHACK, A., 1975. **Exploring the Mind of Ice Age Man**. Washington, USA. National Geographic Society, 89 p.

MASUDA, T., YAMADA, Y., KUCHITSU, N. et IKEUCHI, K., 2006. **Sunlight illumination simulation for archaeological investigation: case study of the Fugoppe Cave**. In : *Proceedings of the Tenth International Conference on Virtual Systems and Multimedia (Softopia Japan)*. Ed. H. Thwaites, IOS Press, Amsterdam, p. 850–859.

MATHWORKS, 2019. [en ligne] **Apply decorrelation stretch to multichannel image - MATLAB decorrstretch**. In : *Matlab documentation*. [Consulté le 26 septembre 2019]. Disponible à l'adresse: https://fr.mathworks.com/help/images/ref/decorrstretch.html

MEDLEY, J., 2018. [en ligne] **Augmented reality for the web**. [Consulté le 10 décembre 2018]. Disponible à l'adresse : https://developers.google.com/web/updates/2018/06/ar-for-the-web.

MESSAOUDI, T., 2017. ***Proposition d'une ontologie de domaine dédiée à l'annotation d'images spatialisées pour le suivi de la conservation du patrimoine culturel bâti***. Thèse de doctorat. Aix-en-Provence, France : Arts et Métiers ParisTech.

MESSAOUDI, T., VÉRON, P., HALIN, G. et DE LUCA, L., 2018. **An ontological model for the reality-based 3D annotation of heritage building conservation state**. In : *Journal of Cultural Heritage*. Vol. 29, p. 100-112. DOI 10.1016/j.culher.2017.05.017.

MILGRAM, P. et KISHINO, F., 1994. **A taxonomy of mixed reality visual displays**. In : *IEICE TRANSACTIONS on Information and Systems*. Vol. 77, n° 12, p. 1321–1329.

MILGRAM, P., TAKEMURA, H., UTSUMI, A. et KISHINO, F., 1995. **Augmented reality: a class of displays on the reality-virtuality continuum**. In : DAS, H. (éd.), *Photonics for Industrial Applications*. Boston, USA. p. 282-292.

MINISTÈRE DE LA CULTURE, 2019. [en ligne] **Protection des grottes ornées**. [Consulté le 27 novembre 2019]. Disponible à l'adresse : /Sites-thematiques/Archeologie/Conserver-proteger/Protection-des-grottes-ornees.

MINISTÈRE DE LA CULTURE et CNRA, 2017. [en ligne] **Programmation nationale de la recherche archéologique**. [Consulté le 5 novembre 2019]. Disponible à l'adresse : https://www.culture.gouv.fr /Sites-thematiques/Archeologie/Etude-recherche/La-programmation-nationale-de-la-recherche-archeologique

MORTET, V., 1900. **Un ancien devis languedocien: marché pour la reconstruction du campanile de l'église de la Dalbade à Toulouse**. In: *Annales du Midi : revue archéologique, historique et philologique de la France méridionale*, Tome 12, n°46, 1900. pp. 209-220. DOI 10.3406/anami.1900.6707.

MOZILLA, 2017. [en ligne] ***webxr-ios***. Swift. Mozilla Mobile. Disponible à l'adresse : https://github.com/mozilla-mobile/webxr-ios.

MUJA, M. et LOWE, D., 2009. **Fast Approximate Nearest Neighbors With Automatic Algorithm Configuration**. In : *Proceedings of the Fourth International Conference on Computer Vision Theory and Applications*. Lisboa, Portugal. SciTePress - Science and and Technology Publications. p. 331-340.

NERURKAR, E., LYNEN, S. et ZHAO, S., 2017. **System and method for concurrent odometry and mapping**. Patent US20170336511A1.

NEWCOMBE, R., IZADI, S., MOLYNEAUX, D., HILLIGES, O., KIM, D., SHOTTON, J., KOHLI, P., FITZGIBBON, A., HODGES, S. et BUTLER, D., 2014. **Mobile camera localization using depth maps**. Patent US8711206B2.

NIGAY, L. et COUTAZ, J., 1996. **Espaces conceptuels pour l'interaction multimédia et multimodale**. In : *Techniques et sciences informatiques, sp. Multimédia et Collecticiel*. Vol. 15, n°9, p. 1195-1225.

NIJBOER, D. et PATTERSON, D., 2016. **Fighting cockpits: in the pilot's seat of great military aircraft from World War I to today**. Minneapolis, Zenith Press. 224 p. ISBN 978-0-7603-4956-4.

NOVAKOVIĆ, P., TASIĆ, N. et HORŇÁK, M., 2017. **Virtual Reconstructions and Computer Visualisations in Archaeological Practice**. Ljubljana University Press, CONPRA series. Slovenia. Vol. 4. ISBN 978-961-237-899-8.

O'BRIEN, E., ROCHE, S-L., MORSE, A., SOURKE, S., TRACEY, L., BYRNE, J., HAWAT, M., COLTON, F., GREENE, S., et KURZMEIER, M., 2019. [en ligne] **3D Cuneiform project - RLM 6A**. [Consulté le 28 novembre 2019]. Disponible à l'adresse : http://dhprojects.maynoothuniversity.ie/2017/3dcuneiform/rlm-6a/.

ORTEU, J-J, 2008. [en ligne] **Calibrage géométrique d'une caméra ou d'un capteur de vision stéréoscopique - Cours**. École des Mines d'Albi, module Imagerie. Disponible à l'adresse: http://www.optique-ingenieur.org/fr/cours/OPI_fr_M04_C01/co/OPI_fr_M04_C01_web.html.

OWEN, S., 2000. **A Survey of Unstructured Mesh Generation Technology**. In : *7th International Meshing Roundtable*. Mai 2000. Vol. 3. p. 239 - 267.

PAILLET, P., 2015. **Art et comportements symboliques au Paléolithique : quelques points de vue actuels**. Toulon, France. Disponible à l'adresse : https://halshs.archives-ouvertes.fr/halshs-01138307.

PALMA, G., CORSINI, M., CIGNONI, P., SCOPIGNO, R. et MUDGE, M., 2010. **Dynamic shading enhancement for reflectance transformation imaging**. In : *Journal on Computing and Cultural Heritage*. Vol. 3, n° 2, p. 1-20. DOI 10.1145/1841317.1841321.

PAMART, A., MORLET, F. et DE LUCA, L., 2019. **A fully automated incremental photogrammetric processing dedicated for collaborative remote-computing workflow**. In : *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. Vol. XLII-2/W9, p. 565-571. DOI 10.5194/isprs-archives-XLII-2-W9-565-2019.

PAMART, A., PONCHIO, F., ABERGEL, V., ALAOUI M'DARHRI, A., CORSINI, M., DELLEPIANE, M., MORLET, F., SCOPIGNO, R. et DE LUCA, L., 2019. **A complete framework operating Spatially-Oriented RTI in a 3D/2D cultural heritage documentation and analysis tool**. In : *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. Vol. XLII-2/W9, p. 573-580. DOI 10.5194/isprs-archives-XLII-2-W9-573-2019.

PANTHER, 2005. [Illustration] **Modélisation de distorsion en coussinet et en barillet créé par une lentille**. Creative Commons 3.0. Disponible à l'adresse: https://commons.wikimedia.org/wiki/File:Lens_distorsion.png

PAPALEO, L., QUERCINI, G., MASCARDI, V., ANCONA, M., TRAVERSO, A. et DE LUMLEY, H., 2011. **Agents and Ontologies for Understanding and Preserving the Rock Art of Mount Bego**. In : *Proceedings of the 3rd International Conference on Agents and Artificial Intelligence*. Rome, Italy : SciTePress - Science and and Technology Publications, p. 288–295.

PAVONI, G., DELLEPIANE, M., CALLIERI, M. et SCOPIGNO, R., 2016. **Automatic Selection of Video Frames for Path Regularization and 3D Reconstruction**. In : *Proceedings GCH 2016*. Genova, Italy.The Eurographics Association. p. 1-10.

PÉLADAN, J., 1910. **Les Manuscrits de Léonard de Vinci - Les 14 Manuscrits de l'Institut de France**. Paris, France : E. Sansot et Cie. 390 p.

PETELER, F., GATTET, E., BROMBLET, P., GUILLON, O., VALLET, J-M et DE LUCA, L., 2015. **Analyzing the evolution of deterioration patterns: A first step of an image-based approach for comparing multitemporal data sets**. In : *Digital Heritage 2015*. IEEE. p. 113–116. DOI 10.1109/DigitalHeritage.2015.7419465

PIERDICCA, R., FRONTONI, E., ZINGARETTI, P., MALINVERNI, E., COLOSI, F. et ORAZI, R., 2015. **Making Visible the Invisible. Augmented Reality Visualization for 3D Reconstructions of Archaeological Sites**. In : DE PAOLIS, L., et MONGELLI, A. (éd.), *Augmented and Virtual Reality*. Springer International Publishing. Lecture Notes in Computer Science. p. 25-37. ISBN 978-3-319-22887-7.

PIERROT-DESEILLIGNY, M., DE LUCA, L. et REMONDINO, F., 2011. **Automated image-based procedures for accurate artifacts 3D modeling and orthoimage generation**. In : *Geoinformatics FCE CTU*. Vol. 6, p. 291–299. DOI 10.14311/gi.6.36.

PIETTE, E., 1907. **L'art pendant l'Âge du renne**. Masson, Paris, France. 112 p.

PIGEAUD, R., 2004. **L'art paléolithique est-il un art pompier ? Ou le triomphe de l'abbé Breuil**. In : *Un siècle de construction du discours scientifique en Préhistoire*. Avignon, France : Société Préhistorique Française avec le concours du Ministère de la Culture et de la Communication. p. 167-184.

PIGEAUD, R., BOUCHARD, M. et LAVAL, E., 2004. **La grotte ornée Mayenne-Sciences (Thorigné-en-Charnie, Mayenne) : un exemple d'art pariétal d'époque gravettienne en France septentrionale**. In : *Gallia préhistoire*. Vol. 46, n° 1, p. 1-154. DOI 10.3406/galip.2004.2040.

PINÇON, G., BOURDIER, C., FUENTES, O. et ABGRALL, A., 2010. **De la manipulation des images 3D**. In : *In Situ,* n° 13. DOI 10.4000/insitu.6814.

PINÇON, G., FUENTES, O. et BOURDIER, C., 2018. **Sortir de la grotte. L'apport de l'étude des abris ornés du paléolithique supérieur**. In : *Les nouvelles de l'archéologie*. N° 154, p. 82-87. DOI 10.4000/nda.5500.

PLISSON, H. et ZOTKINA, L., 2015. **From 2D to 3D at macro- and microscopic scale in rock art studies**. In : *Digital Applications in Archaeology and Cultural Heritage*. Vol. 2, n° 2-3, p. 102-119. DOI 10.1016/j.daach.2015.06.002.

PLOYON, E., SADIER, B., DELANNOY, J-J, JAILLET, S., MONNEY, J., BOCHE, E., et GENESTE, J-M., 2010. **Le SIG comme outil fédérateur de recherche interdisciplinaire: application à la grotte Chauvet-Pont-d'Arc (Vallon-Pont-d'Arc, Ardèche, France)**. In : *Actes des 2èmes Journées d'informatique et archéologie de Paris (JIAP)*. Paris, France, p. 97-112.

PONCHIO, F., CALLIERI, M., DELLEPIANE, M. et SCOPIGNO, R., 2019. **Effective Annotations Over 3D Models**. In : *Computer Graphics Forum*, 39: 89-105. DOI 10.1111/cgf.13664.

PONCHIO, F., POTENZIANI, M. et CIGNONI, P., 2019. [en ligne] **cnr-isti-vclab/relight**. JavaScript. Visual Computing Lab - ISTI - CNR. Disponible à l'adresse : https://github.com/cnr-isti-vclab/relight.

PORTER, S., HUBER, N., HOYER, C., et FLOSS, H., 2016. **Portable and low-cost solutions to the imaging of Paleolithic art objects: A comparison of photogrammetry and reflectance transformation imaging**. In : *Journal of Archaeological Science: Reports*. Vol. 10, p. 859-863. DOI 10.1016/j.jasrep.2016.07.013.

POTENZIANI, M., CALLIERI, M., DELLEPIANE, M., CORSINI, M., PONCHIO, F. et SCOPIGNO, R., 2015. **3DHOP: 3D Heritage Online Presenter**. In : *Computers & Graphics*. Vol. 52, p. 129-141. DOI 10.1016/j.cag.2015.07.001.

QUEYSANNE, B., 2000. **Alberti et Raphaël. Descriptio urbis Romae, ou, Comment faire le portrait de Rome**. Grenoble, France. Plan fixe ; École d'architecture de Grenoble. ISBN 978-2-913419-02-5.

RAY, N., SOKOLOV, D., REBEROL, M., LEDOUX, F. et LÉVY, B., 2017. [en ligne] ***Hexahedral Meshing: Mind the Gap!*** Disponible à l'adresse : https://hal.inria.fr/hal-01551603.

REMONDINO, F., 2011. **Heritage Recording and 3D Modeling with Photogrammetry and 3D Scanning**. In : *Remote Sensing*. Vol. 3, n° 12, p. 1104-1138. DOI 10.3390/rs3061104.

REMONDINO, F., 2014. **3D recording and modelling in archaeology and cultural heritage: theory and best practices**. Oxford, UK. Archaeopress. BAR international series, 2598. ISBN 978-1-4073-1230-9.

REMONDINO, F., DEL PIZZO, S., KERSTEN, T., et TROISI, S., 2012. **Low-Cost and Open-Source Solutions for Automated Image Orientation ? A Critical Overview**. In : IOANNIDES, M., FRITSCH, D., LEISSNER, J., DAVIES, R., REMONDINO, F. et CAFFO, R. (éd.), *Progress in Cultural Heritage Preservation*. Berlin, Heidelberg. Springer. p. 40-54. ISBN 978-3-642-34233-2.

REMONDINO, F. et RIZZI, A., 2010. **Reality-based 3D documentation of natural and cultural heritage sites—techniques, problems, and examples**. In : *Applied Geomatics*. Vol. 2, n° 3, p. 85-100. DOI 10.1007/s12518-010-0025-x.

RIDEL, B., REUTER, P., LAVIOLE, J., MELLADO, N., COUTURE, N. et GRANIER, X., 2014. **The Revealing Flashlight: Interactive Spatial Augmented Reality for Detail Exploration of Cultural Heritage Artifacts**. In : *Journal on Computing and Cultural Heritage*. Vol. 7, n° 2, p. 1-18. DOI 10.1145/2611376.

RIVIÈRE, E., 1897. **La grotte de la Mouthe (Dordogne)**. In : *Bulletins de la Société d'anthropologie de Paris*. Vol. 8, n° 1, p. 302-329. DOI 10.3406/bmsap.1897.5702.

ROBERT, E., EGELS, Y., BOCHE, E., PEYROUX, M., VIALOU, D., PAILLET, P. et VIGEARS, D., 2014. **Applications 3D pour la contextualisation et le relevé d'art en grotte ornée**. Vol. Archeologia e Calcolatori, n° Supplemento 5, p. 88-101.

ROBSON BROWN, K.., CHALMERS, A., SAIGOL, T., GREEN, C., et D'ERRICO, F., 2001. **An Automated Laser Scan Survey of the Upper Palaeolithic Rock Shelter of Cap Blanc**. In : *Journal of Archaeological Science*. Vol. 28, n° 3, p. 283-289. DOI 10.1006/jasc.2000.0574.

RODGER, R., 1998. [en ligne] **The Matrix and Quaternion in OpenGL**. [Consulté le 19 septembre 2019]. Disponible à l'adresse : http://www.flipcode.com/documents/matrfaq.html#Q43.

RÖNNHOLM, P., HONKAVAARA, E., LITKEY, P., HYYPPÄ, H. et HYYPPÄ, J., 2007. **Integration of Laser Scanning and Photogrammetry**. In : *ISPRS Workshop on Laser Scanning 2007 and SilviLaser 2007*. Vol. 36. Espoo, Finland.

RUIZ, J., 2016. **4D, arte rupestre,** Dirección General de Bienes Culturales, Servicio de Patrimonio Histórico. ISBN 978-84-7564-644-2.

RUPNIK, E., DAAKIR, M. et PIERROT-DESEILLIGNY, M., 2017. **MicMac – a free, open-source solution for photogrammetry**. In : *Open Geospatial Data, Software and Standards*. Vol. 2, n° 1. DOI 10.1186/s40965-017-0027-2.

RUSINKIEWICZ, S., BURNS, M. et DECARLO, D., 2006. **Exaggerated shading for depicting shape and detail**. In : *ACM Transactions on Graphics (TOG)*. ACM, New-York, USA. Vol. 25, n°3, p. 1199–1205. DOI 10.1145/1141911.1142015.

SAINT AUBIN, J-P., 1992. **Le relevé et la représentation de l'architecture**. Paris : Inventaire général. Documents et méthodes, 2. ISBN 2-11 087 308-6.

SAKAROVITCH, J., 1998. **Épures d'architecture : de la coupe des pierres à la géométrie descriptive, XVIe - XIXe siècles**. Basel, Birkhäuser. Science networks, 21. ISBN 978-0-8176-5701-7.

SALERI, R., 2018. **Photogrammetric survey of heritage objects: an experimental protocol for optimizing camera placement for aerial surveys**. In : *Proceedings of Digital Heritage 2018*. San Francisco, CA, USA. IEEE. p. 1-7.

SANCHEZ, J., 2019. [en ligne] ***spite/THREE.CubemapToEquirectangular***. JavaScript. Disponible à l'adresse : https://github.com/spite/THREE.CubemapToEquirectangular.

SAUVET, G., 1993. **La composition de l'espace orné**. In : *L'art pariétal paléolithique: techniques et méthodes d'étude*. Paris, France : Ministère de l'Enseignement supérieur et de la Recherche, Comité des travaux historiques et scientifiques. Documents préhistoriques, 5. p. 297-309. ISBN 978-2-7355-0286-8.

SAUVET, G. et TOSELLO, G., 1998. **Le mythe paléolithique de la caverne**. In : *Le Propre de l'homme*. Chap. 3, p. 55-90.

SCALAS, A., MORTARA, M. et SPAGNUOLO, M., 2017. **3D Annotation Transfer**. In : *Eurographics Workshop on Graphics and Cultural Heritage*. p. 5 pages. DOI 10.2312/gch.20171295.

SCHMALSTIEG, D. et HÖLLERER, T., 2016. **Augmented reality: principles and practice**. Addison-Wesley usability and HCI series. Boston Columbus Indianapolis, USA. ISBN 978-0-321-88357-5.

SCHUETZ, M., 2014. [en ligne] **PotreeConverter - Uniform Partitioning of Point Cloud Data into an Octree**. Institute of Computer Graphics and Algorithms, Vienna University of Technology, Vienna, Austria. Disponible à l'adresse : http://potree.org/downloads/converter_documentation.pdf.

SCHUETZ, M., 2016. **Potree: Rendering Large Point Clouds in Web Browsers.** Thèse de Master. Institute of Computer Graphics and Algorithms, Vienna University of Technology, Vienna, Austria.

SEIDL, M. et BREITENEDER, C., 2011. **Detection and Classification of Petroglyphs in Gigapixel Images - Preliminary Results**. In : *VAST: International Symposium on Virtual Reality*. p. 4. DOI 10.2312/pe/vast/vast11s/045-048.

SEIDL, M., WIESER, E. et ALEXANDER, C., 2015. **Automated classification of petroglyphs**. In : *Digital Applications in Archaeology and Cultural Heritage*. Vol. 2, n° 2-3, p. 196-212. DOI 10.1016/j.daach.2015.03.001.

SERNA, S., SCHMEDT, H., RITZ, M., et STORK, A., 2012. **Interactive Semantic Enrichment of 3D Cultural Heritage Collections**. In : *VAST: International Symposium on Virtual Reality*. The Eurographics Association, p. 33–40. DOI 10.2312/VAST/VAST12/033-040.

SHANMUGAM, P. et ARIKAN, O., 2007. **Hardware accelerated ambient occlusion techniques on GPUs**. In : *Proceedings of the 2007 symposium on Interactive 3D graphics and games*. ACM, I3D'07. New-York, USA. p. 73–80. DOI 10.1145/1230100.1230113

SHI, W., KOTOULA, E., AKOGLU, K., YANG, Y., et RUSHMEIER, H., 2016. **CHER-Ob: A Tool for Shared Analysis in Cultural Heritage**. In : *Eurographics Workshop on Graphics and Cultural Heritage*. p. 4 pages. DOI 10.2312/gch.20161404.

SHIH, T. et FAIG, W., 1988. **A solution for space resection in closed form**. In : *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. Vol. 27, n° Part B5, p. 547–556.

SLOAN, P-P., MARTIN, W., GOOCH, A. et GOOCH, B., 2001. **The lit sphere: A model for capturing NPR shading from art**. In : *Graphics interface*. CAN, Ottawa, Ontario, Canada. p. 143–150. DOI 10.5555/780986.781004.

SMEULDERS, A., WORRING, M., SANTINI, S., GUPTA, A. et JAIN, R., 2000. **Content-based image retrieval at the end of the early years**. In : *IEEE Transactions on Pattern Analysis and Machine Intelligence*. Vol. 22, n° 12, p. 1349-1380. DOI 10.1109/34.895972.

SNAVELY, N., SEITZ, S. et SZELISKI, R., 2006. **Photo tourism: exploring photo collections in 3D**. In : *ACM Transactions on Graphics*. Vol. 25, n° 3, p. 835. DOI 10.1145/1141911.1141964.

SOLER, F., MELERO, F., et LUZÓN, M., 2017. **A complete 3D information system for cultural heritage documentation**. In : *Journal of Cultural Heritage*. Vol. 23, p. 49-57. DOI 10.1016/j.culher.2016.09.008.

STATHOPOULOU, E-K. et REMONDINO, F., 2019. **Semantic photogrammetry - boosting image-based 3D reconstruction with semantic labeling**. In : *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. Vol. XLII-2/W9, p. 685-690. DOI 10.5194/isprs-archives-XLII-2-W9-685-2019.

STEFANI, C., 2010. **Maquettes numériques spatio-temporelles d'édifices patrimoniaux : Modélisation de la dimension temporelle et multi-restitutions d'édifice***s*. Thèse de doctorat. Aix-en-Provence: Arts et Métiers ParisTech.

STOLERU, A., 2019. **Towards a domain ontology dedicated to the description of anthropogenic traces in rock art**. In : *DARIAH Annual Event 2019: Humanities Data*. Varsovie, Poland.

STRASDAT, H., MONTIEL, J., et DAVISON, A., 2012. **Visual SLAM: Why filter?** In : *Image and Vision Computing*. Vol. 30, n° 2, p. 65-77. DOI 10.1016/j.imavis.2012.02.009.

STURM, J., BURGARD, W. et CREMERS, D., 2012. **Evaluating Egomotion and Structure-from-Motion Approaches Using the TUM RGB-D Benchmark**. In : *Proc. of the Workshop on Color-Depth Camera Fusion in Robotics at the IEEE/RJS International Conference on Intelligent Robot Systems*.

STURM, P., 2015. [en ligne] **Cours de Vision par Ordinateur** INRIA Rhône-Alpes. Disponible à l'adresse: https://team.inria.fr/steep/files/2015/03/poly_3D.pdf.

TAMAAZOUSTI, M., 2013. **L'ajustement de faisceaux contraint comme cadre d'unification des méthodes de localisation : application à la réalité augmentée sur des objets 3D**. Thèse de doctorat. Clermont-Ferrand, France : Université Clermont-Ferrand 2.

TARINI, M., PIETRONI, N., CIGNONI, P., PANOZZO, D., et PUPPO, E., 2010. **Practical quad mesh simplification**. In : *Computer Graphics Forum*. Vol. 29, n° 2, p. 407-418. DOI 10.1111/j.1467-8659.2009.01610.x.

TEWES, J., 2017. **How Light Field Technology Will Change the Future of Augmented and Mixed Reality**. In : *AWE Europe 2017 - the largest conference for AR+VR*. Keynote. Munich, Germany.

THIVET, M., VERRIEZ, Q., et VURPILLOT, D., 2019. **Aspectus : développement d'une plateforme collaborative en ligne d'analyse des données 3D pour l'archéologie et le patrimoine**. In : *In Situ*, n° 39. DOI 10.4000/insitu.21934.

TOSELLO, G., et FRITZ, C., 2004. **Grotte Chauvet-Pont d'Arc: Approche structurelle et comparative du Panneau des Chevaux**. In : *L'Art du Paléolithique Supérieur*. Liège, Belgique, p. 69-86.

TRIGGS, B., MCLAUCHLAN, P., HARTLEY, R., et FITZGIBBON, A., 2000. **Bundle Adjustment — A Modern Synthesis**. In : TRIGGS, B., ZISSERMAN, A. et SZELISKI, R. (éd.), *Vision Algorithms: Theory and Practice*. Berlin, Heidelberg : Springer. p. 298-372. ISBN 978-3-540-67973-8.

TYLECEK, R. et FISHER, R., 2018. **Consistent Semantic Annotation of Outdoor Datasets via 2D/3D Label Transfer**. In : *Sensors*. Vol. 18, n° 7, p. 2249. DOI 10.3390/s18072249.

ULLMANN, S., 1959. **Précis de sémantique française.** A. Francke. Bibliotheca Romanica, n°9, 350 p.

URIBE, M., et WHEATLEY, D., 2013. **Rock art and digital technologies: the application of reflectance transformation imaging (RTI) and 3D laser scanning to the study of late bronze age iberian stelae**. In : *Menga: Revista de prehistoria de Andalucía*. n° 4, p. 187–203.

VALLET, J-M., DE LUCA, L., et FEILLOU, M., 2012. **Une nouvelle approche spatio-temporelle et analytique pour la conservation des peintures murales sur le long terme**. In : *In Situ,* n° 19. DOI 10.4000/insitu.9829.

VERGNE, R., BOUSSEAU, A., THOLLOT, J., VANDERHAEGHE, D., BARLA, P. et GRANIER, X., 2007. **Utilisation du rendu expressif pour l'illustration et l'exploration de données archéologiques**. In : *Virtual Retrospect 2007: Archéologie et Réalité Virtuelle*. Pessac, France : Ausonius. p. 101–106.

VERGNE, R., PACANOWSKI, R., BARLA, P., GRANIER, X., et SCHLICK, C., 2010. **Radiance scaling for versatile surface enhancement**. In : *Proceedings of the 2010 ACM SIGGRAPH symposium on Interactive 3D Graphics and Games*. ACM, New-York, USA. p. 143–150.

VIAL, S., 2016. **La fin des frontières entre réel et virtuel : vers le monisme numérique**. In : *Frontières numériques et artefacts*. p. 135-146. ISBN 978-2-343-07623-2.

VIALLON, M., et CHATENET, M., 2008. **Construire l'espace au XVIe siècle : actes du XIVe colloque du Puy-en-Velay**. Publications de l'Université de Saint-Etienne, France. ISBN 978-2-86272-481-2.

VILAS-ESTEVEZ, B., VÁZQUEZ-MARTÍNEZ, A., et CARRERO-PAZOS, M., 2017. **Going further: (Re) Discovering rock art carvings with photogrammetric techniques in Galicia (North-West Iberian Peninsula)**. In : *Handbook of Research on Emerging Technologies for Digital Preservation and Information Modeling*. IGI Global. p. 175–200. ISBN 978-1-5225-0680-5.

VOLOSHIN, V., 2018. [en ligne] **Visualizing ARKit Sessions**. [Consulté le 29 août 2019]. Disponible à l'adresse : https://www.ittybittyapps.com/blog/2018-12-03-visualizing-arkit-sessions/.

W3C®, 2019. [en ligne] **WebXR Device API, W3C First Public Working Draft**. [Consulté le 19 août 2019]. Disponible à l'adresse : https://www.w3.org/TR/2019/WD-webxr-20190205/.

W3C, 2018. [en ligne] **WebXR Device API, Editor's Draft**. [Consulté le 10 décembre 2018]. Disponible à l'adresse : https://immersive-web.github.io/webxr/.

W3C, 2019. [en ligne] **Semantic Web - W3C**. [Consulté le 28 novembre 2019]. Disponible à l'adresse : https://www.w3.org/standards/semanticweb/.

WANG, Z., SHI, W., AKOGLU, K., KOTOULA, E., YANG, Y. et RUSHMEIER, H., 2018. **CHER-Ob: A Tool for Shared Analysis and Video Dissemination**. In : *Journal on Computing and Cultural Heritage*. Vol. 11, n° 4, p. 1-22. DOI 10.1145/3230673.

WILL, E., 1984. **La maquette de l'adylon du temple de Niha**. In : *Actes du colloque de Strasbourg*. Strasbourg, p. 277-281.

WINDELS, F., 1948. **Lascaux « Chapelle Sixtine » de la Préhistoire**. Centre d'Études et de Documentation préhistorique, Montignac-sur-Vézère, France. 141 p.

YADAV, N., 2018. [en ligne] **Diagram of augmented reality displays**. Disponible à l'adresse: https://blog.prototypr.io/understanding-display-techniques-in-augmented-reality-c258b911b5c9.

YOUNES, G., ASMAR, D., SHAMMAS, E. et ZELEK, J., 2017. **Keyframe-based monocular SLAM: design, survey, and future directions**. In : *Robotics and Autonomous Systems*. Vol. 98, p. 67-88. DOI 10.1016/j.robot.2017.09.010.

YU, C-H., GROZA, T. et HUNTER, J., 2013. **Reasoning on Crowd-Sourced Semantic Annotations to Facilitate Cataloguing of 3D Artefacts in the Cultural Heritage Domain**. In : SALINESI, C., NORRIE, M., et PASTOR, Ó. (éd.), *Advanced Information Systems Engineering*. Berlin, Heidelberg. Springer. p. 228-243. ISBN 978-3-642-38708-1.

YU, C-H., et HUNTER, J., 2013. **Documenting and sharing comparative analyses of 3D digital museum artifacts through semantic web annotations**. In : *Journal on Computing and Cultural Heritage*. Vol. 6, n° 4, p. 1-20. DOI 10.1145/2532630.2532634.

ZEPPELZAUER, M., POIER, G., SEIDL, M., REINBACHER, C., SCHULTER, S., BREITENEDER, C. et BISCHOF, H., 2016. **Interactive 3D Segmentation of Rock-Art by Enhanced Depth Maps and Gradient Preserving Regularization**. In : *Journal on Computing and Cultural Heritage*. Vol. 9, n° 4, p. 1-30. DOI 10.1145/2950062.

ZHU, Q., WANG, X., KEOGH, E., et LEE, S-H., 2011. **An efficient and effective similarity measure to enable data mining of petroglyphs**. In : *Data Mining and Knowledge Discovery*. Vol. 23, n° 1, p. 91-127. DOI 10.1007/s10618-010-0200-z.

ZOMAHOUN, D., 2015. **Emergsem : une approche d'annotation collaborative et de recherche d'images basée sur les sémantiques émergentes**. Thèse de doctorat. Dijon, France. Université de Bourgogne.

ZOTKINA, L., 2016. **Le bison de La Grèze (Marquay, Dordogne, France) sous le microscope**. In : *PALEO. Revue d'archéologie préhistorique*, n° 27, p. 307-320.

Table des sigles et abréviations

API	*Application Programming Interface* (Interface de Programmation Applicative)
AR	*Augmented Reality* (Réalité Augmentée)
AV	*Augmented Virtuality* (Virtualité Augmentée)
BBA	*Bundle Bloc Adjustment* (Ajustement de Faisceaux)
CH	*Cultural Heritage* (Patrimoine Culturel)
CHE	*Cultural Heritage Entity* (Entité à valeur patrimoniale)
CNP	Centre National de la Préhistoire
CNRTL	Centre National de Ressources Textuelles et Lexicales
EDL	*Eye-Dome Lighting*
FOV	*Field Of View* (Champ de Vision)
GCP	*Ground Control Point* (Points d'Appui)
GUI	*Graphical User Interface* (Interface Graphique Utilisateur)
HMD	*Head Mounted Display* (Dispositif Fixé à la Tête)
HTML	*HyperText Markup Language* (littéralement : Langage de Balisage Hypertexte)
ICP	*Iterative Closest Point*
IHM	Interface Homme-Machine
IMU	*Inertial Measurement Unit* (Centrale Inertielle)
NPR	*Non-Photorealistic Rendering* (Rendu Non-Photoréaliste)
MNT	Modèle Numérique de Terrain
MR	*Mixed Reality* (Réalité Mixte)
PBR	*Physically-Based Rendering* (Rendu Physique Réaliste)
PCA	*Principal Component Analysis* (Analyse en Composantes Principales)
PLY	*Polygon File Format*
RGB	*Red, Green, Blue* (Rouge, Vert, Bleu)
SDK	*Software Development Kit* (Kit de Développement Logiciel)
SFM	*Structure From Motion* (littéralement : Structure Acquise à Partir d'un Mouvement)
SIG	Système d'Information Géographique
SLAM	*Simultaneous Localization And Mapping* (Localisation et Cartographie Simultanées)
SR	*Spatial Resection* (Résection Spatiale)
VR	*Virtual Reality* (Réalité Virtuelle)
XML	*Extensible Markup Language* (littéralement : Language de Balisage Extensible)

Table des illustrations

Figure 1 : Statue de Gudea, prince de Lagash dite « statue de l'architecte au plan ». Photographies © RMN-Grand Palais (Musée du Louvre) / René-Gabriel Ojéda. 9

Figure 2 : Perspective cavalière d'un projet pour le palais des Tuileries par Jacques Androuet du Cerceau en 1579. Illustration : (Babelon, 1989). 12

Figure 3 : Cartographie suivant la méthode d'Alberti. Illustration : (Cataldi et al., 2000). 14

Figure 4 : Schéma présentant les trois moments interreliés du relevé. Illustration : auteur. 15

Figure 5 : À gauche : Breuil relevant une figure pariétale. À droite : André Glory relevant les gravures de Lascaux, accompagné d'Alain Roussot. Illustrations : (Groenen, 2018) 17

Figure 6 : Les différents modes de projection appliqués au relevé d'art pariétal. À gauche : projection sur une surface développable. Au centre : projection orthographique. À droite : projection polyédrique. Illustration : (Aujoulat, 1993a). 19

Figure 7 : Vue d'ensemble du système de visée utilisé pour le relevé indirect. Illustration : (Chabredier, 1966) 19

Figure 8 : Enregistrement d'une coupe de la Salle des Taureaux à Lascaux par la méthode des profils lumineux de Norbert Aujoulat. Illustration : © Ministère de la Culture / Centre National de la Préhistoire / Norbert Aujoulat. 21

Figure 9 : Relevé du bison mugissant du Grand Plafond de la grotte d'Altamira (Cantabrie) par Henri Breuil. À gauche : paroi de la grotte. Au centre : esquisse au crayon. À droite : reproduction au pastel de la figure relevée. Illustration : (Groenen, 2018). 22

Figure 10 : Deux relevés du Mégacéros gravé de Pech-Merle. A. Celui de Breuil, B. Celui de Lorblanchet. Illustrations : (Pigeaud, 2004). 23

Figure 11 : À gauche : Relevé d'un auroch de la grotte d'El Castillo réalisé par Henri Breuil. À droite : l'auroch dans son contexte. Illustration : (Groenen, 2018) 24

Figure 12 : Différentes approches pour la représentation du relief. À gauche : représentation des concrétions en pointillés, extrait du relevé de la grotte Huchard par L. Chabredier et P. Ayrolles. Illustration : (GRAPP, 1993). Au centre : représentation par courbes de niveau, extrait du relevé de la grotte Carriot par M. Lorblanchet. Illustration : (GRAPP, 1993). À droite, utilisation de conventions graphiques de relevé analytique, extrait du relevé d'un mammouth de la grotte de Domme par B. et G. Delluc. Illustration : (Delluc et al., 1986b). 25

Figure 13 : Utilisation d'une photographie comme fond graphique du relevé. Relevé d'un fragment de la grotte de Teyjat par P. Paillet, Photographie de J.-C. Domenech. Illustration : (Ministère de la Culture, CNRA, 2017) 25

Figure 14 : Relevé analytique plastique du panneau des Chevaux de la grotte Chauvet-Pont d'Arc par Carole Fritz et Gilles Tosello. Illustration : © C.Fritz et G.Tosello. 28

Figure 15 : Relevé analytique technique du panneau des Bouquetins du Roc-aux-Sorciers par Ludmila Iakovleva et Geneviève Pinçon. Illustration : (Fuentes et al., 2019) 30

Figure 16 : Relevé analytique technique (ombré) de l'entité graphique n°1 de la frise sculptée du Cap-Blanc, par Aurélie Abgrall, Camille Bourdier, Olivier Huard, Éric Le Brun et Geneviève Pinçon. Illustration : (Bourdier et al., 2011). 31

Figure 17 : Comparaison des deux familles de relevé selon différents critères liés aux trois moments de Saint-Aubin. Il apparaît clairement que les deux approches ne répondent pas aux mêmes besoins. Illustration : auteur. 32

Figure 18 : Schéma des différentes étapes du relevé numérique. Contrairement au relevé *classique*, ce processus est séquentiel. Illustration : auteur. 35

Figure 19 : Les distorsions a) en coussinet, b) en barillet, c) sans distorsion. Illustration : (Panther, 2005)..... 36

Figure 20 : Bruit numérique sur une photographie (facsimilé du panneau des Chevaux). Illustration : auteur 36

Figure 21 : Détail du facsimilé du panneau des Chevaux (griffure). À gauche : lumière directe. À droite : lumière rasante. Illustration : auteur. 38

Figure 22 : Principaux domaines du spectre électromagnétique. Le visible n'en représente qu'une faible portion. Illustration : auteur 38

Figure 23 : Bison peint du Salon Noir (Grotte de Niaux, Ariège, France). À gauche : photographie VIS. À droite : photographie UV. Endommagée en 1978, la figure n'est pas facilement visible aujourd'hui, mais en lumière UV de nombreux détails sont à nouveau perceptibles car les traces de manganèse, conservées sur la paroi de la grotte absorbent les rayons UV tandis que la couche de calcite les reflète. Illustration : (Fritz, Tosello, 2007) 39

Figure 24 : Mammouth dans la grotte de Rouffignac. À gauche : Photographie VIS. À droite : Photographie IRr. Le trait traversant au niveau de la tête, préalablement identifié comme du noir de carbone par spectrométrie Raman, apparaît plus clair que le reste de la figure, dont les traits sont composés d'oxyde de manganèse. Illustration : (Beck et al., 2014). 40

Figure 25 : Capture d'écran d'une image RTI (facsimilé du panneau des chevaux) visualisée dans le logiciel RTIviewer. 41

Figure 26 : Taxonomie non exhaustive des techniques d'acquisition 3D. Illustration : auteur, d'après (Jacquot, 2014 ; Loriot, 2009 ; Remondino, 2011). 42

Figure 27 : Principe du système de balayage du scanner laser. Illustration : auteur. 42

Figure 28 : Capture d'écran du nuage de points de la frise sculptée de l'abri du Cap-Blanc. Données : © AGP/CNP. 43

Figure 29 : Approximation du fonctionnement d'un appareil photo. Illustration : auteur, d'après (Héno, Chandelier, 2014). 44

Figure 30 : Position 3D d'un point apparaissant sur une (A) et deux (B) photographies. Illustration : auteur, d'après (De Luca, 2009). 44

Figure 31 : Deux représentations différentes d'une même image. À partir de l'image gauche, l'être humain est capable de reconnaître immédiatement le contenu d'une image numérique comme une représentation logique et continue, alors que l'image droite représentant les mêmes données d'intensité est bien moins intelligible. Photographie © CNP/MC. 47

Figure 32 : Exemples de traitements manuels d'une image. À gauche : image originale (photographie © CNP/MC). Au centre, traitement en fausses couleurs. À droite : seuillage. Illustration : Auteur..... 48

Figure 33 : À gauche : image d'entrée (Cueva Raton, Californie) et répartition des pixels dans l'espace colorimétrique. À droite : image résultant de son traitement par *decorrelation stretching*. Illustration : d'après Jon Harman. 49

Figure 34 : Méthode proposée par Zeppelzauer et al. (2016) pour la segmentation interactive de pétroglyphes. Illustration : (Zeppelzauer et al., 2016). 50

Figure 35 : À gauche : projection perspective. À droite : projection parallèle. Facsimilé du panneau des Chevaux de la grotte Chauvet Pont-d'Arc. Illustration : auteur. 51

Figure 36 : Les modes de projection panoramique. À gauche : Projection en faces de cube. Au centre : Projection cylindrique. À droite : Projection sphérique. Illustration : auteur, à partir d'une image panoramique de l'abri du Cap-Blanc © CNP/AGP. 52

Figure 37 : Le *web-viewer* 3DHop, utilisé pour le projet 3D Cuneiform. Illustration : (O'brien et al., 2019). 53

Figure 38 : Maillage et texturage d'un nuage de points du facsimilé. Illustration : auteur. 54

Figure 39 : Exemples de cartes de texture caractérisant des attributs géométriques ou visuels de l'objet d'étude. De gauche à droite : couleurs, normales, profondeur, occlusion ambiante, rugosité. 55

Figure 40 : Détail du facsimilé, rendu avec l'occlusion ambiante. 56

Figure 41 : Nuage de points d'un détail du facsimilé, rendu avec le *shader* EDL. 56

Figure 42 : Détail du facsimilé rendu à l'aide de XShade. 57

Figure 43 : Détail du facsimilé rendu à l'aide du *shader* Radiance Scaling. 57

Figure 44 : Simulation de raccord réalisée avec des blocs du Roc-aux-Sorciers. À gauche : modèles 3D virtuels. À droite : impressions 3D. Illustration : (Bourdier et al., 2015). ©CNP / DRAC Poitou-Charentes / AGP. 58

Figure 45 : Comparaison de deux modèles 3D produits à deux mois d'intervalle. L'image révèle des écarts correspondant à des pertes de matière. Abrigo del Buen Aire (Espagne). Illustration : (Ruiz, 2016) 59

Figure 46 : Types d'annotations sémantiques d'images 2D. De gauche à droite : globale, ponctuelle, rectangulaire (boîte englobante), région (groupe de pixels). Photographie : Rhinocéros ZD1-20 du panneau des chevaux. ©CNP. Illustration : auteur, d'après (Tylecek, Fisher, 2018). 62

Figure 47 : Outil d'annotation proposé par Seidl et al. (2015) pour constituer le jeu de données de référence. Illustration : (Seidl et al., 2015). 64

Figure 48 : Relevé sur modèle 3D sur Blender avec l'outil *Grease Pencil*. Illustration : © O. Fuentes / J.Lepelé / CNP. 65

Figure 49 : Résultats expérimentaux de l'approche proposée par Stathopoulou et Remondino (2019). Gauche : Image originale. Centre : Image segmentée sémantiquement. Droite : Nuage de points épars après transfert des labels. Illustration : (Stathopoulou, Remondino, 2019). 67

Figure 50 : Approche de propagation 2D/3D d'annotations, proposée par Manuel (2016). Illustration : (Manuel, 2016). 68

Figure 51 : Exemple issu de l'approche proposée par Messaoudi (2017), deux instances d'une annotation liées au même concept. La description basée sur une ontologie permet d'établir un modèle de corrélation multidimensionnel. Illustration : (Messaoudi, 2017). 68

Figure 52 : Modèle 3D du mammouth du Patriarche (grotte de Rouffignac), et relevé infographique correspondant. Illustration : (Robert et al., 2014). 69

Figure 53 : Logiciel Cumulus, proposé par Robert et al. (2014) pour l'intégration et la consultation de données archéologiques et analytiques à partir du modèle 3D. Illustration : (Robert et al., 2014). 69

Figure 54 : Comparaison des spécificités du relevé numérique et des méthodes de relevé *classiques*. Illustration : auteur. 72

Figure 55 : Schéma de notre approche intégrée. Le relevé d'art pariétal doit être considéré comme un cycle dont les différentes phases sont centrées sur l'objet d'étude, éventuellement à travers sa copie numérique. Illustration : auteur. 73

Figure 56 : Relevé de tracés gravés (grotte de Cussac) à l'aide du *plugin* Meshlab. Illustration : (Feruglio et al., 2013). 74

Figure 57 : Deux approches possibles pour la prise en compte de données 3D dans des SIG. En haut : par relations. En bas : par attribut altimétrique. Illustration : auteur. 75

Figure 58 : Principe d'affectation de la carte d'élévation comme texture du modèle 3D de la façade et exploitation dans un SIG. Illustration : (Campanaro et al., 2016). 77

Figure 59 : Architecture logicielle de CHER-Ob. Illustration : (Shi et al., 2016). 77

Figure 60 : À gauche : exemple de requête qu'Agata peut traiter. À droite : principe d'annotation proposé. Illustration : (Soler et al., 2017). 78

Figure 61 : Structure hiérarchique définie pour la Fontaine de Neptune et segmentation subséquente. Illustration : (Apollonio et al., 2018). 79

Figure 62 : Le référencement spatial d'images autour d'une maquette numérique segmentée leur permet de recevoir des annotations par projection 3D → 2D. Illustration : (Busarayat, 2010). 80

Figure 63 : Principe de l'indexation exploitée par Aïoli pour propager des annotations par la relation 2D → 3D → 2D. Illustration : auteur, d'après (Manuel, 2016). 82

Figure 64 : Interface d'Aïoli, le *viewer* 3D permet de visualiser les images orientées autour d'un nuage de points annoté. 83

Figure 65 : Schéma de l'expérience de Brunelleschi. Illustration : (Argan, Wittkower, 1990). 85

Figure 66 : Continuum Réalité – Virtualité. Illustration : auteur, d'après (Milgram et al., 1995). 86

Figure 67 : Schéma des différentes catégories de dispositifs d'affichage de réalité augmentée. Illustration : auteur. 88

Figure 68 : Schéma des principes d'affichage AR. Illustration : auteur, d'après (Yadav, 2018). 89

Figure 69 : Schéma des différents modes de visions. Illustration : auteur, d'après (Aukstakalnis, 2016). 90

Figure 70 : Les différentes activités visuelles mobilisent des zones variables du champ de vision. Illustration : auteur. 91

Figure 71 : Application de réalité augmentée spatiale *The revealing flashlight*. Illustration : (Ridel et al., 2014). 93

Figure 72 : Visualisation AR du modèle 3D du site archéologique de Chan Chan (Pérou). Illustration : (Pierdicca et al., 2015). 93

Figure 73 : Principe de l'application MARCH. Illustration : (Choudary et al., 2009). 93

Figure 74 : Application de réalité augmentée pour le site Cova dels Cavals. Illustration : (Blanco-Pons et al., 2019). 93

Figure 75 : Relevé par vidéoprojection. Illustration : (Pinçon et al., 2018). 94

Figure 76 : Vue de l'application MAAP pour Hololens. Illustration : (Barbier et al., 2017). 94

Figure 77 : Schéma du principe du recalage. Il s'agit d'établir une relation entre les poses successives de la caméra du dispositif et celles de la caméra virtuelle agissant sur le rendu de la scène 3D, afin de rendre les deux images superposables. Illustration : auteur. 95

Figure 78 : Transformations successives du modèle sténopé. Illustration : auteur, d'après (Orteu, 2008). 97

Figure 79 : Marqueurs installés dans le cabinet de travail de Charles V (château de Vincennes) pour une application de réalité augmentée, 2009.Installation : © CMN/CNRS/AGP/Axyz. Photographie : © Mr~Poussnik. 100

Figure 80 : Données produites par SLAM. En vert : poses successives des caméras. En gris : cartographie 3D de l'environnement. Illustration : (Engel et al., 2014). 102

Figure 81 : Diagramme résumant le fonctionnement d'une méthode Keyframe-based Visual Monocular SLAM. Illustration : auteur, d'après (Younes et al., 2017). 103

Figure 82 : Composants d'un système de réalité augmentée. Illustration : auteur. 105

Figure 83 : Schéma du relevé *classique* : un moment privilégié d'observation du réel, qui aboutit à une représentation graphique destinée à illustrer un propos scientifique. Les données résultantes présentent un faible potentiel d'interrogation (section 1.3). Illustration : auteur. 109

Figure 84 : Schéma du relevé numérique, centré sur des copies virtuelles de l'objet d'étude. Pour les acteurs, l'enjeu est de réussir à intégrer des connaissances aux représentations numériques. Les pratiques actuelles mènent à un effet de dispersion des données et connaissances (section 2.5). Illustration : auteur. 110

Figure 85 : Schéma de l'approche intégrée proposée par Aïoli : la donnée 3D est au cœur du système d'information et permet de fédérer les ressources liées à l'objet d'étude. Illustration : auteur. 111

Figure 86 : Schéma de notre approche : un système d'information dédié au relevé d'art pariétal centré sur l'objet réel, permettant de spatialiser des observations. Le nuage de points 3D permet de fédérer les ressources liées à l'objet d'étude, et joue le rôle d'intermédiaire entre ces ressources et l'objet réel. Illustration : auteur. 112

Figure 87 : Schéma de notre approche. Le nuage de points agit comme intermédiaire entre l'objet réel et les données spatialisées associées, elles-mêmes liées à des descripteurs géométriques et sémantiques. Illustration : auteur. 113

Figure 88 : Hiérarchie proposée par Aïoli. Illustration : auteur. 117

Figure 89 : Architecture applicative de la plateforme Aïoli. Illustration : auteur. 118

Figure 90 : Une image maîtresse et les images de caractérisation morphologique produites par reprojection 3D → 2D. 119

Figure 91 : Principe du processus de propagation d'annotations. Illustration : auteur, d'après (Manuel, 2016). 120

Figure 92 : Structure proposée pour le *viewer* hybride regroupant les ressources 2D et 3D au sein d'une même scène. Illustration : auteur. 122

Figure 93 : Une implémentation du *viewer* selon la structure annoncée. Le projet comprend un calque « figures » doté d'une annotation, visible en 3D sous forme de nuage de points, et en 2D sur chaque vue maîtresse. 122

Figure 94 : Gauche : Extrait d'un fichier XML de calibration. Droite : Modélisation de l'appareil photographique selon le modèle sténopé, avec les paramètres du fichier de calibration. 123

Figure 95 : Extrait d'un fichier XML d'orientation produit par le module TAPAS lors du calcul photogrammétrique. 124

Figure 96 : Orientation de la caméra selon les paramètres extrinsèques calculés par TAPAS. Illustration : auteur. 125

Figure 97 : Structure proposée de modèle de caméra à textures multiples permettant de visualiser les images maîtresses, les cartes de caractérisation morphologiques, et les annotations 2D. Illustration : auteur. 125

Figure 98 : Construction d'une annotation sous forme de *shape* à partir des coordonnées pixellaires des points de contours. La structure du document JSON permet de tenir compte d'éventuelles discontinuités dans les contours en traçant des sous-régions. 126

Figure 99 : Plaquage de la texture panoramique sur une sphère en tenant compte des limites du champ de balayage. 127

Figure 100 : Fichier texte issu de l'export des paramètres de position et d'orientation des stations depuis Faro Scene. 127

Figure 101 : Visualisation de données issues de lasergrammétrie (©CNP/AGP), comprenant le nuage de points total et deux stations. À gauche : image obtenue en se positionnant sur la première station, avec la sphère en opacité nulle. À droite : même positionnement et sphère à 50% d'opacité. La transparence montre que la texture panoramique de la sphère est positionnée de manière cohérente par rapport au nuage de points. Illustration : auteur. 128

Figure 102 : Notre approche permet de visualiser les données d'un projet dans un *viewer* unique (nuage de points SIFT, images maîtresses, et une annotation 2D et 3D). 129

Figure 103 : Schématisation du problème posé par le recalage de deux scènes présentées dans des repères arbitraires. Illustration : auteur. 131

Figure 104 : Processus proposé pour le calcul de la matrice de passage permettant le recalage AR d'un projet. Flèches pleines : déclenchement de processus, flèches en pointillés : transfert de données. Illustration : auteur. 133

Figure 105 : Contrainte épipolaire. Illustration : auteur, d'après (Larnaout, 2014). 135

Figure 106 : Processus proposé pour diminuer le temps dédié au recalage. Illustration : auteur. 137

Figure 107 : Modèle de caméra permettant de prendre en compte des images de caractérisation géométrique. Illustration : auteur. 141

Figure 108 : Orientation directe d'une image par rapport à un jeu d'images déjà orientées. Illustration : auteur. 141

Figure 109 : Orientation indirecte d'une image. L'image candidate est orientée de manière relative par rapport à une image intermédiaire. L'image intermédiaire est orientée de manière directe. Illustration : auteur. 142

Figure 110 : Principe de l'algorithme de *decorrelation stretching*. Illustration : auteur, d'après (Dangampola, 2015). 144

Figure 111 : *Decorrelation stretching* d'une carte de normales provenant d'un projet Aïoli. À gauche : image originale. Au centre : décorrélation dans l'espace colorimétrique LDS. À droite : décorrélation dans l'espace colorimétrique YYE. Illustration : auteur. 145

Figure 112 : Exemple d'extrait du fichier PLY d'un nuage de points doté d'attributs stockés comme des couleurs. 146

Figure 113 : Rendu du nuage de points avec une texture issue d'un traitement par *decorrelation stretching*. 147

Figure 114 : À gauche : coloration du nuage de points à partir d'un *shader* de dégradé d'élévation. À droite : Modification de l'axe du dégradé pour révéler les variations morphologiques de la paroi. ... 147

Figure 115 : Rendu composite du nuage de points du facsimilé, basé sur l'utilisation conjointe des couches RGB et du dégradé de profondeur. 148

Figure 116 : Détail du panneau des chevaux selon différents modes de rendu. À gauche : RGB. Au centre : *shader* EDL. À droite : *shader* EDL et lumière directionnelle rasante. Ce dernier mode de rendu révèle bien les traces de griffures. 148

Figure 117 : Schéma du traitement envisagé. Le traitement du flux vidéo pourrait faciliter la compréhension de l'environnement en AR, par exemple à l'aide d'algorithmes de détection de contours. Illustration : auteur. 149

Figure 118 : Exemples de traitements d'images modifiant la perception de l'objet étudié. À gauche : solarisation de l'image (inversion partielle des densités). Au centre : détection des contours par passage par zéro du Laplacien. À droite : seuillage. 149

Figure 119 : Utilisation de notre application de *decorrelation stretching* vidéo pour apprécier visuellement les variations de matériaux sur le facsimilé. Illustration : auteur. 150

Figure 120 : Processus proposé pour fluidifier le traitement vidéo en exploitant les images-clés. Illustration : auteur. 150

Figure 121 : Schéma du processus d'annotation 2D d'une image orientée par lancer de rayons, et la prévisualisation 3D du tracé subséquent (voir section suivante). Illustration : auteur. 152

Figure 122 : Processus proposé pour l'annotation d'une ressource 2D d'un projet. Illustration : auteur. 152

Figure 123 : Prévisualisation 3D d'une annotation 2D. La prévisualisation ne correspond pas tout à fait au tracé du fait du manque d'informations 3D dans la zone du cheval. Illustration : auteur. 153

Figure 124 : Processus proposé pour l'annotation d'un nuage de points 3D d'un projet. Illustration : auteur. 154

Figure 125 : À gauche : principe de l'annotation 2D basée sur la *frame* active du flux vidéo. À droite : exemple d'implémentation, basé sur le fonctionnement actuel d'Aïoli. Illustration : auteur. 156

Figure 126 : À gauche : principe de l'annotation dépendant des déplacements de l'utilisateur à l'aide d'une Spline. À droite : exemple d'implémentation, à partir des démonstrations de WebARonARKit. Illustration : auteur. 156

Figure 127 : À gauche : Principe de l'annotation par *hit-test,* dépôt d'une géométrie solide sur une ancre issue du suivi. À droite : Exemple d'implémentation, à partir des démonstrations de WebXR Viewer - polyfill. Illustration : auteur. 157

Figure 128 : Processus proposé pour l'annotation semi-automatique basée sur des algorithmes de segmentation. Illustration : auteur. 158

Figure 129 : Potentiel d'un outil de sélection par seuillage appliqué aux attributs colorimétriques ou aux normales. T correspond à la valeur de tolérance. Illustration : auteur, d'après (Manuel et al., 2018). 160

Figure 130 : Outil baguette magique 3D via la fonction de coloration du *shader* (les points sélectionnés sont colorés en rouge). À gauche : sur la texture RGB du nuage de points. À droite : sur la texture décorrélée. Illustration : auteur. 161

Figure 131 : Schéma du principe de l'outil *baguette magique* 3D. En haut : initialisation. En bas : actualisation de la sélection en cas de déplacement de la vue. Illustration : auteur. 162

Figure 132 : Adaptation de la couleur et de l'opacité d'une région 3D pour l'harmonisation des matériaux d'annotation 2D et 3D. La prise en charge des propriétés de contours et des textures de remplissage est plus complexe. 164

Figure 133 : Schéma du principe du processus de rendu automatique d'une légende graphique à partir des données de projet. Illustration : auteur. 165

Figure 134 : Comparaison du relevé multimodal et des différentes approches de relevé d'art pariétal. Illustration : auteur. 166

Figure 135 : Navigateurs prenant en charge le WebXR, de manière expérimentale. 167

Figure 136 : Représentation schématique du fonctionnement d'une application basée sur les API WebXR et WebGL Illustration : auteur. 168

Figure 137 : Composition de la page HTML selon le mode de visualisation adopté. Illustration : auteur. 169

Figure 138 : Fonctionnement global du *viewer*, unifiant les fonctionnements de Three.JS et Potree.JS pour rendre des ressources 3D diverses au sein d'une unique scène 3D. Illustration : auteur. 170

Figure 139 : Processus de rendu de la scène contenant les ressources d'un projet. Illustration : auteur. 171

Figure 140 : Répartition des points du nuage du facsimilé dans l'octree. 172

Figure 141 : Déroulement général du processus de conversion d'un nuage de points en octree avec PotreeConverter. Illustration : auteur. 173

Figure 142 : Affichage sélectif des nœuds de l'octree basé sur le cône de vision de la caméra. Les couleurs correspondent aux différents niveaux de détails liés à la proximité de la caméra. Illustration : auteur, d'après (Schuetz, 2016). 174

Figure 143 : Rendu des points sous forme de carrés ou de cercles par Potree.JS. 174

Figure 144 : Tous les attributs de points sont communiqués au *shader*, qui peut les traduire en informations colorimétriques pour le rendu. 175

Figure 145 : Ces attributs peuvent être facilement manipulés et exploités pour offrir de nouveaux modes de visualisations. 176

Figure 146 : Chargement du nuage de point d'un projet dans le *viewer*, en respectant la structure hiérarchique préétablie. 177

Figure 147 : Principe proposé pour l'utilisation d'une image RTI comme texture dynamique d'un plan 3D. Illustration : auteur. 178

Figure 148 : Capture d'écran d'une image RTI spatialisée et annotée dans Aïoli. Illustration : auteur. 178

Figure 149 : Résultats de l'implémentation de l'algorithme de *decorrelation stretching* en variant les paramètres. À gauche : image originale. Au centre : image décorrélée, avec m_{target}=120. À droite : image décorrélée avec m_{target}=180. 179

Figure 150 : Capture d'écran. Visualisation d'une ressource complémentaire non spatialisable : ancrage d'un DIV à un point 3D. 180

Figure 151 : Les différents repères impliqués dans le processus d'annotation d'une image. Illustration : auteur. 180

Figure 152 : Exports 2D issus d'une scène 3D. À gauche : projection perspective. Au centre : projection orthographique. À droite : extrait d'image panoramique équirectangulaire. 182

Figure 153 : Diagramme du processus de rendu permettant plusieurs modalités de visualisation (*desktop* ou XR). Illustration : auteur. 184

Figure 154 : Diagramme du processus de rendu du flux vidéo dans le canvas. Illustration : auteur. 185

Figure 155 : Diagramme de processus de sortie du mode XR. Illustration : auteur. 186

Figure 156 : Capture d'écran de l'application en mode XR. À ce stade, la scène apparaît de manière stable d'en l'environnement réel, mais n'est pas positionnée de manière cohérente. 186

Figure 157 : Diagramme du processus de recalage basé sur un calcul de résection spatiale. Illustration : auteur. 187

Figure 158 : Résultat. Capture d'écran d'un projet visualisé en réalité augmentée après un recalage réussi. 188

Figure 159 : Exemple de couple d'images extrait pour l'évaluation. À gauche : image de référence. À droite : image AR. 190

Figure 160 : Comparaison des deux images par SURF et FLANN. La distance entre chaque paire de points homologues donne une indication concernant l'écart du nuage de points à sa position réelle. 191

Figure 161 : Vues AR 0 à 5 extraites pour l'évaluation d'un recalage. 192

Table 1 : Données quantitatives obtenues à l'aide de notre protocole d'évaluation. 192

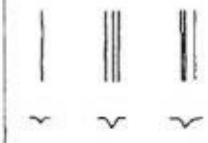
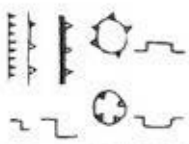
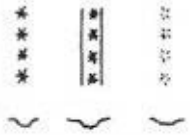
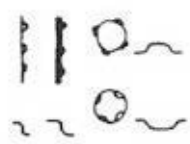
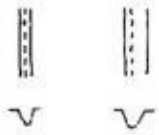
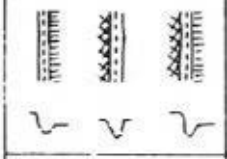
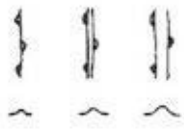
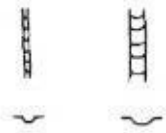
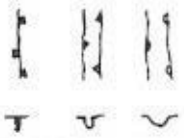
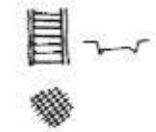
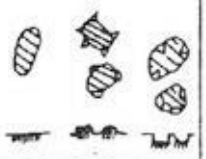
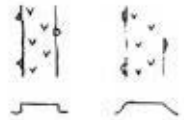
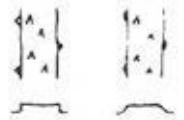
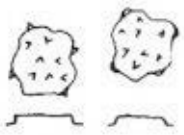
Figure 162 : Points homologues détectés sur les deux images du couple 3. Nous pouvons très nettement identifier un faux positif, caractérisé par le segment oblique vert. 192

Figure 163 : Couple d'images n°6. 193

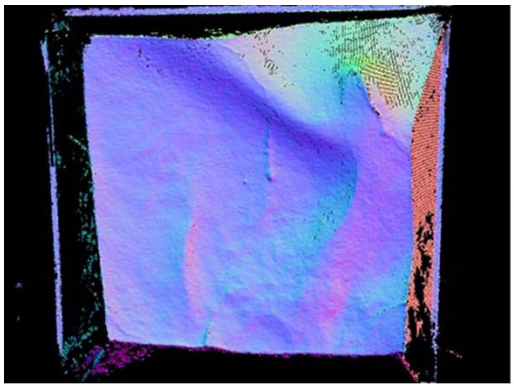
Figure 164 : Les modalités d'interaction en réalité augmentée ouvrent de nombreuses perspectives en termes de traçabilité. A. Les poses du dispositif d'affichage donnent des indications sur les zones d'intérêt de l'observateur. B. L'utilisation de la caméra frontale permet de déterminer le point d'intérêt de l'utilisateur sur l'écran. C. Un lancer de rayon basé sur les informations A. et B. permet de connaître précisément à chaque instant le point 3D observé. À droite : schéma d'une carte de chaleur synthétisant ces informations. Illustration : auteur. 204

Annexes

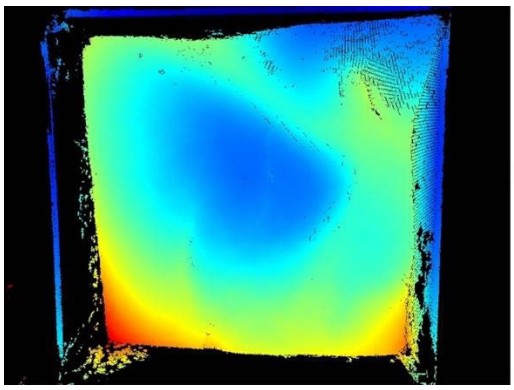
Annexe 1 : Conventions graphiques de relevé analytique proposées par (Delluc et al., 1986a)

	trait à section angulaire (fin, large, asymétrique)		bord vif mince ou épais; relief ou dépression à bord vif
	ligne de cupules; trait piqueté; cupules peu visibles		bord mousse mince ou épais; relief ou dépression à bord mousse
	trait à section recticurviligne (symétrique et asymétrique)		crêtes aiguës
	même trait à 1 bord abaissé; à 1 bord modelé; à 1 bord modelé et 1 bord abaissé		bourrelets mousses
	traits à section courbe		fissure étroite ou non; large sillon à bords mousses
	trait incisé - raclé; égrisage		enclave rocheuse de nature différente: plane; en relief; en dépression
	pigment noir; pigment rouge; pigment jaune		stalactite, coulée ou draperie
	trace historique ou moderne: linéaire; diffuse; ponctuelle		stalagmite
	trace animale: linéaire; diffuse		nappe de calcite

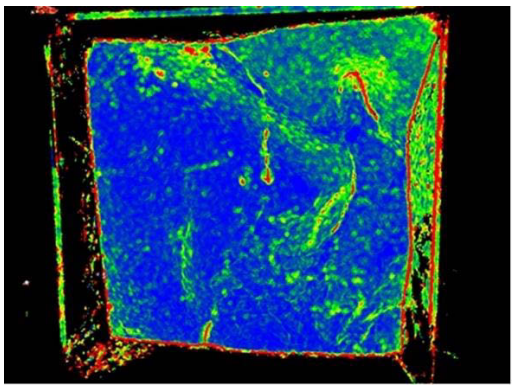
Annexe 2 : Descripteurs géométriques calculés par la plateforme Aïoli



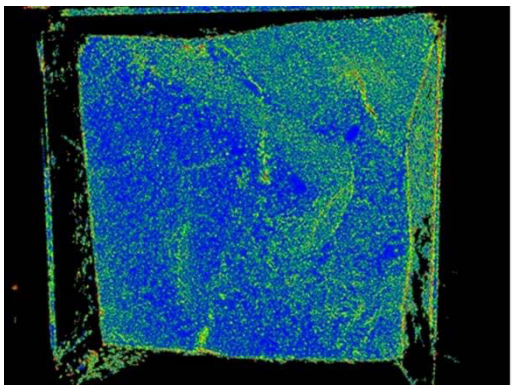
Cartes de normales : ce sont des images dans lesquelles chaque triplet de valeurs RGB fournit les composantes de la direction de la normale du texel correspondant. La méthode exploitée par Aïoli pour les produire consiste à estimer le meilleur plan tangent à la surface passant par les voisins du point étudié puis déduire les trois composantes du vecteur normal à ce plan.



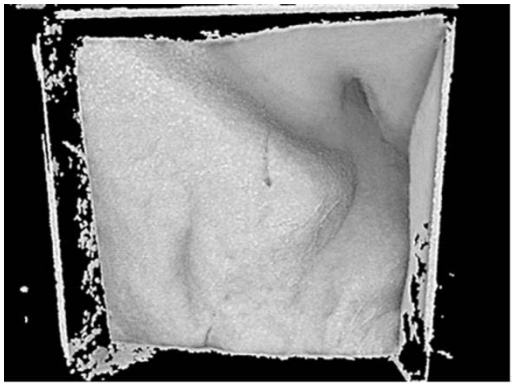
Cartes de profondeur : Pour créer des cartes de profondeur, Aïoli calcule, pour chaque point 3D visible sur l'image, la distance séparant ce point du centre optique de la caméra.



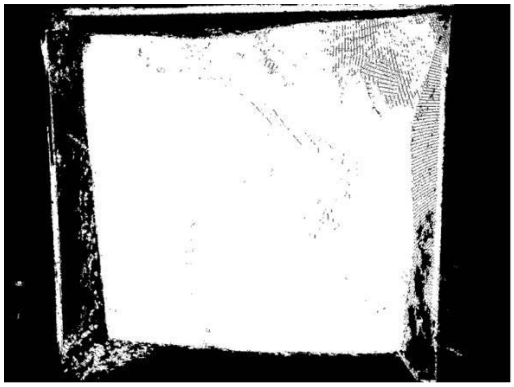
Cartes de courbure : La courbure est une mesure quantitative permettant de décrire la courbe d'un objet. Par exemple, la courbure d'une droite est nulle, et celle d'un cercle est positive et constante. Les valeurs de courbure peuvent être estimées à partir des données 3D à l'aide de méthodes gaussiennes, de méthodes moyennes, ou par les normales, puis être représentées sous forme de cartes à l'aide d'un dégradé.



Cartes de rugosité : La rugosité en un point d'un objet 3D est une valeur obtenue à partir de la mesure de la distance entre ce point et le meilleur plan passant par ses voisins. Les cartes de rugosité sont donc des images dans lesquelles l'intensité de chaque pixel indique la valeur de rugosité du texel correspondant. Ainsi, une surface parfaitement lisse produit une carte de rugosité monochrome.



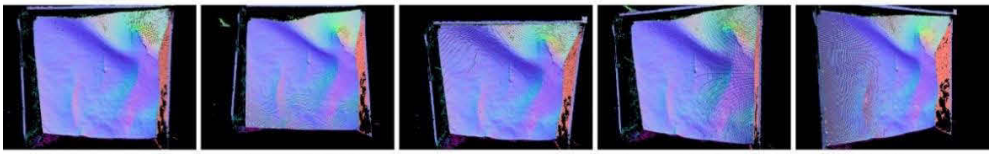
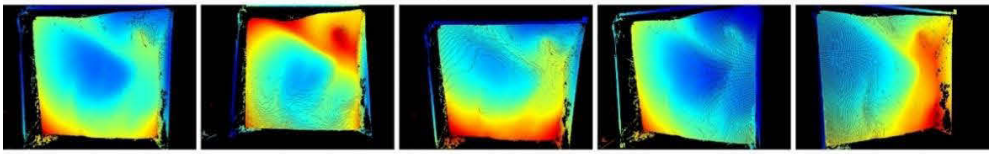
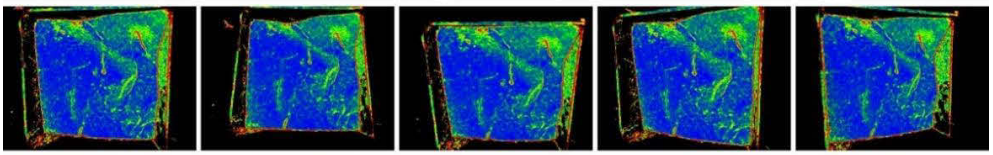
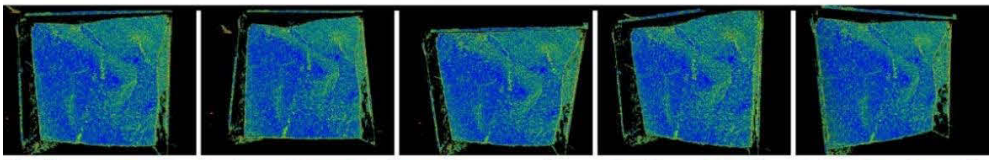
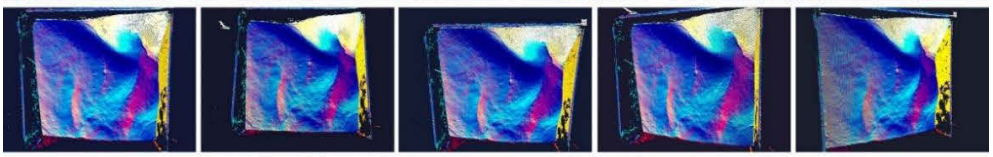
Cartes d'occlusion ambiante : L'occlusion ambiante est une méthode de rendu permettant de tenir compte des effets de la lumière ambiante pour contribuer au réalisme d'une scène 3D tout en simplifiant les calculs par rapport aux algorithmes d'illumination globale. Cette technique consiste à évaluer la quantité de lumière ambiante qui parvient à chaque point de la scène, afin de simuler l'assombrissement des zones les plus difficiles d'accès.



Cartes de positions des indices : Ces images qualifient la présence ou non d'un point 3D indexé. Elles se présentent donc comme des masques en noir et blanc. Les pixels noirs signifient que le nuage de points ne présente pas de point indexé sur l'axe de projection, les pixels blancs signifient qu'il en existe au moins un. Ces images ne sont pas exploitables pour faciliter l'observation de stigmates mais plutôt pour rendre compte de la composition du nuage de points.

Le processus de propagation d'annotation repose sur l'utilisation de points indexés, ou autrement dit, sur l'exhaustivité du nuage de points. Ainsi si un utilisateur souhaitait réaliser une annotation sur une zone noire, elle renverrait une région nulle. Ces images fournissent donc des indications sur les zones non couvertes du projet et incidemment sur la nécessité d'ajouter des photographies maîtresses supplémentaires pour pallier les manques.

Annexe 3 : Textures additionnelles liées aux images maîtresses

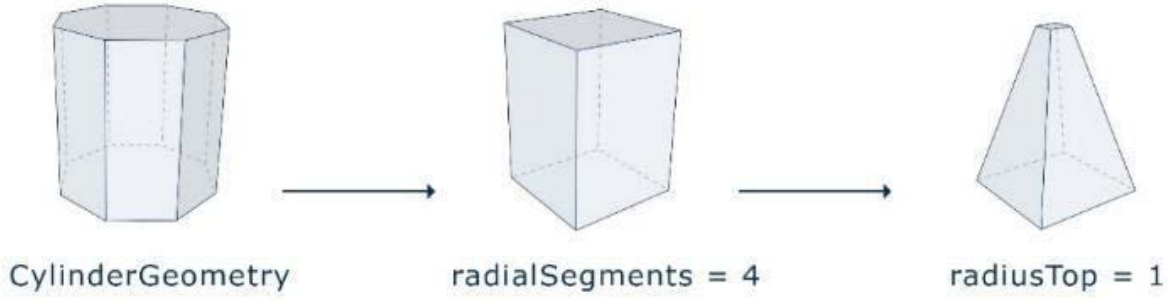
Images maîtresses	
Normales	
Positions des indices	
Profondeur	
Occlusion ambiante	
Courbure	
Rugosité	
Décorrélation des images maîtresses (LAB)	
Décorrélation des images maîtresses (YRD)	
Décorrélation des cartes de normales (LDS)	
Décorrélation des cartes de normales (YBK)	

Annexe 4 : Fonction d'initialisation du *viewer*

```
FONCTION initialisation

Entrée : la balise <div> de la page HTML dans laquelle doit être inséré le canvas de rendu

DÉBUT

    // création des éléments nécessaires au rendu WebGL
    créer canvas
    créer mainCamera
    créer renderer
    créer viewer
    créer lumiereAmbiante

    //affectation du canvas HTML (destination du rendu) au renderer
    insérer canvas dans <div>
    renderer.canvas ← canvas

    // création des éléments de la hiérarchie du viewer
    créer groupePointclouds
    créer calques3D
    ajouter calques3D à groupePointclouds

    créer groupeObjets
    créer groupeImages
    ajouter groupeImages à  groupeObjets

    créer groupeLumieres
    ajouter lumiereAmbiante à groupeLumieres

    // affectation des sous-scenes au viewer
    viewer.PointcloudScene ← groupePointclouds
    viewer.ObjectScene ← groupeObjets
    viewer.LightScene ← groupeLumieres

    RETOURNER 1

FIN
```

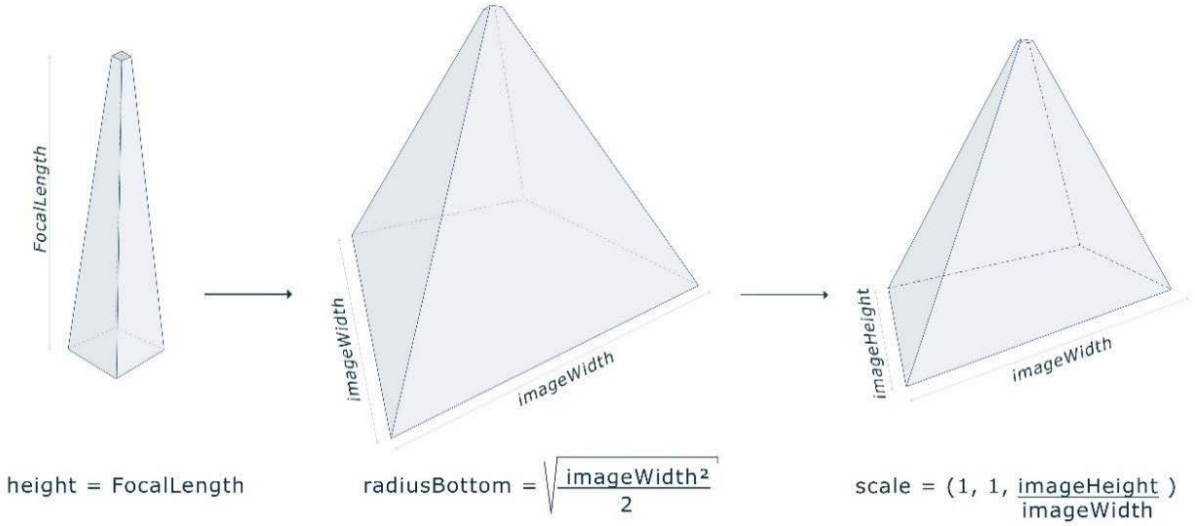
Annexe 5 : Création de pyramides de vue dans Three.JS selon les données de calibration provenant de MicMac

Pour modéliser les caméras, nous pouvons nous baser sur une géométrie solide de type pyramide. Dans Three.JS, les pyramides sont des cas particuliers des géométries de type cylindre, pour lesquels le rayon du sommet est minimal et le nombre d'arrêtes du périmètre de la base égale 4.



a. Génération d'une pyramide de vision correspondant au modèle de caméra.

La base de la pyramide correspond au plan image du modèle de caméra, et son sommet à son centre optique. La hauteur de la pyramide, déterminée par le paramètre `height` du constructeur `THREE.CylinderGeometry` doit donc être égale à la distance focale. Celle-ci est donnée par la balise `<F>` du fichier XML de calibration.



b. Paramétrage de la pyramide de vue selon les paramètres intrinsèques calculés par MicMac.

Les dimensions de la base doivent être égales aux dimensions `(imageWidth, imageHeight)` de l'image maîtresse. Or, dans Three.JS les dimensions de la base sont gérées par un unique paramètre nommé `radiusBottom`. Dans notre cas, la géométrie créée a une base carrée, `radiusBottom` correspond donc à la longueur de la demi-diagonale. Une approche possible pour conférer à la base les dimensions de l'image maîtresse est de calculer dans un premier temps le `radiusBottom` permettant d'obtenir une base carrée de côté `imageWidth`. Puis, d'appliquer dans un second temps un changement d'échelle non uniforme à la géométrie afin que deux côtés opposés soient égaux à `imageHeight`.

Annexe 6 : Stockage de la géométrie des annotations 2D issues de propagation


```
"regions":
{
    "5c88d838ae5e3":
    {
        "images":
        {
            "ITER0__0":
            {
                "subregions":
                {
                    "boundaries":[[950,2554],[954,2558],[954,2562],[954,2566],[954,2570],[9
                    "holes": []
                }
            },
            "ITER0__1":
            {
                "subregions":
                {
                    "boundaries": [[1226,2366],[1230,2370],[1230,2374],[1234,2378],[1234,23
                    "holes": []
                }
            },
            {
               ...
            }
        },
        "name": "Annot",
        "materialId": "5bfd5901d795d"
    },
```


Extrait d'un fichier JSON contenant les données issues d'une propagation d'annotation.

La propagation d'une annotation produit des listes de coordonnées 2D définissant, pour chaque image du projet, les contours de la géométrie correspondant à la région tracée en entrée. Ces informations sont stockées dans des fichiers JSON, chargés par le *viewer* lors du chargement d'annotations.

L'entrée `"images"` fournit la liste des différentes images sur lesquelles la région apparaît. Pour chacune, les données géométriques de l'annotation sont accessibles à partir des entrées `"subregions"`. Si la région est continue sur l'image, le document JSON ne comprend qu'une seule sous-région. Au contraire, en cas de discontinuité (par exemple liée à des phénomènes d'occlusion), elle en comprend de multiples. Chaque sous-région comprend deux entrées :

- une entrée `"boundaries"` correspondant au tableau des points de contours ordonnés, exprimés en coordonnées pixellaires dans le repère image ;

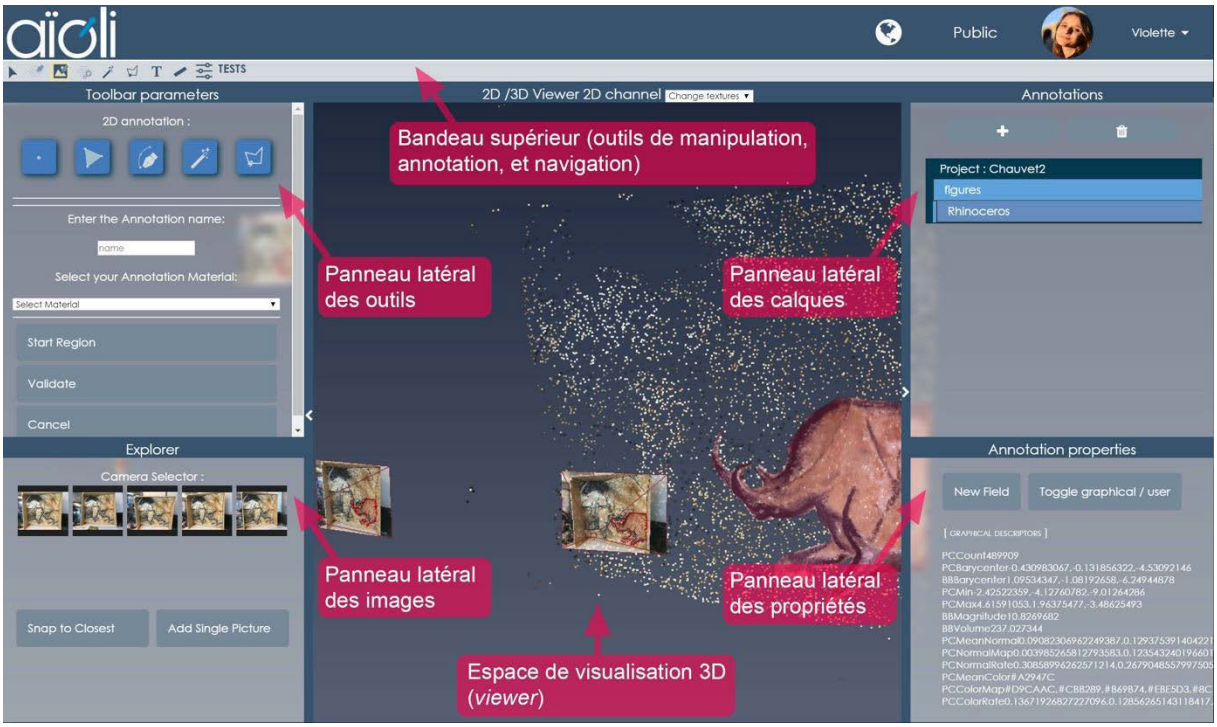
- une entrée `"holes"` correspondant aux points de contours intérieurs des trous éventuels de la région, exprimés en coordonnées pixellaires dans le repère image.

L'entrée `"materialId"` renvoie à l'identifiant de l'élément de la charte graphique à utiliser pour représenter la région conformément aux attentes de l'utilisateur.

Annexe 7 : Fonction de chargement d'une région 2D dans le *viewer*

```
FONCTION loadAnnotation (Charge une annotation 2D dans le viewer)

Entrée:     Json params, le document contenant les données issues de la
propagation de l'annotation ; String regionID, l'identifiant de la région

DÉBUT

SI regionId existe dans params.regions
   déclarer String region ← params.region[regionId];
   déclarer String materialID ← region.materialId
   déclarer String calqueID ← region.calqueId

   //pour chaque image sur laquelle la région est visible
   POUR CHAQUE imageID de region.images
       déclarer Object sousregions ← region.images[imageID].subregions
       déclarer Array contours ← sousregions.boundaries
       déclarer Array trous ← sousregions.holes

       récupérer dans la scène le Object3D planImage dont ID  = imageID
       récupérer  dans  planImage.children  le  Object3D  calque  dont  ID  =
       calqueID

       déclarer Array path
       déclarer Array pathHoles

       // pour chaque point de contrôle du contour
       POUR CHAQUE i de contours
         // ajouter les coordonnées x et y du point dans le path
         insérer dans path un Vector2(contours[i][0], contours[i][1])
       FIN_POUR

       // pour chaque point de contrôle du contour du trou
       POUR CHAQUE j de trous
         // ajouter les coordonnées x et y du point dans le tracé
         insérer dans pathHoles un Vector2(trous[j][0], trous[j][1])
       FIN_POUR

       Créer Shape géométrie à partir de path et pathHoles
       Récupérer matériau dans la base de données à partir de materialID

       SI matériau non défini
         créer matériau par défaut
       FIN_SI

       Créer Mesh annotation à partir de géométrie, matériau
       annotation.scale ← planImage.scaleRatio
       Insérer annotation dans calque

       RETOURNER VRAI
   FIN_POUR

SINON
   RETOURNER FAUX
FIN_SI

FIN
```

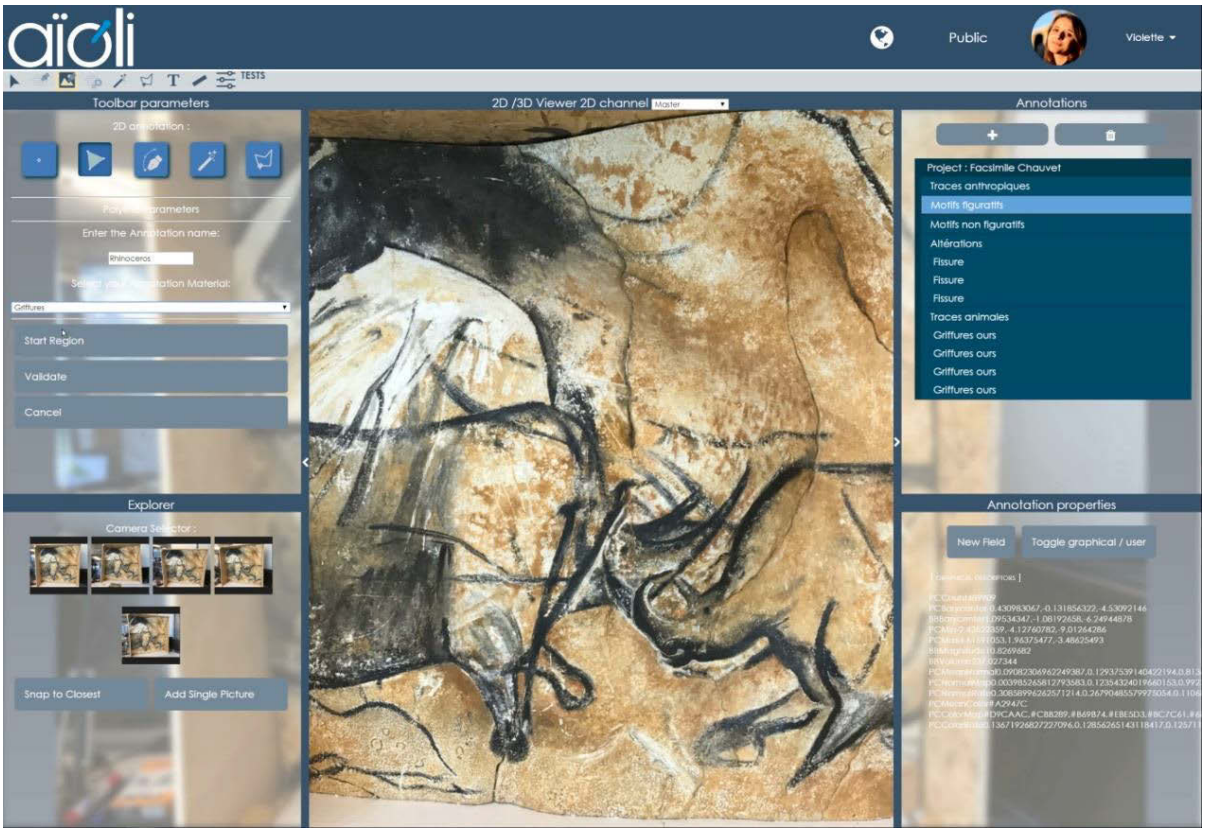
Annexe 8 : Interface de la preuve de concept



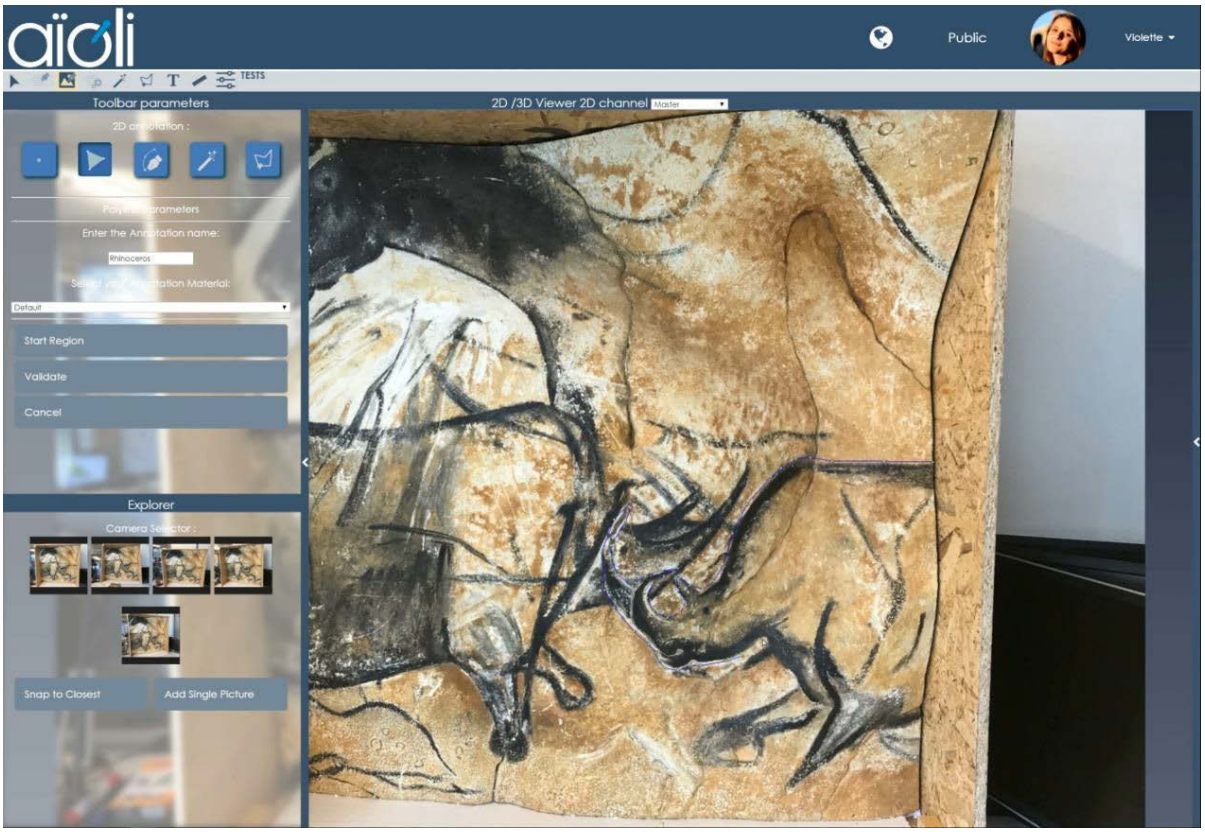
L'interface de la preuve de concept comprend six éléments distincts :

- **Le bandeau supérieur :** il comprend les différents onglets menant aux outils de manipulation, annotation, et navigation ;
- **Le panneau latéral des outils :** il comprend les éléments GUI permettant d'accéder aux différents outils et les paramétrer ;
- **Le panneau latéral des images : i**l comprend les vignettes de toutes les images maîtresses du projet et permet leur sélection ;
- **Le panneau latéral des calques :** il offre une représentation hiérarchique des calques et annotations existantes ;
- **Le panneau latéral des propriétés :** il présente les descripteurs sémantiques associés à l'annotation ou au calque actif ;
- **L'espace de visualisation :** il s'agit du *viewer* hybride 2D/3D.

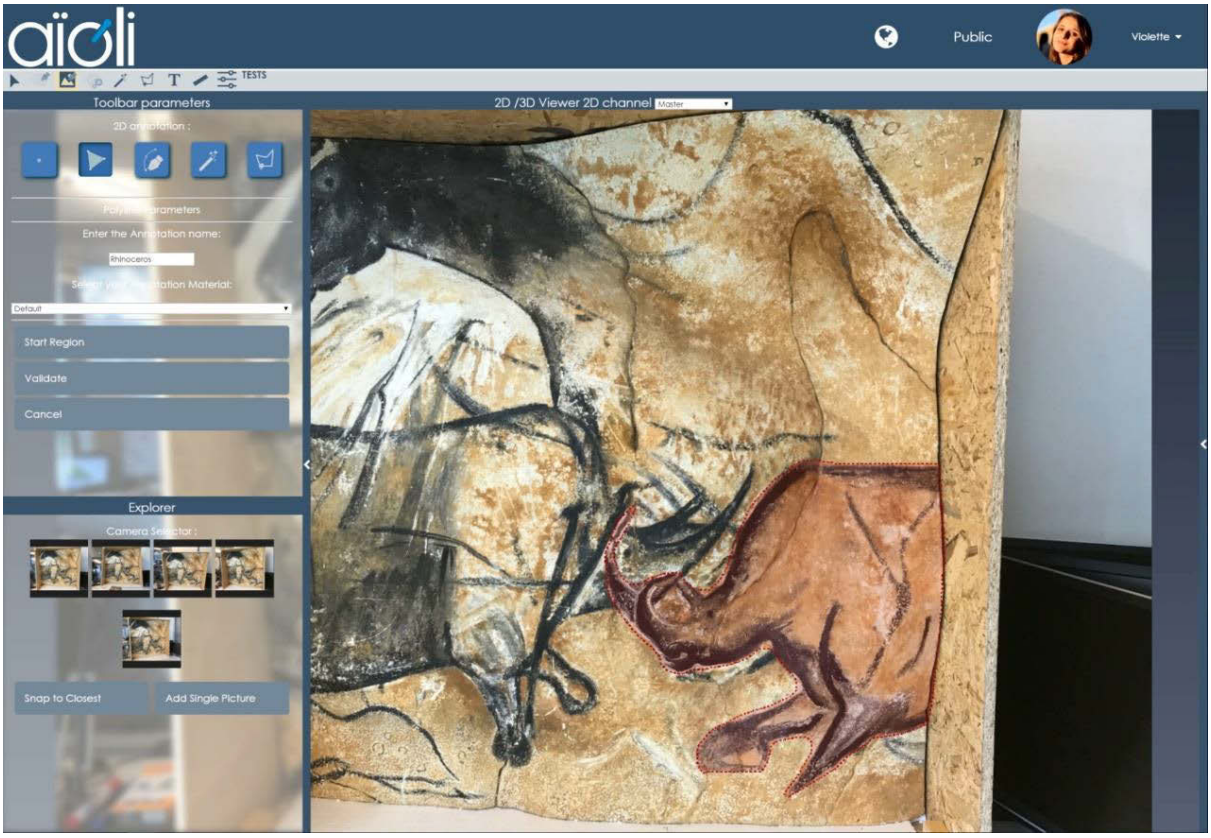
Annexe 9 : Captures d'écran de la preuve de concept (viewer hybride 2D/3D)



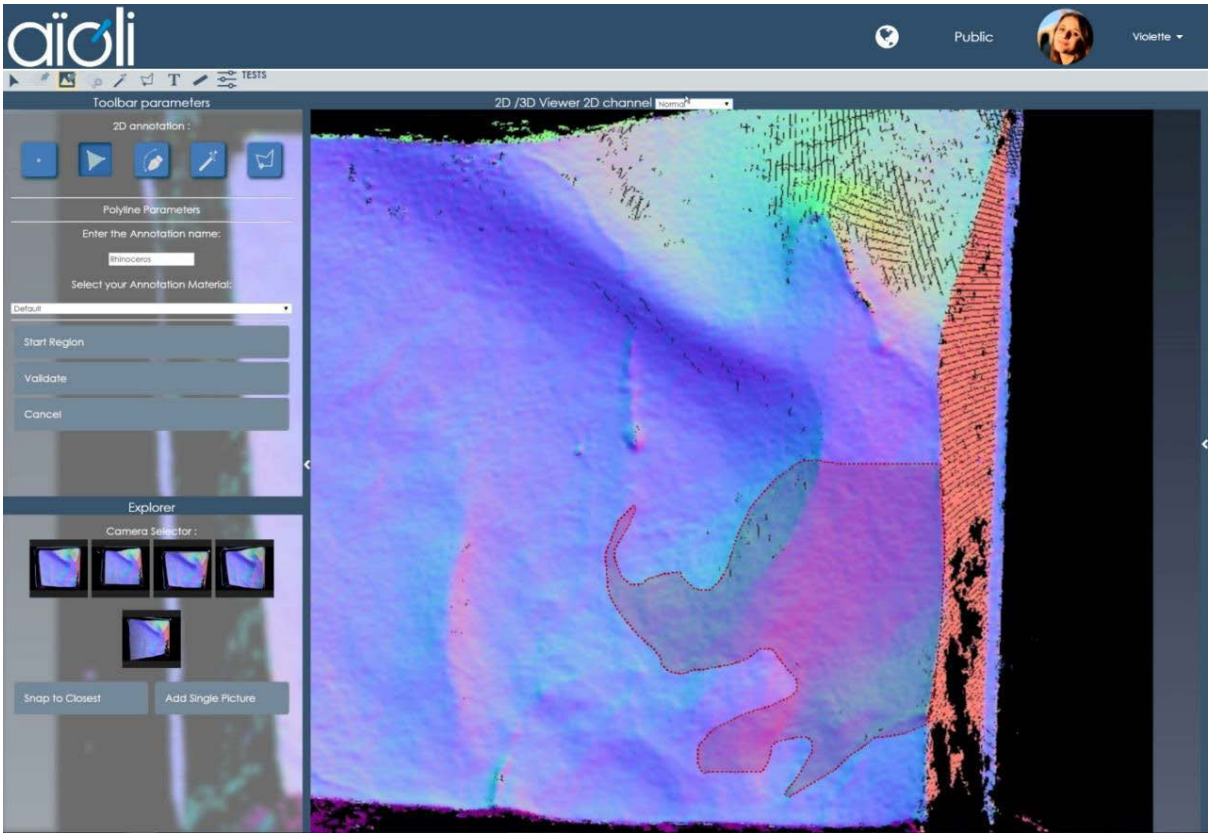
a. La caméra virtuelle est placée sur une image maîtresse. Préparation de l'annotation de cette image (choix du calque, de l'outil de dessin, et du matériau assigné à la future région).



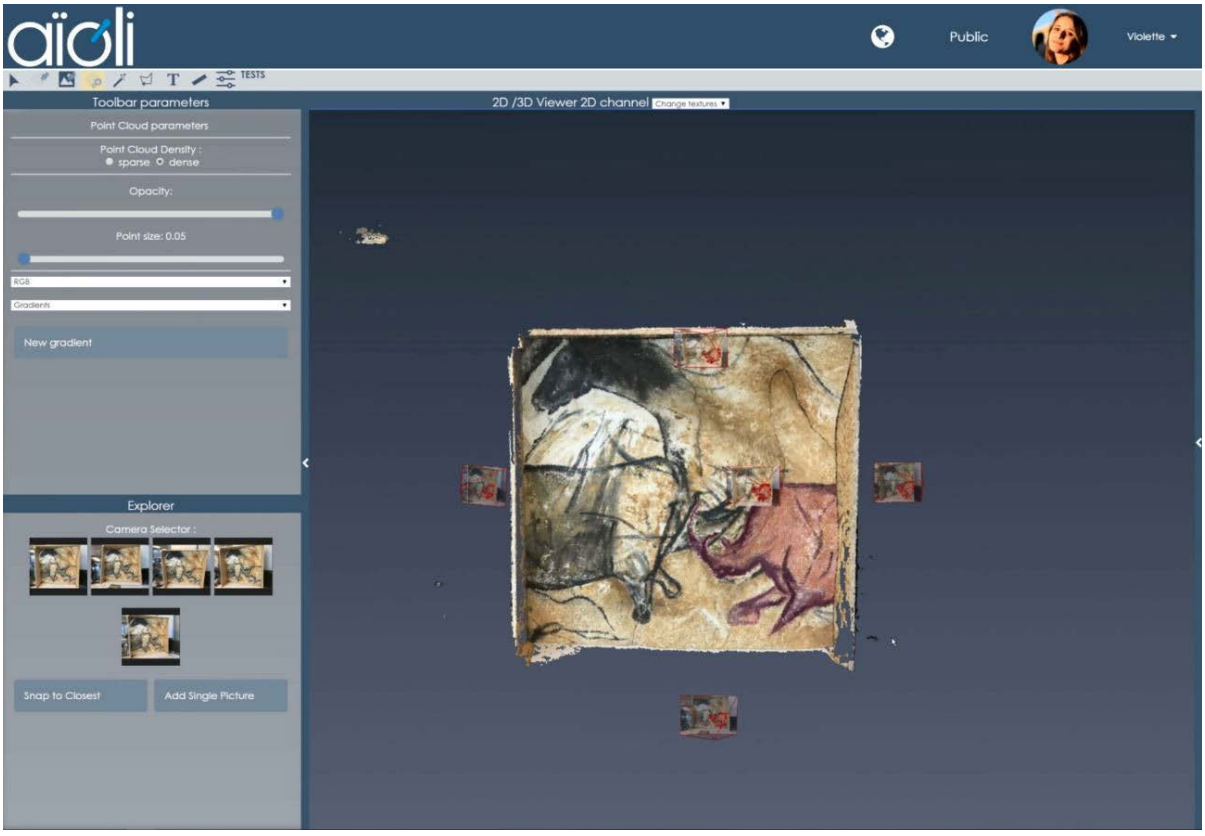
b. Tracé de l'annotation sur l'image à l'aide de l'outil polyligne.



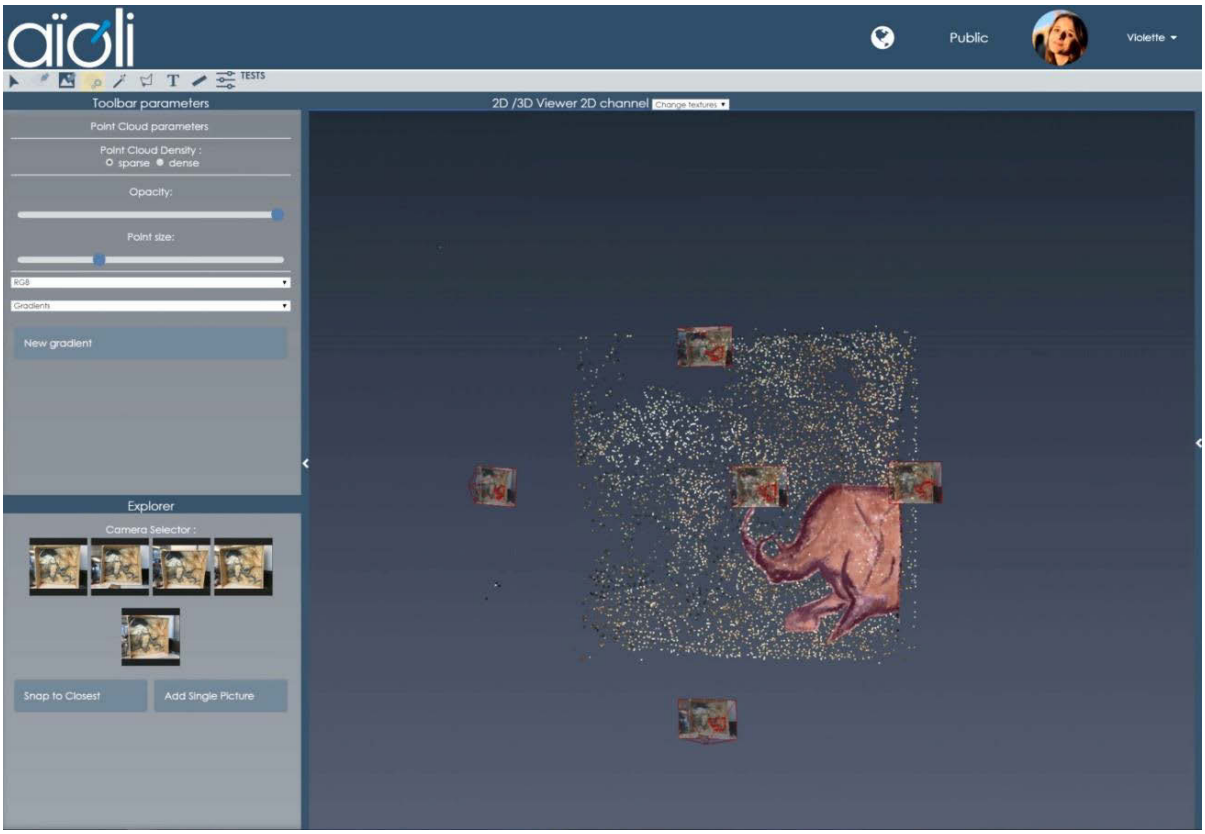
c. Visualisation de la région résultante à l'issue de la propagation.



d. Visualisation de la région résultante sur une image complémentaire (carte de normales).

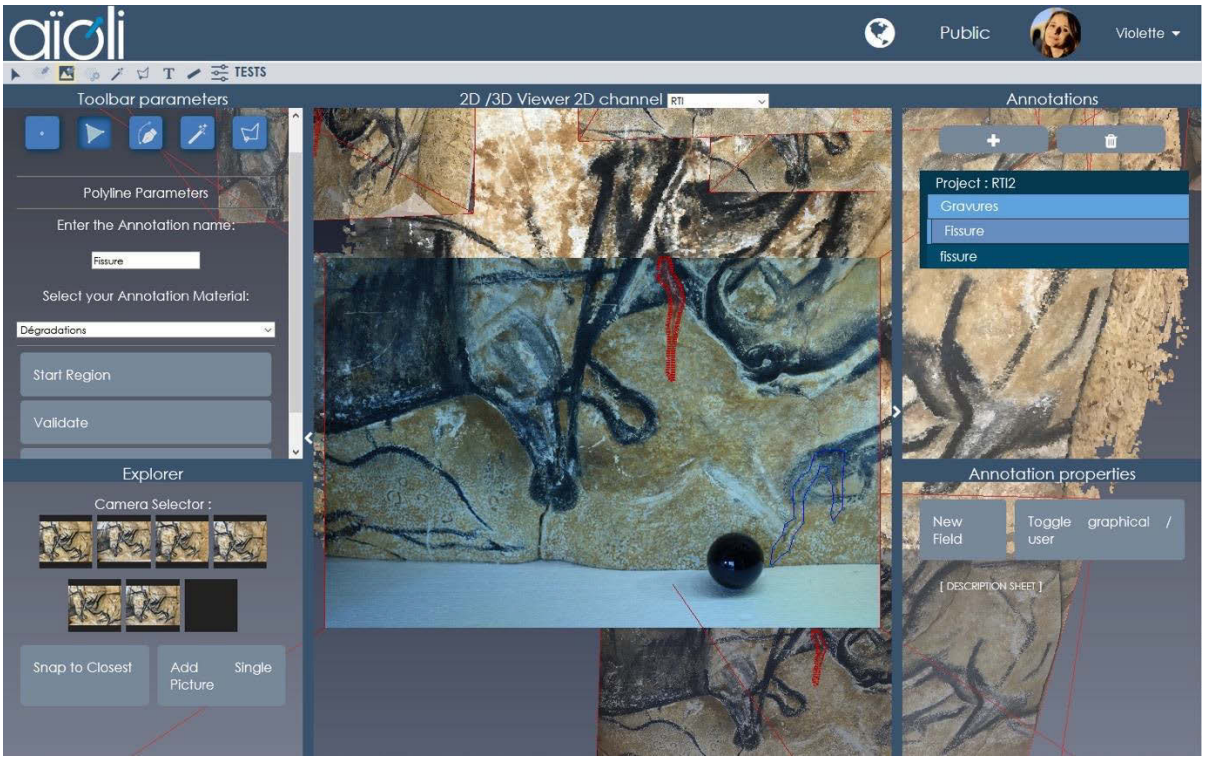


e. Visualisation de la région résultante en 3D, avec le nuage dense du projet et les images maîtresses spatialisées.

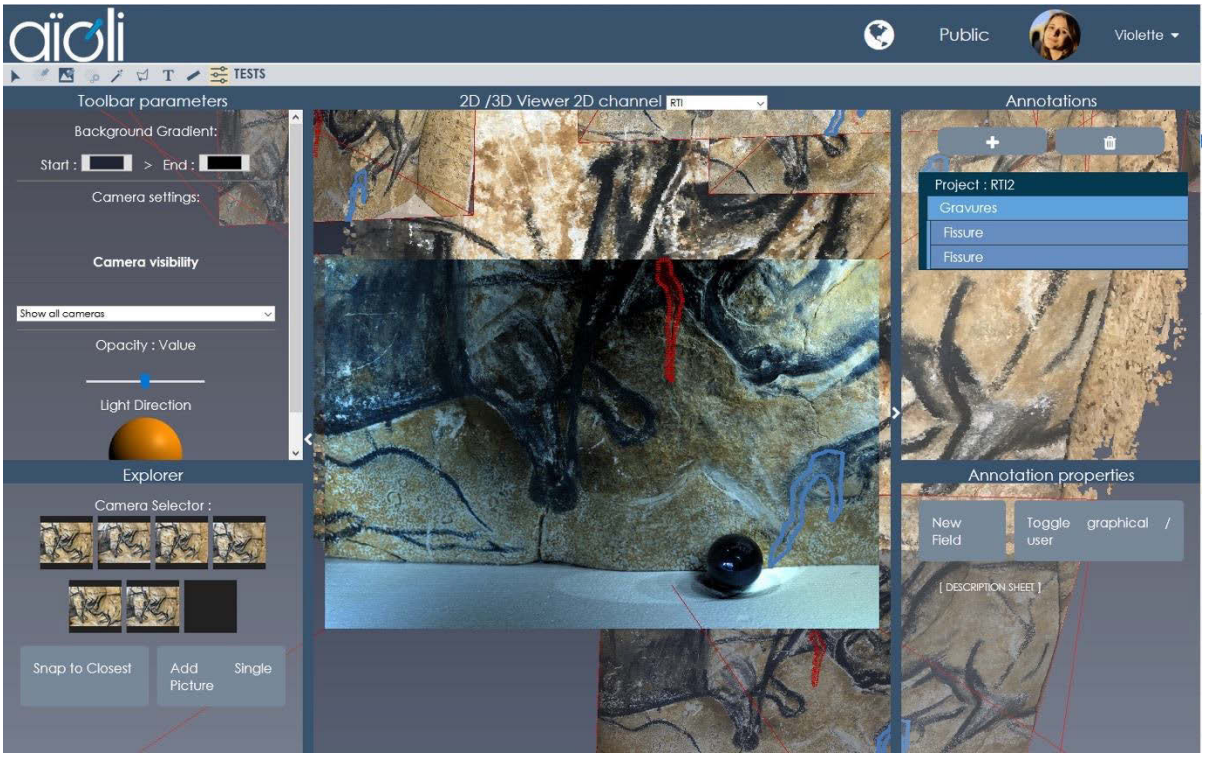


f. Visualisation de la région résultante en 3D, avec le nuage épars du projet et les images maîtresses spatialisées.

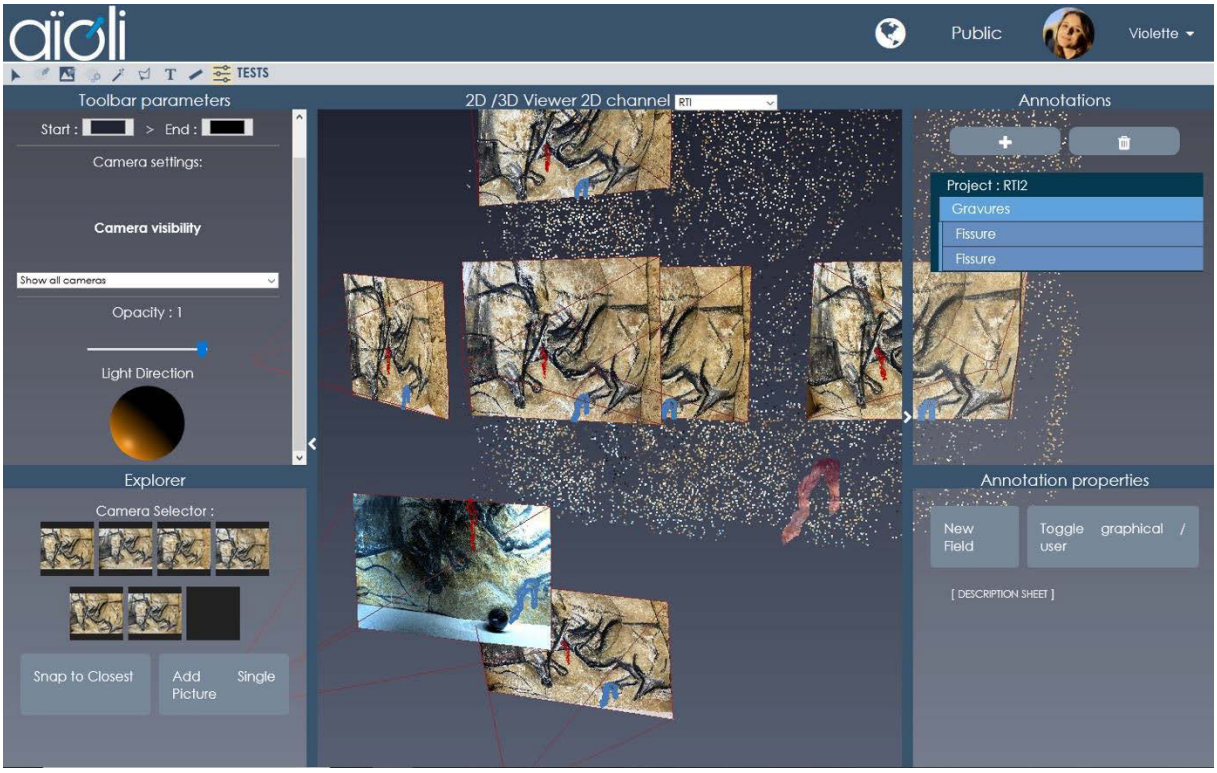
Annexe 10 : spatialisation et annotation d'une image RTI



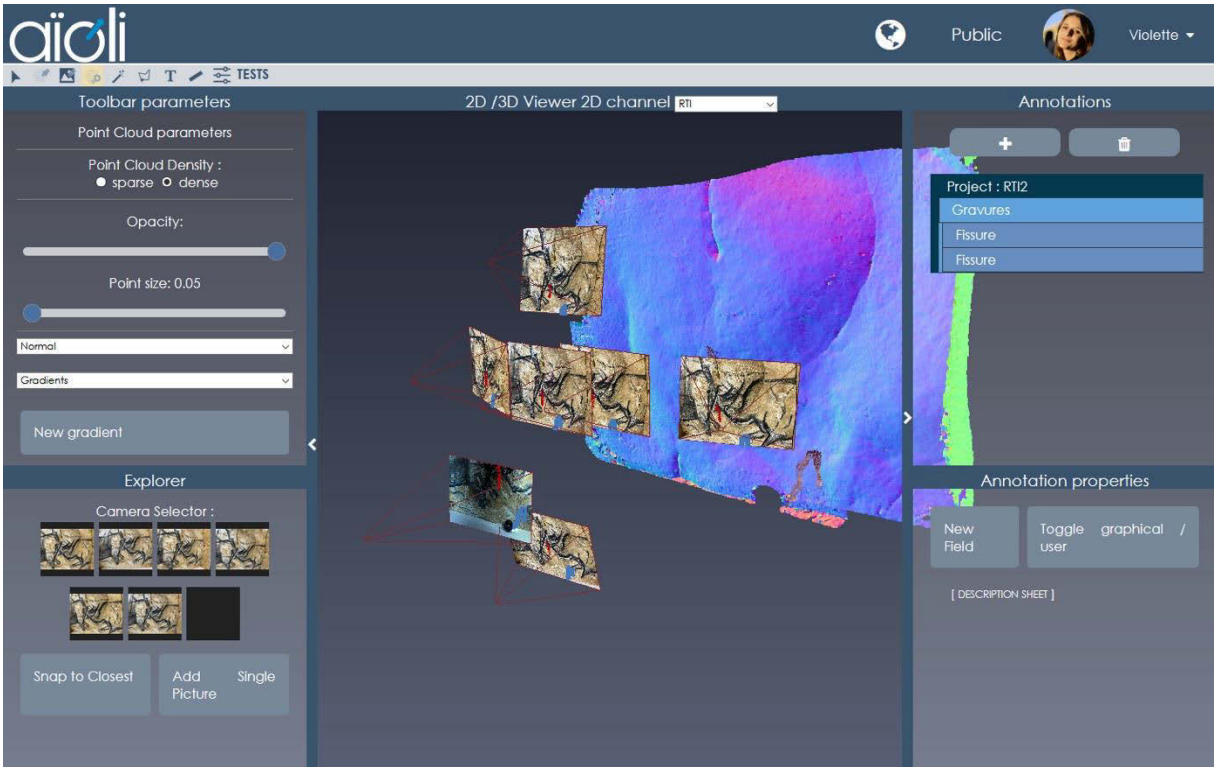
a. Capture d'écran de la preuve de concept lors de l'annotation d'une image RTI.



b. Capture d'écran de la preuve de concept après la propagation d'une annotation réalisée sur une image RTI.

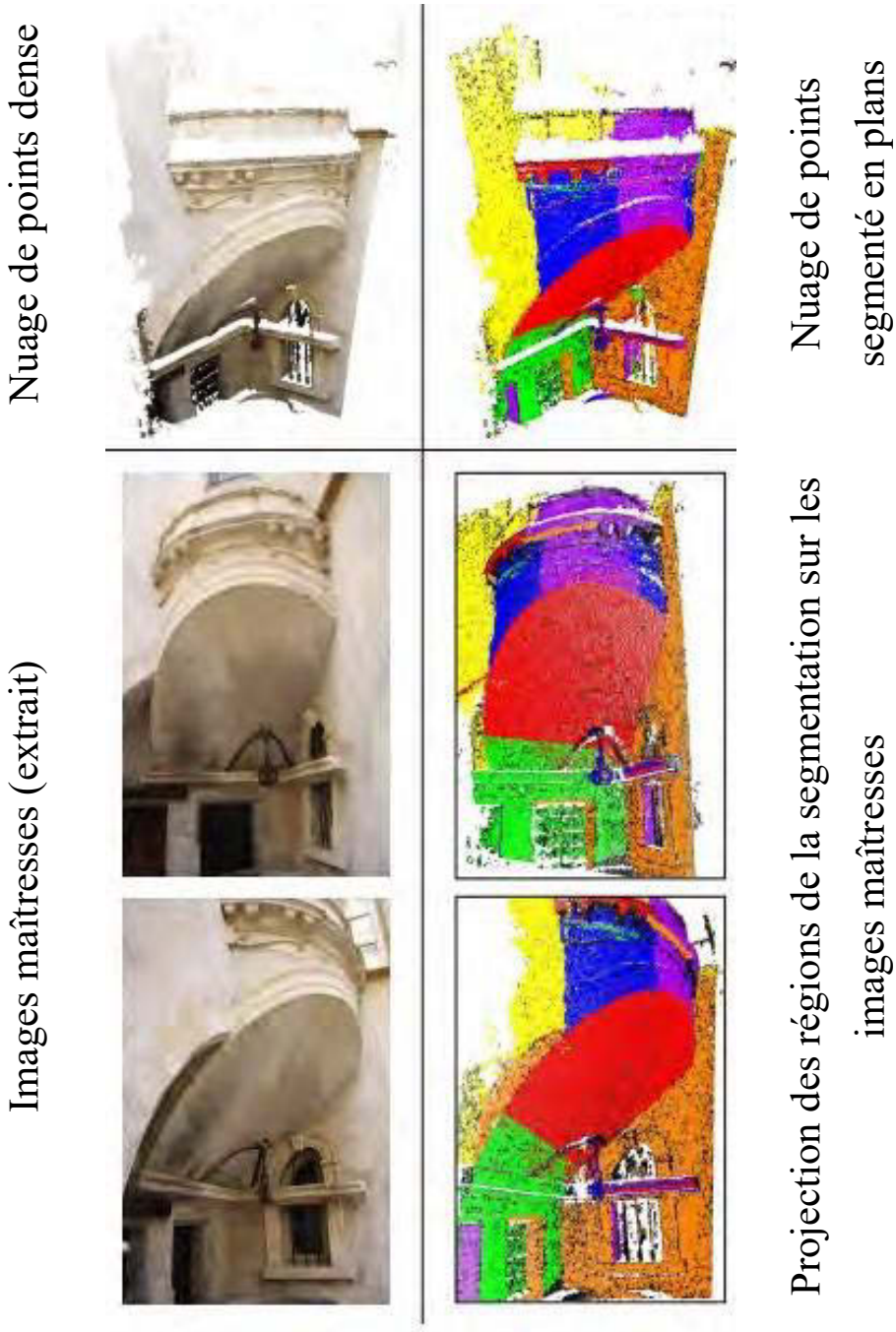


c. Capture d'écran de la preuve de concept. Image RTI orientée au sein d'un projet Aïoli. Les annotations réalisées sur l'image RTI ont été propagées aux autres ressources 2D et 3D du projet.



d. Capture d'écran de la preuve de concept. Image RTI orientée au sein d'un projet Aïoli, nuage de points du projet texturé par normales.

Annexe 11 : Segmentation en plans du nuage de points d'un projet (trompe Philibert de l'Orme, Lyon) et projection des régions résultantes sur les images maîtresses



À partir des indices des points 3D, chacune de ces régions peut directement être propagée sous forme d'annotation à l'ensemble des images du projet.

Annexe 12 : Sélection « baguette magique » de points 3D via la fonction de coloration du shader

```
FONCTION magicwand (sélection du point selon sa couleur)

Entrée : Vecteur d'entiers inputcolor contenant les attributs RGB du point,
Entier id correspondant à l'index du point testé,
Vecteur d'entiers ref contentant les attributs RGB du point de référence,
Réel T indiquant la tolérance de l'outil de sélection,
Tableau d'entiers selected contenant les indices des points sélectionnés

DÉBUT

      Déclarer un vecteur d'entiers outputColor

      Déclarer des réels Rmin, Rmax, Gmin, Gmax, Bmin, Bmax

      // calcul des valeurs R G et B minimales acceptables

      SI ref.r - T < 0 ALORS Rmin ← 0 SINON Rmin ← ref.r - T FIN_SI

      SI ref.g - T < 0 ALORS Gmin ← 0 SINON Gmin ← ref.g - T FIN_SI

      SI ref.b - T < 0 ALORS Bmin ← 0 SINON Bmin ← ref.b - T FIN_SI


      // calcul des valeurs R G et B maximales acceptables

      SI ref.r - T > 255 ALORS Rmax ← 255 SINON Rmax ← ref.r - T FIN_SI

      SI ref.g - T < 255 ALORS Gmax ← 255 SINON Gmax ← ref.g - T FIN_SI

      SI ref.b - T < 255 ALORS Bmax ← 255 SINON Bmax ← ref.b - T FIN_SI


      // vérifier si le point actif est éligible

      SI Rmin < inputcolor.r < Rmax ET Gmin < inputcolor.g < Gmax ET Bmin <
inputcolor.b < Bmax

            Insérer id dans selected

            // colorer le point sélectionné en rouge
            outputColor ← (255,0,0)

      SINON

            // Conserver la couleur originale du point
            outputColor ← inputcolor

      FIN_SI


      RETOURNER outputColor

FIN
```

Annexe 13 : Sélection « baguette magique » de points 3D via les rendus 2D successifs

FONCTION InitMagicWand (initialisation de l'outil de sélection de points 3D selon leur couleur)

Entrée : Vecteur 2 input : coordonnées 2D du clic de l'utilisateur,

Réel T indiquant la tolérance de l'outil de sélection

DÉBUT

déclarer vecteur 3 d'entiers refColor

déclarer un entier refIndex

Lancer un rayon de la caméra vers input

SI le rayon intersecte un point 3D Hit

RefColor ← Hit.color

RefIndex ← Hit.index

Sélection par baguette magique 2D sur le canvas de rendu à partir du point input

Affichage du contour des points sélectionnés à l'écran

Lancer l'écoute des évènements *change* dans la scène 3D

SINON

Retourner faux

FIN_SI

FIN

FONCTION onChange (fonction déclenchée en cas de changement dans la scène 3D après initialisation de la baguette magique)

Entrée : Vecteur 3 d'entiers refColor : couleur du point 3D de référence,
Entier RefIndex indiquant l'index du point 3D de référence,
Tableau de vecteurs d'entiers selection contenant les points du contour de la sélection en cours

DÉBUT

Récupérer la position de la caméra

Supprimer l'affichage du contour des points sélectionnés

Vider selection

Récupérer le point Ref via RefIndex

Lancer un rayon de Ref vers la caméra

Déclarer vecteur 2 newInput ← résultat du lancer de rayon (coordonnées écran du point 3D de référence)

Sélection par baguette magique 2D sur le canvas de rendu à partir du point newInput

Affichage du contour des points sélectionnés à l'écran

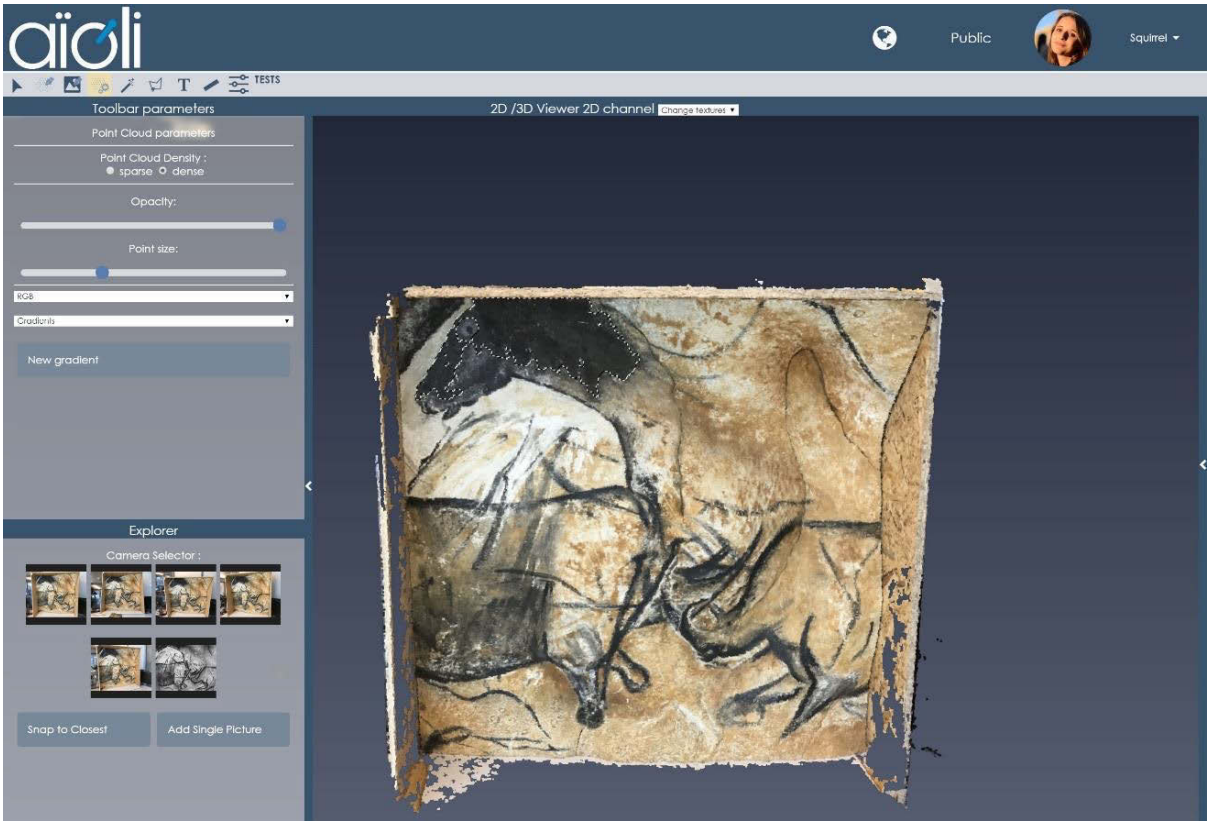
FIN

FONCTION validateMagicWand (validation de la sélection réalisée par baguette magique)

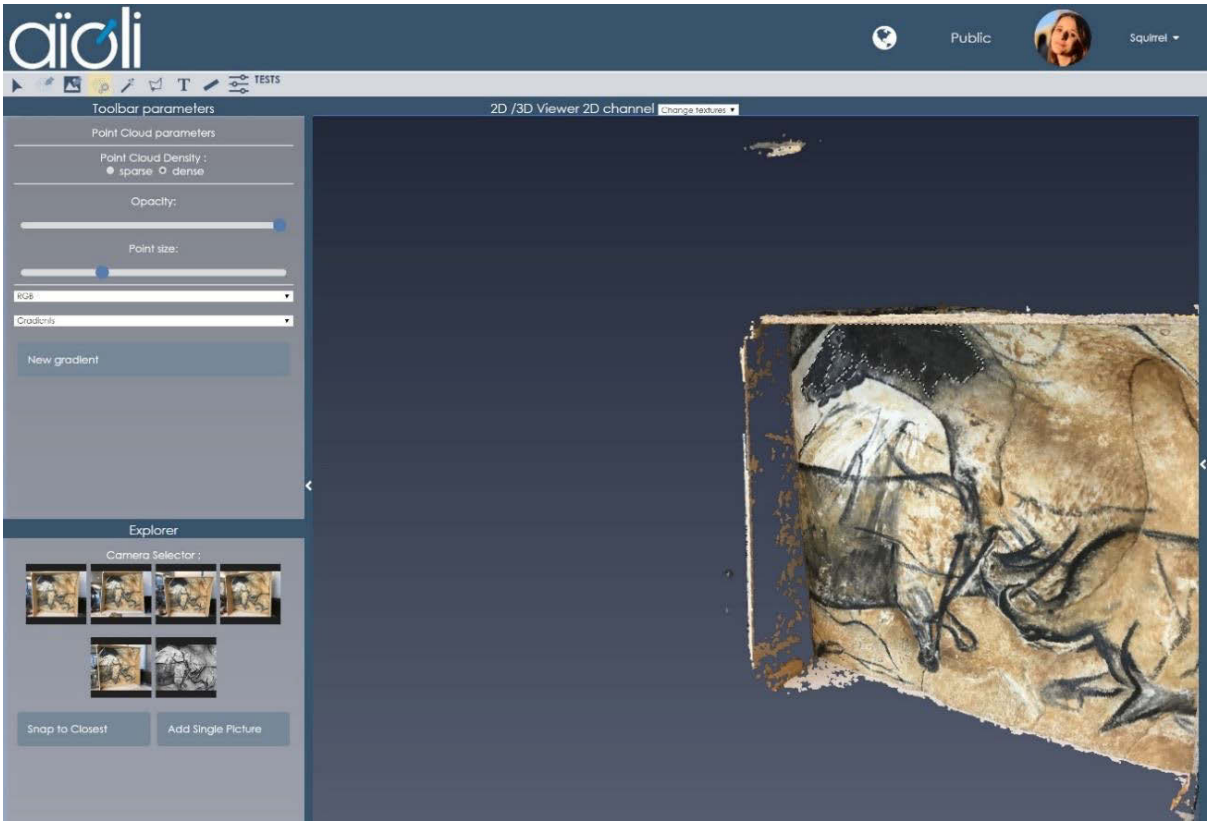
Entrée : Tableau de vecteurs d'entiers selection contenant les points du contour de la sélection en cours

```
DÉBUT
    Déclarer vecteur 3 previousHit
    Déclarer tableau de vecteurs 3 finalSelection
    Arrêter d'écouter les évènements change dans la scène 3D
    POUR CHAQUE point DANS selection
        Lancer un rayon du point vers la caméra
        SI le rayon intersecte un point 3D Hit
            Insérer un sprite dans la scène 3D à Hit.position
            SI previousHit
                Insérer un segment dans la scène 3D entre Hit.position et
                previousHit
            FIN_SI
            previousHit ← Hit.position
            Insérer Hit.position dans finalSelection
        FIN_SI
    FIN_POUR
    Supprimer l'affichage du contour des points sélectionnés
    Lancer la propagation de la région définie par finalSelection
FIN
```

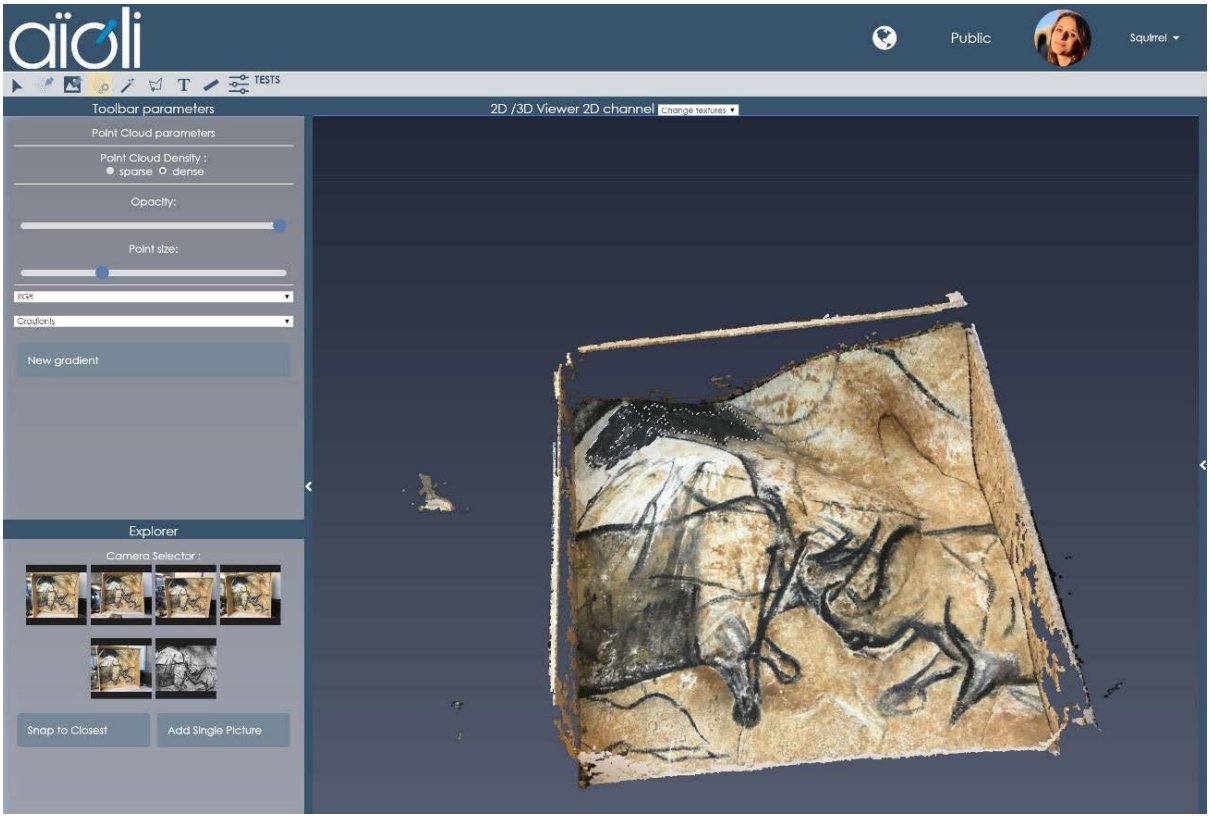
Annexe 14 : Captures d'écran de la fonction de sélection « baguette magique » via les rendus 2D successifs



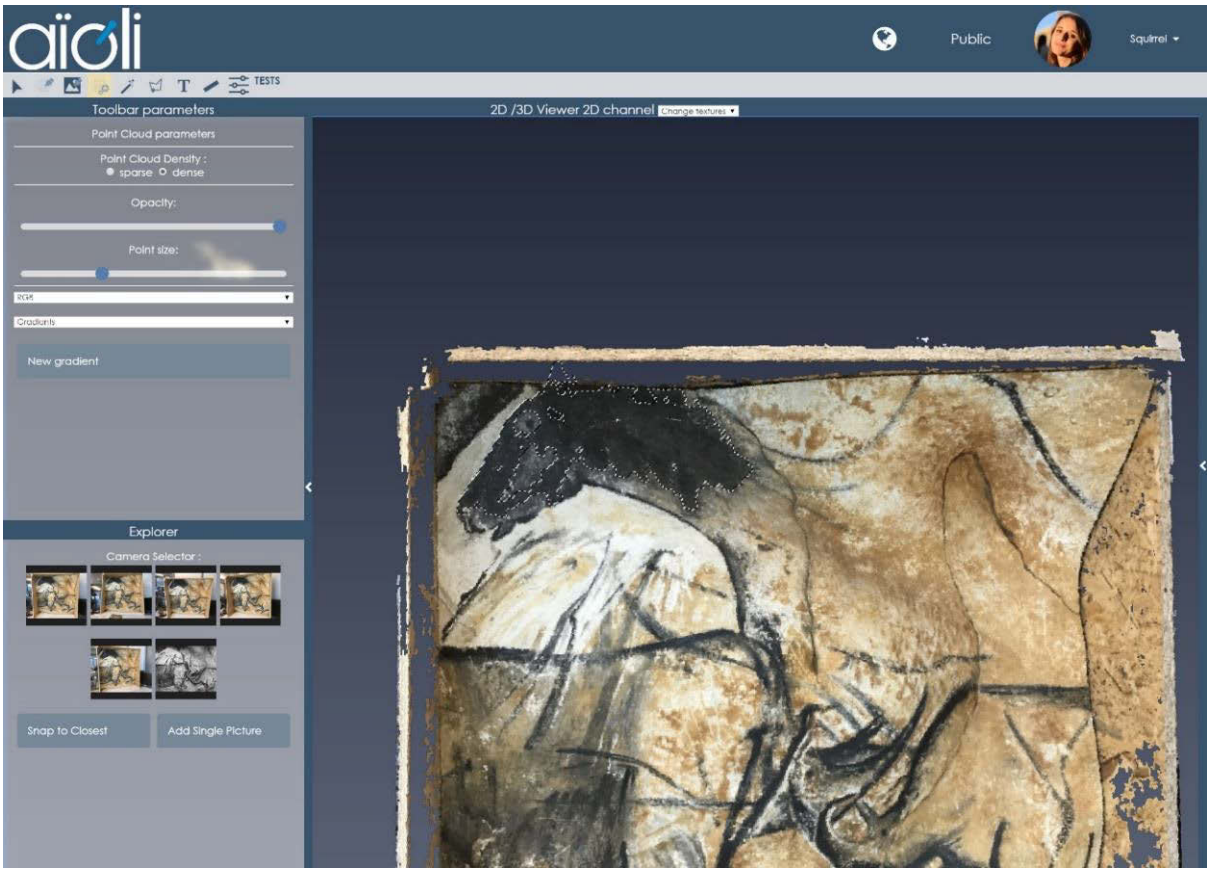
a. Sélection d'une région en 3D.



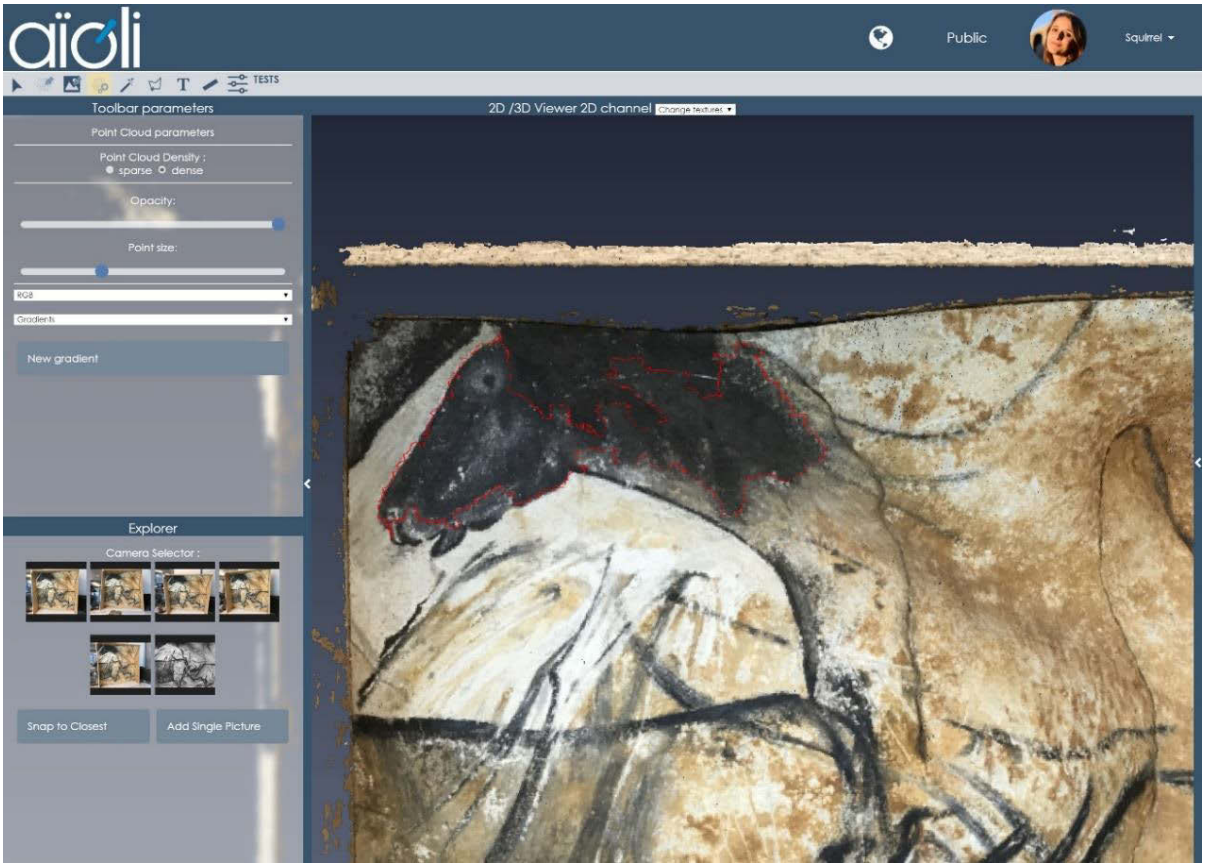
b. Actualisation de la sélection après un déplacement dans la scène 3D.



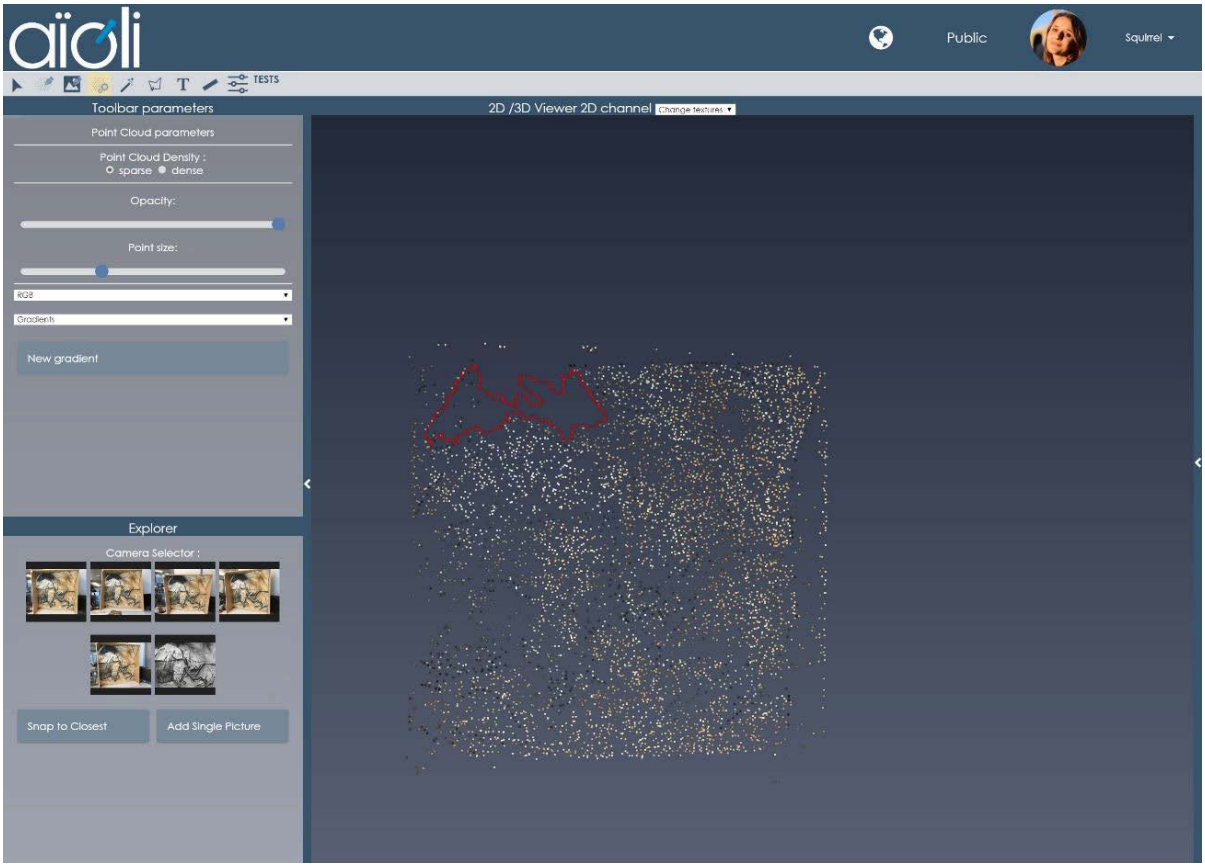
c. Actualisation de la sélection après un déplacement dans la scène 3D.



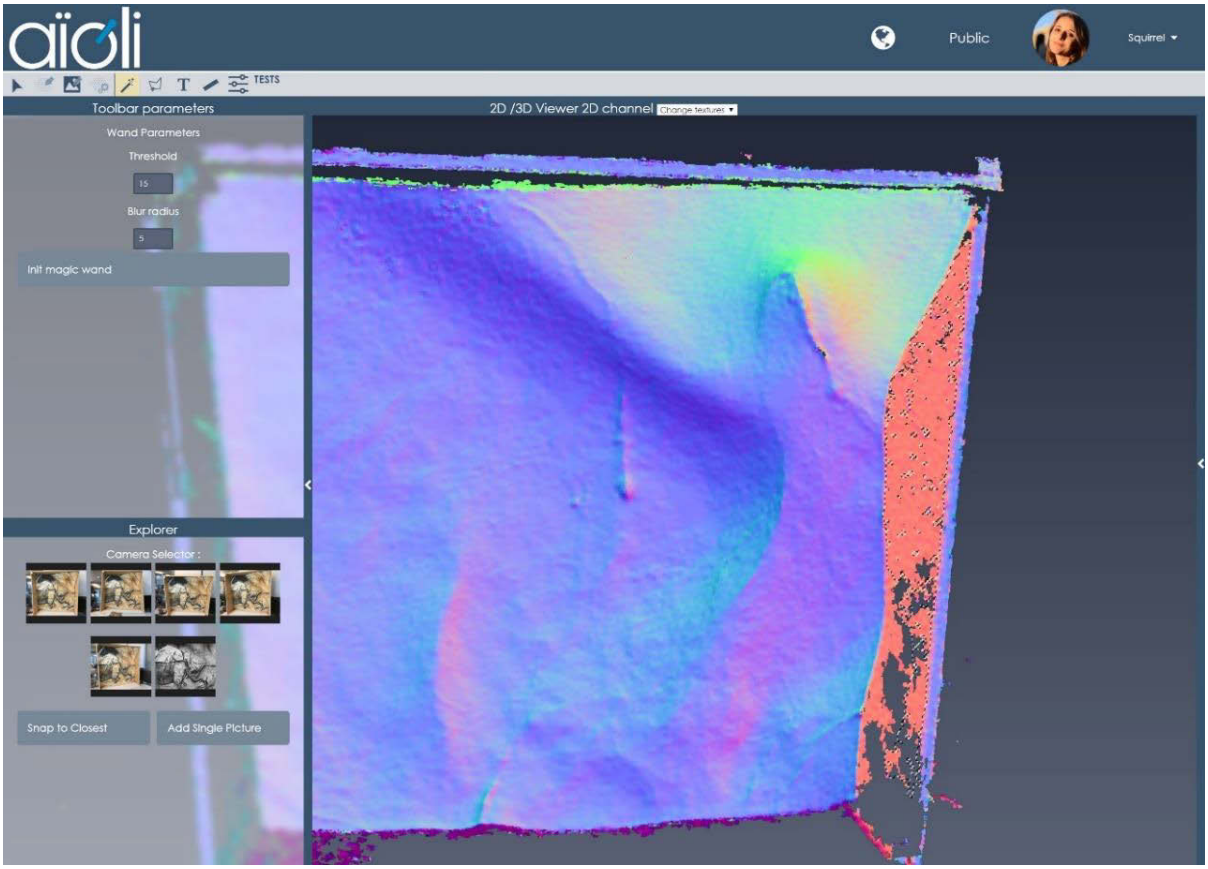
d. Actualisation de la sélection après un déplacement dans la scène 3D. Nous pouvons constater des erreurs : la région ne correspond pas tout à fait au tracé initial et comprend des points qui n'existent pas en 3D.



e. Validation de la sélection. La prévisualisation 2D est remplacée par les segments 3D rouges insérés dans la scène 3D, délimitant le contour de la région finale.

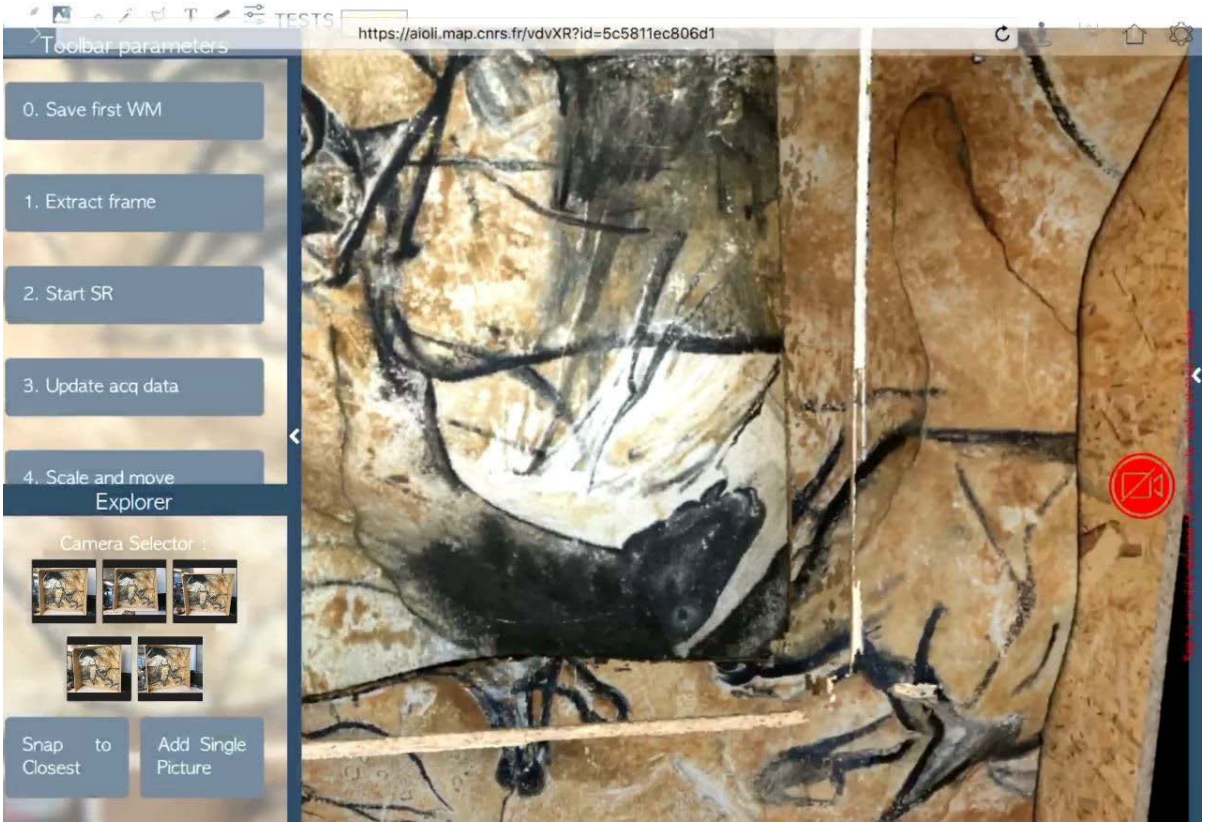


f. Prévisualisation 3D de la région sélectionnée avant propagation.

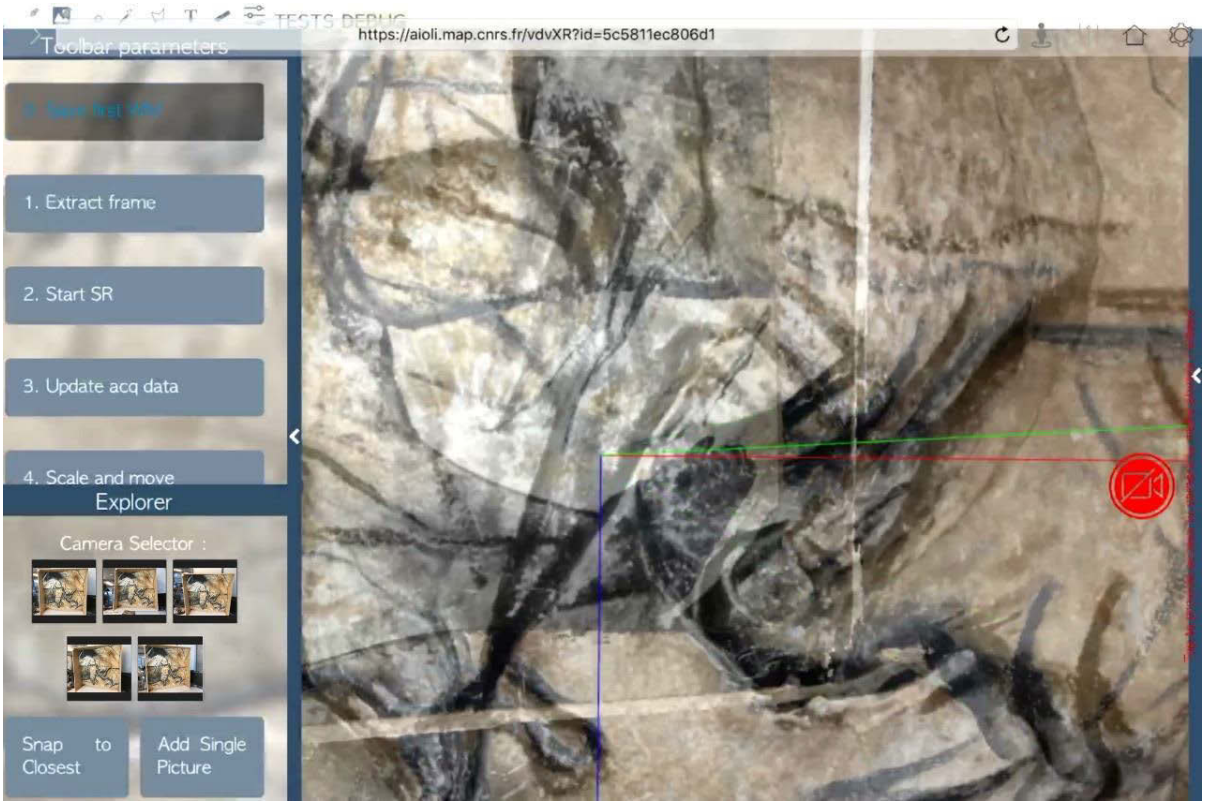


g. La sélection par « baguette magique » fonctionne quelle que soit la texture active. Sur les normales, elle permet de sélectionner rapidement des points coplanaires.

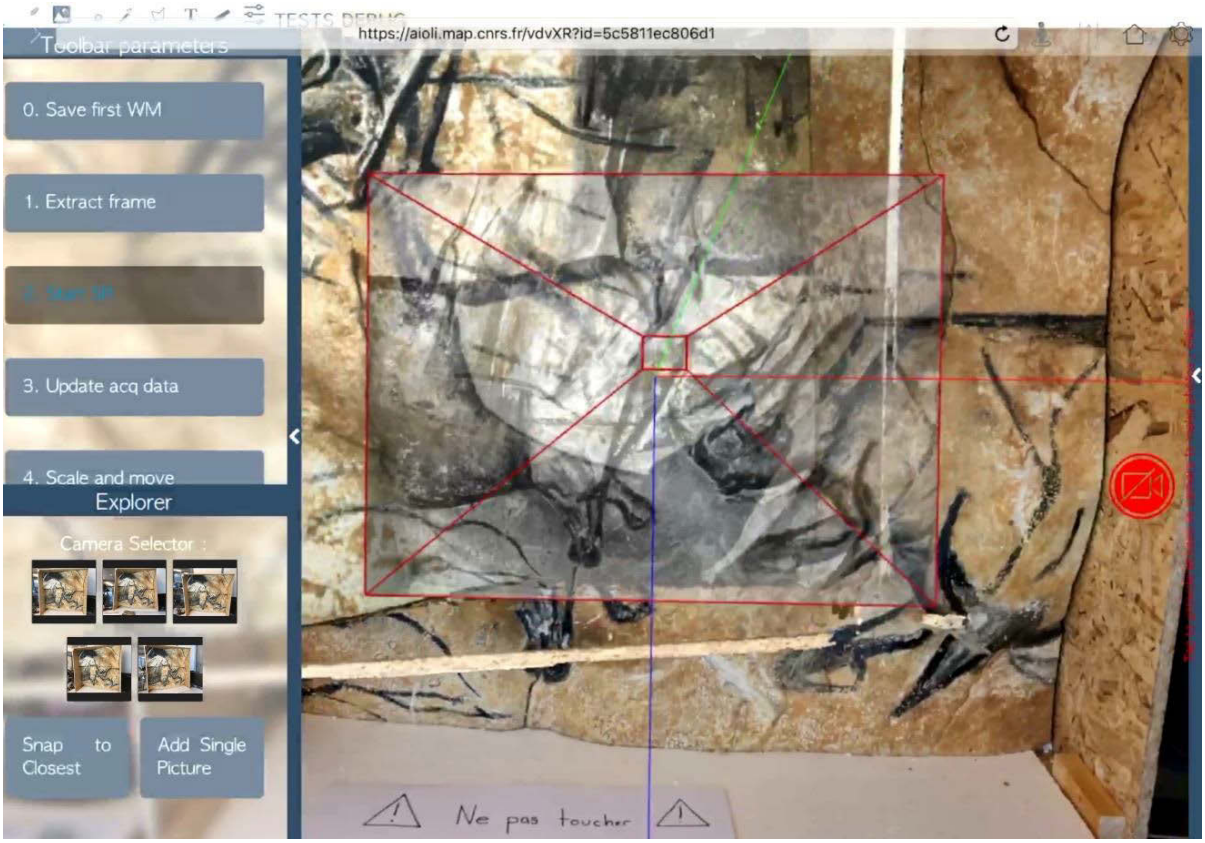
Annexe 15 : Captures d'écran de la preuve de concept durant la procédure de recalage. Une vidéo de démonstration est disponible à l'adresse *www.aioli.cloud/experiments/*



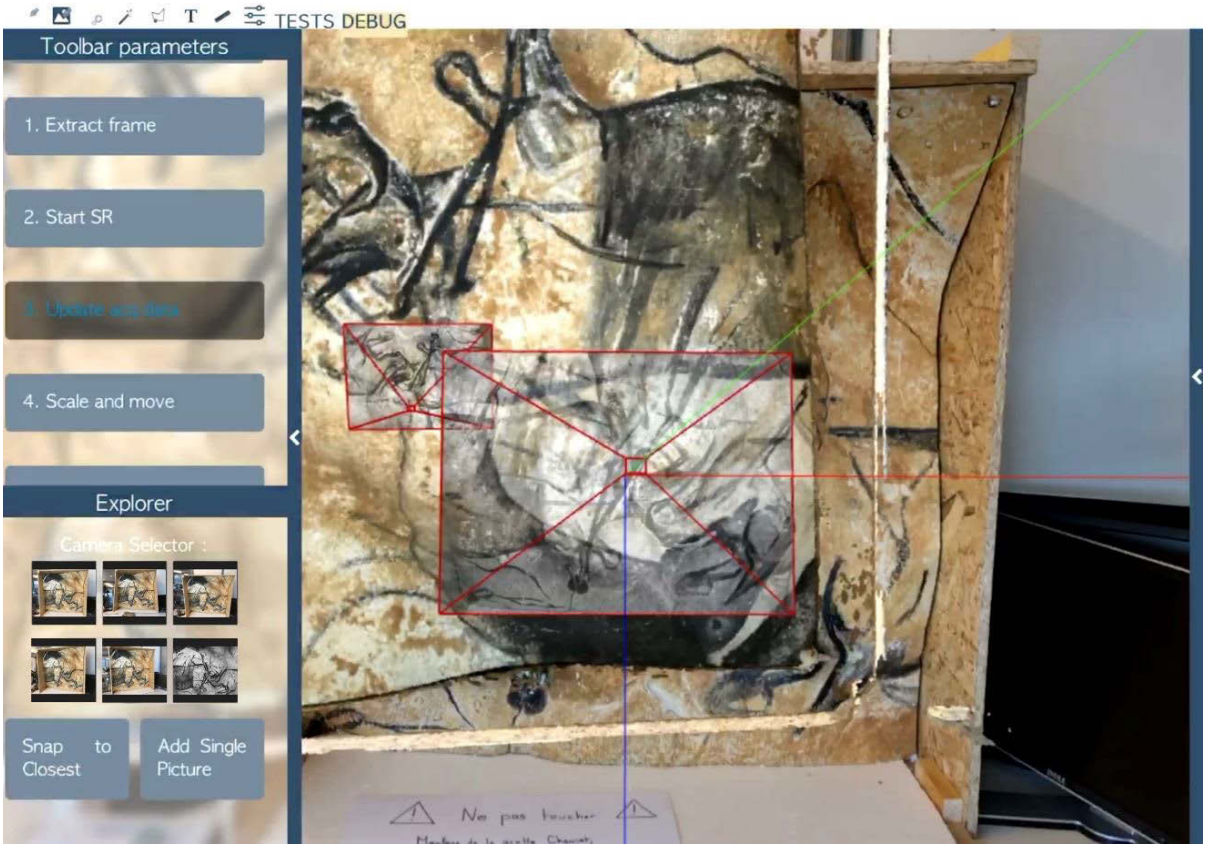
a. Chargement des données du projet. Elles ne sont pas alignées de manière cohérente sur la réalité.



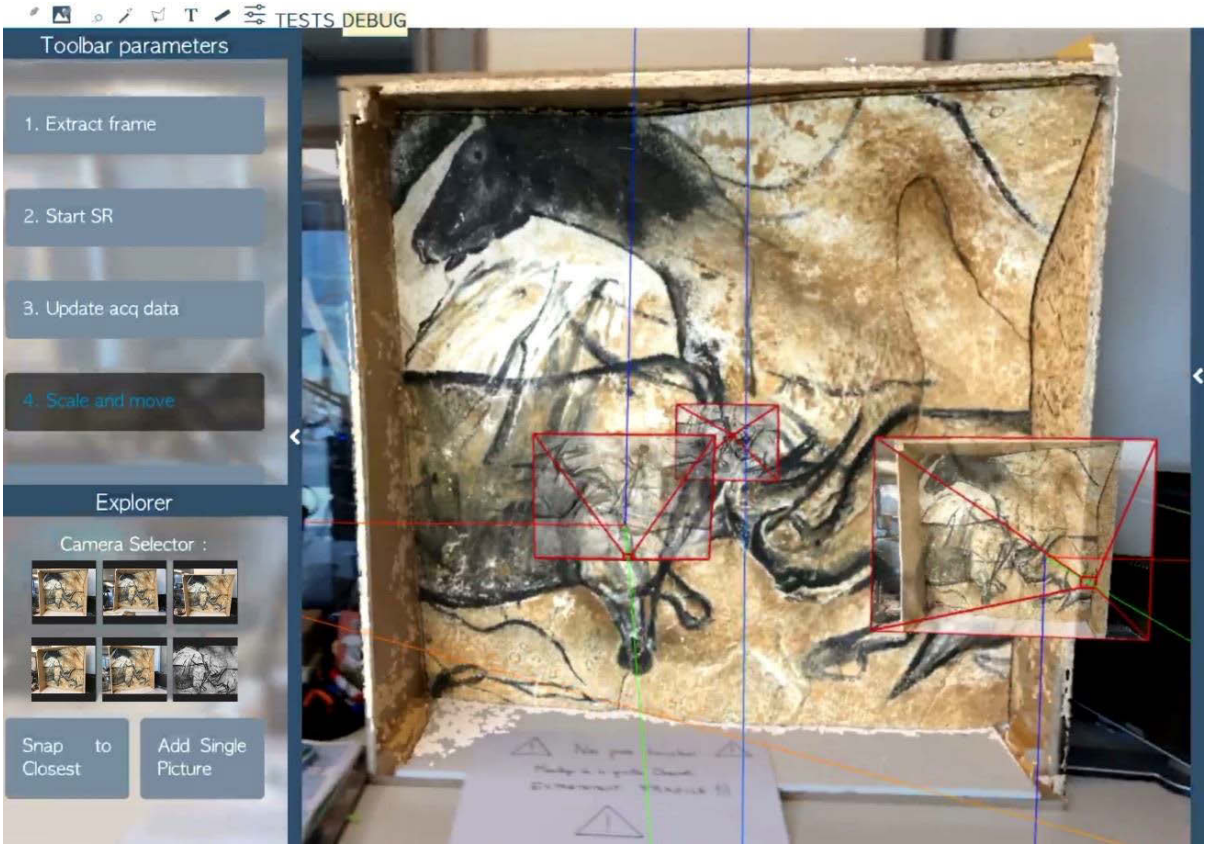
b. Extraction d'une *frame* du flux vidéo. Celle-ci conserve sa pose, calculée lors du suivi AR.



c. *Frame* extraite du flux vidéo. Elle conserve sa pose, calculée par SLAM lors du suivi AR.

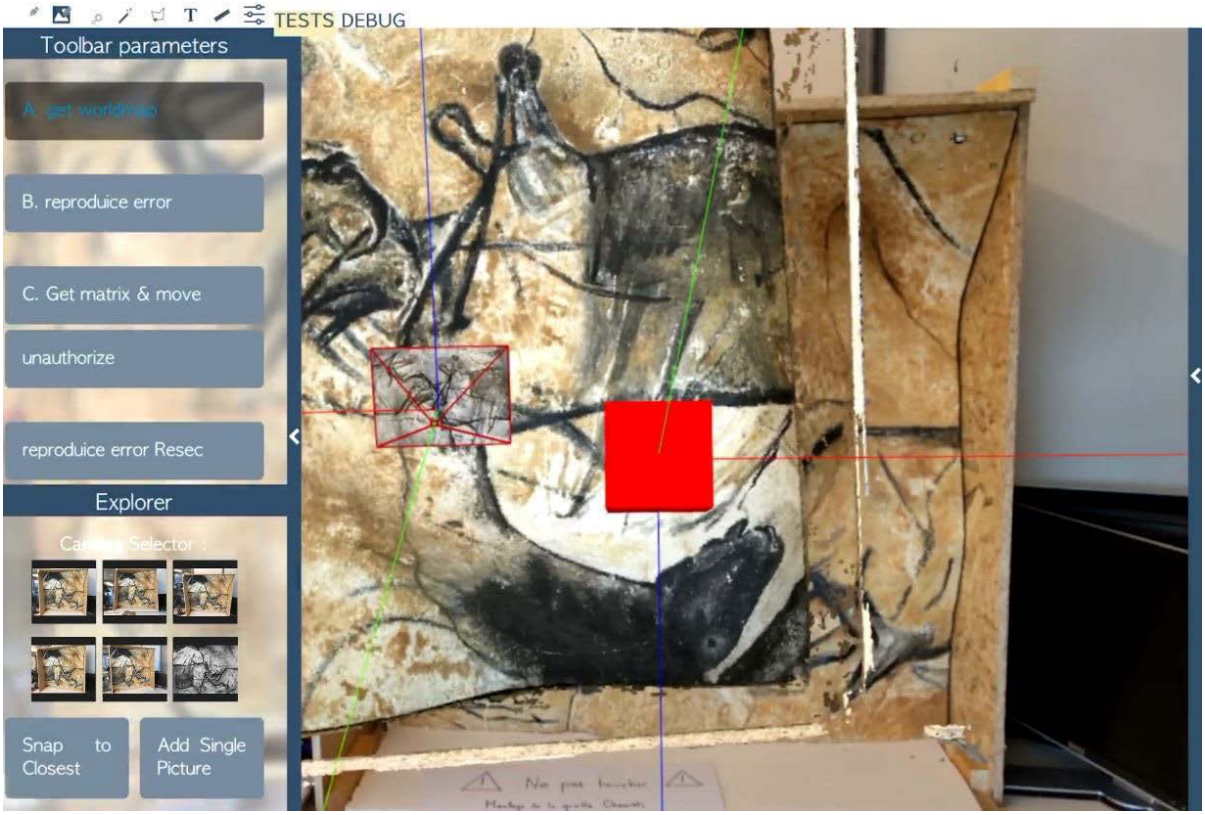


d. Mise à jour des données du projet. La *frame* extraite a été spatialisée autour du nuage de points du projet. Elle est chargée comme une image maîtresse, mais nous disposons toujours de sa pose AR.

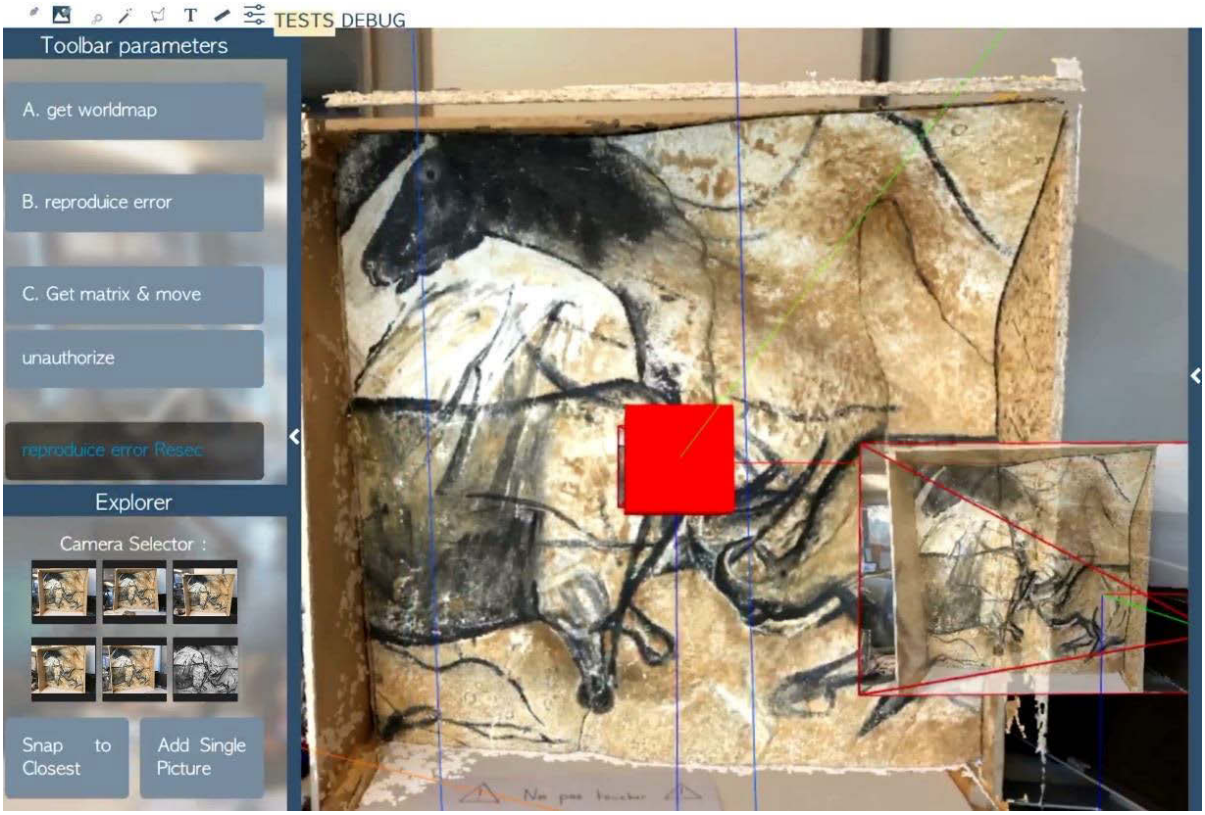


e. Recalage des données du projet après calcul de la matrice de passage entre la pose AR de la *frame* extraite et sa pose relative aux images maîtresses du projet.

Annexe 16 : Captures d'écran de la preuve de concept durant la procédure de recalage d'un projet précalibré par persistance AR



a. Chargement des ancres créées lors de la précalibration. L'ancre rouge donne la pose de la *frame* extraite à T_0.



b. Recalage du projet à partir de l'ancre et de la pose de l'image maîtresse resec.png.

Liste des publications

Publications dans des revues à comité de lecture

ABERGEL, V., JACQUOT, K., DE LUCA, L., VERON, P., 2019. **Towards a SLAM-based augmented reality application for the 3D annotation of rock art**. In : *Journal of Virtual Reality*. Vol. 19 n°2, p. 17-26. DOI 10.20870/IJVR.2019.19.2.2910.

ABERGEL, V., PINÇON, G., KONIK, S., JACQUOT, K., 2019. **Harmonisation et diffusion des ressources numériques 3D des grottes ornée**. In : *In Situ*. N°39, DOI 10.4000/insitu.21550.

ABERGEL, V., SALERI, R., BERGEROT, L., DE LUCA, L., 2016. **Personal hand-held devices and virtual reality passive technologies for digital heritage browsing: Lesson learnt with the KIVI project**. In : *SCIRES-it*. DOI 10.2423/i22394303v6n1p67.

Publications dans des actes de congrès avec comité de lecture

ABERGEL, V., SALERI, R., LEQUAY, H., et DE LUCA, L., 2015. **An immersive visualization kit for online 3D objects databases**. In : *Proceedings of the 2nd Digital Heritage International Congress (DigitalHERITAGE)*. Vol. 1, p 305-308. Grenade, Espagne. IEEE, DOI: 10.1109/DigitalHeritage2015.7413888.

MANUEL, A., M'DARHRI, A., ABERGEL, V., ROZAR, F. et DE LUCA, L., 2018. **A semi-automatic 2D/3D annotation framework for the geometric analysis of heritage artefacts**. In : *Proceedings of the 3rd Digital Heritage International Congress (DigitalHERITAGE)*. San Francisco, CA, USA. IEEE, p. 1-7.

PAMART, A., PONCHIO, F., ABERGEL, V., ALAOUI M'DARHRI, A., CORSINI, M., DELLEPIANE, M., MORLET, F., SCOPIGNO, R. et DE LUCA, L., 2019. **A complete framework operating Spatially-Oriented RTI in a 3D/2D cultural heritage documentation and analysis tool**. In : *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. Vol. XLII-2/W9, p. 573-580. DOI 10.5194/isprs-archives-XLII-2-W9-573-2019.

Posters et communications sans actes

ABERGEL, V., JACQUOT, K., DE LUCA, L., VERON, P., 2019. **Annotation sémantique multimodale : application au cas des grottes ornées**. In : *Rencontres Doctorales en Architecture et Paysage.* Poster. Lyon, France.

ABERGEL, V., 2019. **Annotation sémantique multimodale : application au cas des grottes ornées**. In : *Séminaire de l'UMR MAP.* [en ligne]. Poster. Nancy, France. Disponible à l'adresse : https://hal.archives-ouvertes.fr/hal-02181533.

ABERGEL, V., 2019. **Les relevés en architecture et l'application Aïoli**. In : *Réseau SOCle : Les méthodes du relevé d'art préhistorique.* Conférencier invité. Préfecture de Dordogne, Périgueux, France.

DE LUCA, L., ABERGEL, V., ALAOUI M'DARHRI, et MANUEL, A., 2018. **Hands on Aïoli: a reality-based 3D annotation platform for collaborative documentation of CH artifacts**. In : *The 3rd Digital Heritage International Congress (DigitalHERITAGE)*. San Francisco, CA, USA.

Violette ABERGEL

Relevé numérique d'art pariétal : définition d'une approche innovante combinant propriétés géométriques, visuelles et sémantiques au sein d'un environnement de réalité mixte

Résumé

Les progrès de ces dernières décennies dans les domaines de l'informatique et de la métrologie ont favorisé l'essor d'outils numériques de mesure de l'existant, marquant ainsi une évolution décisive dans de nombreuses disciplines. Dans le domaine du relevé d'art pariétal, leur introduction a rendu possible la collecte massive de données 2D et 3D en réponse à des besoins variés d'étude, de suivi, de documentation, d'archivage, ou encore de valorisation. Ces données constituent autant de nouveaux supports précieux pour la compréhension des objets d'étude, en particulier en ce qui concerne leur caractérisation morphologique. Cependant, malgré leurs nombreux potentiels, elles demeurent souvent sous-exploitées par manque d'outils facilitant leur manipulation, leur analyse, et leur enrichissement sémantique dans des contextes d'études pluridisciplinaires. En outre, ces méthodes tendent à reléguer l'engagement cognitif et analytique de l'observateur derrière son instrument de mesure, causant une rupture profonde entre les moments d'étude sur site et l'ensemble des traitements réalisés hors site, ou autrement dit, entre les environnements de travail réels et virtuels. Cette thèse propose de répondre à ces problèmes par la définition d'une approche intégrée permettant la fusion des aspects géométriques, visuels et sémantiques du relevé au sein d'un environnement de réalité mixte multimodal. A la croisée des domaines des systèmes d'informations patrimoniaux et de la réalité mixte, notre objectif est de garantir la continuité des travaux menés *in situ* et *ex situ*. Ce travail a abouti au développement d'une preuve de concept fonctionnelle permettant la consultation de données numériques 2D et 3D issues de relevé et leur annotation sémantique en réalité augmentée via une interface web.

Patrimoine culturel, Description sémantique, Réalité mixte, Spatialisation 3D, Systèmes d'informations

Résumé en anglais

The advances of the last decades in the fields of computer science and metrology have led to the development of efficient measurement tools allowing the digitization of the environment, bringing a decisive evolution in many fields. In the case of rock art surveying, their introduction has allowed a massive gathering of 2D and 3D data, meeting various needs for study, monitoring, documentation, archiving, or dissemination. These data provide new and valuable supports for the understanding of the objects of study, in particular concerning their morphological characterization. However, in spite of their great potentials, they often remain under-exploited due to the lack of tools facilitating their manipulation, analysis, and semantic enrichment in multidisciplinary study contexts. Moreover, these methods tend to relegate the cognitive and analytical engagement of the observer behind the measurement tool, causing a deep break between on-site study moments and all off-site processing, or in other words, between real and virtual work environments. This thesis proposes to address these problems by defining an integrated approach allowing the fusion of the geometric, visual and semantic aspects of surveying within a single multimodal mixed reality environment. At the crossroads of the fields of heritage information systems and mixed reality, our goal is to ensure an informational continuity between in situ and ex situ analysis activities. This study led to the development of a functional proof of concept allowing the visualization of 2D and 3D digital data from surveys and their semantic annotation in augmented reality through a web interface.

Cultural heritage, Semantic description, Mixed reality, 3D spatialization, Information systems