HAL
open science

Visualisation scientifique de grands volumes de données : Pour une approche perceptive

Christian Boucheny

To cite this version:

Christian Boucheny. Visualisation scientifique de grands volumes de données : Pour une approche perceptive. Informatique [cs]. Université Joseph-Fourier - Grenoble I, 2009. Français. NNT : . tel-00438464

HAL Id: tel-00438464

https://theses.hal.science/tel-00438464

Submitted on 4 Dec 2009

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.

Visualisation scientifique interactive de grands volumes de données : Pour une approche perceptive

Christian BOUCHENY

THÈSE

présentée pour l'obtention du titre de
Docteur de l'Université Joseph Fourier, spécialité Informatique,

préparée au sein du
Projet EVASION, Laboratoire Jean Kuntzmann, UMR 5524
et du
Laboratoire de Physiologie de la Perception et de l'Action, UMR 7152
et de
Electricité de France, Recherche et Développement

soutenue le 13 février 2009 *devant le jury composé de :*

Président	Georges-Henri	COTTET
Rapporteurs	Bruno	LEVY
	Pascal	MAMASSIAN
Examinateurs	Claude	PUECH
	Charles	HANSEN
Co-directeurs de thèse	Georges-Pierre	BONNEAU
	Jacques	DROULEZ
Encadrant EDF	Guillaume	THIBAULT

SOMMAIRE

Remerciements **5**

Résumé **7**

Abstract **9**

Liste des abréviations **11**

Introduction **13**
Contexte et motivations 15
Organisation du document 16

1 La visualisation scientifique à EDF **17**
1 Simulation et visualisation 18
2 Domaines et échelles 22
3 Moyens matériels et logiciels 25
4 Pratiques et difficultés 28
5 Conclusion et objectifs 33

2 Perception visuelle et informatique graphique **35**
1 Rendre plus vite 36
2 Rendre mieux 46
3 Place de la perception en visualisation scientifique 55
4 Perspectives, orientation du travail 59

3 Evaluation des rendus volumiques **63**
1 Les techniques de rendu volumique direct 65
2 Perception de la transparence 69
3 Expérience 1 : perception statique en DVR 72
4 Expérience 2 : DVR et effet cinétique de profondeur 83
5 Expérience 3 : apport de la perspective linéaire 89

6 Conclusion et perspectives 97

4 EyeDome Lighting **101**
1 Perception de l'éclairage et modèles infographiques 102
2 Algorithme de l'EyeDome Lighting 115
3 Extension de l'EyeDome Lighting 127
4 Conclusion et perspectives 132

5 Focus+Contexte interactif avec extrusion conique **139**
1 Techniques variées de Focus+Contexte 140
2 Algorithme d'extrusion conique sur le GPU 144
3 Rendu complet en écorché 151
4 Discussion et perspectives 157

Conclusion et perspectives **161**
Contributions 161
Perspectives 162

A Calcul de la translation compensatoire **167**

Table des figures **171**

Table des matières **175**

Bibliographie **181**

Remerciements

Remercier est un exercice délicat. On souhaiterait n'oublier personne et exprimer en quelques mots ce que les autres nous ont apporté à différentes étapes de ce long voyage que représente un travail de thèse. Un mélange de relations personnelles et professionnelles, parfois les deux, avec des moments très forts. Un deux trois, je prends ma respiration, et me lance pour adresser tous mes remerciements.

A mes parents pour leur soutien indéfectible, eux qui m'ont permis de grandir et d'avancer même dans les moments les plus difficiles.

A mes deux relectrices pointilleuses, qui ont absorbé toute ma prose avec énergie malgré leur « naïveté scientifique », pour rendre ce tapuscrit plus lisible.

A tous ceux qui m'ont progressivement guidé vers cette thèse en éveillant ma curiosité sur certains sujets et en m'encourageant à me lancer dans un tel projet. A Philippe Fuchs et Guillaume Moreau, qui m'ont fait découvrir l'univers de la Réalité Virtuelle, et m'ont initié les premiers à la démarche de recherche. Au professeur Alain Berthoz pour avoir fait partager son enthousiasme scientifique au jeune ingénieur que j'étais alors en évoquant notamment le « chant des neurones », me permettant de plonger dans le vaste monde des sciences cognitives. A Angelo Arleo qui, après m'avoir fait grandir scientifiquement, et alors que je doutais du chemin à prendre, m'a encouragé et motivé pour me lancer dans le grand bain et suivre le panneau « doctorat », à force de discussions patientes et constructives. Bien sûr, merci à mes directeurs de thèse et encadrants industriels, Georges-Pierre Bonneau, Jacques Droulez, Guillaume Thibault et Stéphane Ploix, pour m'avoir proposé ce sujet, m'avoir communiqué leur enthousiasme et m'avoir soutenu dans les moments délicats. Une thèse pluridisciplinaire, avec de plus un encadrement industriel, peut se révéler particulièrement délicate à mener, avec notamment le risque d'un tiraillement permanent entre des préoccupations très différentes. Merci de m'avoir laissé la liberté de naviguer dans ces différentes eaux plus ou moins troubles, tout en m'apportant les lumières relatives à chacun des domaines concernés. Administrativement, merci également à EDF et à l'ANRS pour avoir financé ces travaux de thèse via une bourse CIFRE.

Aux ingénieurs-chercheurs d'EDF R&D pour le temps consacré lors de différents entretiens qui m'ont permis de mieux cerner les problématiques inhérentes à leur métier et à leurs pratiques (pour n'en citer que quelques uns : Christophe Péniguel, Isabelle Rupp, Yvan Fournier, Alexandre Douce, Olivier Diard). Merci à Christophe Mouton pour son aide précieuse dans la récupération de données de travail au début de cette thèse, et pour sa disponibilité quand il y avait un problème technique à

résoudre. A Anne Marie pour ses encouragements quand je trouvais le temps de la rédaction un peu long. A Guillaume pour m'avoir également fortement soutenu et encouragé dans cette période agitée, et pour m'avoir entrouvert de nouvelles portes. J'adresserai enfin ici une mention du courage à Louis Bimaï, pour sa patience à essayer de comprendre mon code afin de réaliser une première intégration de l'EyeDome Lighting dans VTK au cours de son stage.

Et pour conclure cette salve, je tiens également à adresser mes remerciements à Grenoble pour avoir attendu que je reparte pour faire tomber la neige, au moins ai-je pu randonner et manger en plein air en plein hiver, à défaut de skier. Aux amis du théâtre de la distanciation, avec qui j'ai bien ri malgré la Grand'Peur. A Anne Pierson bien entendu pour son aide et sa disponibilité en toutes circonstances, à Marie-Paule Cani pour m'avoir accueilli au sein de l'équipe-projet EVASION, à tous les collègues d'EVASION et ARTIS pour les bons moments passés ensemble ; qu'il fut dur de rentrer sur Paris ! A Claire et Matthieu pour une année de colocation très agréable dans notre petite maison échirolloise, la chaleur de l'atmosphère annihilant bien vite l'humidité ambiante. Et enfin à Sébastien, pour son amitié et son soutien sans faille, et aux amis de longue date, qui se reconnaîtront dans ces mots de Francis Bacon : « L'amitié double les joies et réduit de moitié les peines ».

Résumé

L'explosion de la puissance de calcul permet actuellement de réaliser des simulations physiques comportant jusqu'à plusieurs milliards d'éléments. Pour extraire l'information importante de tels volumes de données, les ingénieurs doivent disposer d'outils de visualisation permettant d'explorer et d'analyser interactivement les champs calculés. Cette thèse vise à améliorer les visualisations réalisées en prenant en compte les caractéristiques de la perception visuelle chez l'homme, notamment en ce qui concerne la perception de l'espace et du volume au cours de visualisations 3D denses. D'abord, trois expériences de psychophysique ont montré que les rendus volumiques, reposant sur l'accumulation ordonnée de transparences, génèrent des difficultés importantes de perception de la profondeur. Cela est particulièrement vrai lors de visualisations statiques, mais l'ajout de mouvement dans la scène ainsi que l'amplification de la projection perspective permettent de lever en partie ces ambiguïtés. Ensuite, deux algorithmes améliorant la perception de l'espace lors de la visualisation de structures tridimensionnelles complexes ont été développés. Leur implémentation sur GPU permet des rendus interactifs indépendamment de la nature géométrique des données visualisées. L'EyeDome Lighting, un nouvel ombrage non photoréaliste basé sur l'image de profondeur, améliore la perception des formes et des profondeurs dans des scènes 3D complexes. Par ailleurs, une nouvelle technique d'écorché dépendant du point de vue a été implémentée. Elle autorise la visualisation d'objets normalement occultés tout en rendant compte de la structure des surfaces masquantes.

Abstract

With the fast increase in computing power, numerical simulations of physical phenomena can nowadays rely on up to billions of elements. To extract relevant information in the huge resulting data sets, engineers need visualization tools permitting an interactive exploration and analysis of the computed fields. The goal of this thesis is to improve the visualizations performed by engineers by taking into account the characteristics of the human visual perception, with a particular focus on the perception of space and volume during the visualization of dense 3D data. Firstly, three psychophysics experiments have shown that direct volume rendering, a technique relying on the ordered accumulation of transparencies, provide very ambiguous cues to depth. This is particularly true for static presentations, while the addition of motion and exaggerated perspective cues help to solve part of these difficulties. Then, two algorithms have been developed to improve depth perception during the visualization of complex 3D structures. They have been implemented on the GPU, to achieve interactive renderings independently of the geometric nature of the analysed data. EyeDome Lighting is a new non-photorealistic shading technique that relies on the projected depth image of the scene. This algorithm enhances the perception of shapes and relative depths in complex 3D scenes. Also, a new fast view-dependent cutaway technique has been implemented, which permits to access otherwise occluded objects while providing cues to understand the structure in depth of masking objects.

Liste des abréviations

API	Application Programming Interface
BRDF	Bidirectional Reflectance Distribution Function
CAO	Conception Assistée par Ordinateur
CPU	Central Processing Unit
CT	Computed Tomography
CUDA	Compute Unified Device Architecture
DRR	Digitally Reconstructed Radiograph
DVR	Direct Volume Rendering
EDL	EyeDome Lighting
FBO	Frame Buffer Object
F+C	Focus+Contexte
FT	Fonction de Transfert
GCD	Gaze Contingent Display
GLSL	OpenGL Shading Language
GPU	Graphics Processing Unit
GPGPU	General-Purpose Computing on Graphics Processing Unit
HMD	Head-Mounted Display
HPC	High Performance Computing
IHM	Interface Homme-Machine
IRM	Imagerie par Résonance Magnétique nucléaire
KDE	Kinetic Depth Effect
LOD	Level Of Detail
MIP	Maximum Intensity Projection
NPR	Non-Photorealistic Rendering
SfM	Structure from Motion
SSAO	Screen Space Ambient Occlusion
SVH	Système Visuel Humain
TDL	TorchDome Lighting
VDP	Visible Differences Predictor
VTK	Visualization Toolkit

Introduction

> Pour voir une chose il faut la comprendre [...] Si nous avions une vision réelle de l'univers, peut-être pourrions-nous le comprendre.
>
> Jose Luis Borges

La visualisation scientifique vise à produire des représentations visuelles exploitables pour analyser et comprendre des données complexes, que celles-ci soient le fruit d'une simulation numérique ou d'une acquisition par un appareil de mesure (IRM, scanner...). Il s'agit d'une branche de l'informatique graphique dont elle partage certaines méthodes, mais avec un support, un objet représenté, qui est de nature bien spécifique, parfois non tangible. De plus, les modes de représentation utilisés peuvent être soumis à certaines conventions directement liées au domaine d'application auquel la visualisation est consacrée.

Le recours à un support visuel en tant que moyen d'accéder à une connaissance sur la nature n'est pas nouveau en soi. Par exemple, le dessin scientifique était déjà pratiqué par Léonard de Vinci, pour qui l'observation et la représentation des tourbillons de l'eau étaient un moyen de comprendre ce phénomène naturel. Comme le souligne Daniel Arasse, « les différences entre dessin scientifique et dessin préparatoire à l'œuvre d'art s'estompent fortement. Tels que les pratique Léonard, en effet, la ligne a pour fonction dans l'un et l'autre cas de faire accéder à la figure une forme, un contour, invisibles ou mal visibles 'en réalité', latents dans le flux continu des apparences. » ([Ara03], p.284).

La visualisation scientifique constitue le pendant informatique du dessin scientifique. En tant que telle, il s'agit d'une discipline relativement récente, née à la fin des années 1980, et profondément motivée par la nécessité des différentes communautés scientifiques de disposer d'outils efficaces pour exploiter des masses de données de plus en plus importantes. Au contraire de la visualisation d'informations dont l'objet d'étude est constitué de données abstraites, la visualisation scientifique s'intéresse, par essence, à des structures spatiales qui possèdent une extension physique propre. Cependant, elle doit donner corps à des objets ne possédant pas nécessairement de support matériel, comme des champs de pressions ou de vitesses d'un fluide.

La nature d'une visualisation dépend fortement du domaine scientifique concerné, tant du point de vue de la nature des objets à analyser que des outils utilisables en pratique pour atteindre le but recherché. En médecine, le praticien visualisant une angiographie a une connaissance forte des vais-

seaux sanguins présents à l'image, et l'examen consiste par exemple à rechercher une anomalie dans la structure physique de ces vaisseaux. La visualisation doit ici aider à distinguer clairement les tissus numérisés. Elle ne porte que sur un unique champ scalaire de densité, les structures principales étant relativement bien segmentées autour d'une valeur caractéristique. La visualisation doit permettre de bien segmenter visuellement les éléments clés, et permettre de modifier interactivement le point de vue pour aider à comprendre les relations spatiales au sein du volume. Si l'on souhaite analyser les résultats d'une simulation numérique de mécanique des fluides relative au refroidissement d'une centrale, la situation est profondément différente. Certes, l'ingénieur connaît la structure physique sur laquelle le calcul a été réalisé, et peut avoir une première idée des principaux phénomènes en jeu. Cependant, il doit analyser les différents champs enregistrés (scalaires, vectoriels voire tensoriels) afin de détecter d'éventuelles anomalies ou singularités, et poursuivre son exploration dans le volume et dans l'espace des données afin d'appréhender la structure et les causes de ces phénomènes.

Dans tous les cas, l'interactivité est un point clé de la visualisation. Pour mieux comprendre ce qui est observé, le praticien ou l'ingénieur doivent pouvoir changer rapidement le point de vue, c'est-à-dire naviguer dans la scène affichée, et accéder aussi vite que possible à des représentations explicites des champs volumiques à analyser. Cette interactivité est d'autant plus critique que les volumes de données à explorer sont importants, et donc que l'espace à explorer est large. Mais une augmentation du volume des données à traiter et afficher constitue naturellement un frein à cette interactivité, demandant des efforts de calculs beaucoup plus conséquents.

Les gros volumes de données auxquels la visualisation doit faire face aujourd'hui sont une conséquence directe de la croissance exponentielle de la puissance de calcul disponible à des fins de simulation numérique. Grâce au développement de supercalculateurs pouvant effectuer aujourd'hui quasiment un million de milliards d'opérations par seconde, il est possible de modéliser de plus en plus finement les structures physiques que l'on désire étudier, et d'étendre les simulations à des modèles de plus en plus larges. Il n'est plus nécessaire de se restreindre à une sous-partie judicieusement choisie et de tenter d'extrapoler les résultats obtenus. Là où l'on devait se contenter de discrétisations limitées à quelques centaines de primitives géométriques en deux dimensions, il est désormais envisageable de réaliser des simulations relatives à des millions et même des milliards d'éléments finis, et ce sur plusieurs dizaines à centaines de pas de temps. Les nombreux champs physiques calculés représentent alors des volumes considérables de données qu'il faut être capables de stocker d'abord, puis d'exploiter par une analyse visuelle interactive.

Par ailleurs, et du moins jusqu'à aujourd'hui, la visualisation se heurte à un problème important : la projection sur un écran bidimensionnel de quantités importantes de données denses dans le volume. Les images produites peuvent être difficiles à comprendre, notamment en ce qui concerne les formes et profondeurs relatives des différents éléments représentés, le nombre d'indices visuels de profondeur étant nécessairement restreint. De plus, les occultations nécessairement introduites obligent à limiter le nombre d'objets effectivement projetés en même temps. Pour explorer tout le volume occupé par un champ donné, il est souvent nécessaire d'afficher séquentiellement plusieurs représentations et de reconstruire mentalement le lien entre ce qui a été ainsi observé, ce qui constitue une tâche cognitive particulièrement complexe.

Les connaissances relatives à la perception visuelle chez l'homme peuvent aider à apporter des solutions pratiques à certains des problèmes mentionnés. Face à la quantité de données à afficher, on peut envisager d'utiliser des représentations géométriques simplifiées moins gourmandes en temps de calcul. Les différents niveaux de détails peuvent être réglés en tenant compte des particularités du système visuel, de façon à prévenir toute perte effective d'information au cours de l'affichage. Par ailleurs, comprendre précisément comment notre système visuel interprète et combine différentes informations afin de générer une perception stable de l'espace peut nous aider à proposer des modes de représentation plus efficaces, en sélectionnant les indices les plus pertinents compte tenu des res-

sources calculatoires disponibles. A l'heure où les simulations numériques produisent des volumes de résultats de plus en plus importants, une collaboration entre les domaines de recherche de la visualisation scientifique et de la perception visuelle paraît fondamentale pour répondre efficacement aux exigences grandissantes d'exploration visuelle de ces données complexes.

Contexte et motivations

Les travaux présentés dans cette thèse ont été réalisés dans le cadre d'une convention industrielle de formation par la recherche (CIFRE) financée par EDF R&D. Ils résultent d'une collaboration avec le groupe de visualisation scientifique à EDF R&D, le projet EVASION du laboratoire Jean Kuntzmann (INRIA Rhône-Alpes, UJF, CNRS) et le laboratoire de physiologie de la perception et de l'action (Collège de France, CNRS).

A EDF, la simulation numérique connaît un développement important, et complète de plus en plus les expériences réalisées sur des maquettes à taille réduite. Elle permet d'essayer de mieux comprendre les causes de phénomènes observés sur le terrain, comme le vieillissement prématuré de certaines structures, et de proposer des réponses adaptées. Elle a également pour vocation d'avancer de nouvelles solutions optimisées, par exemple pour évacuer la chaleur du cœur d'une centrale, et de prédire les conséquences de certains événements, comme un choc thermique soudain dans un circuit de refroidissement. Les domaines physiques concernés sont multiples : thermique, physique des matériaux, mécanique des structures, mécanique des fluides, électromagnétisme, sédimentologie, neutronique... Des simulations couplées entre différents domaines peuvent parfois être réalisées, par exemple en thermohydraulique.

Des entretiens réalisés au début de la thèse avec plusieurs ingénieurs de la R&D ont permis de faire ressortir un ensemble de difficultés relatives à la visualisation de leurs résultats de calculs. Parmi celles-ci, la compréhension des structures spatiales affichées et la possibilité de naviguer rapidement dans les données et le volume occupent une place importante. L'interactivité est apparue comme une condition indispensable à la visualisation, les ingénieurs préférant parfois éviter de recourir à des représentations plus claires mais nécessitant un temps de précalcul de plusieurs minutes. Par ailleurs, il ressort des cas observés que les maillages sur lesquels reposent les calculs, et donc la visualisation, incluent des primitives géométriques diverses : tétraèdres, hexaèdres, pentaèdres... Il ne saurait donc être question de proposer des solutions de visualisation spécifiquement adaptées à un type particulier de primitives[1].

Notre motivation principale au cours de cette thèse a concerné l'amélioration de la perception de la forme, de la profondeur et du volume dans des scènes complexes de visualisation. Nous avons d'une part cherché à évaluer des méthodes de rendu volumique visant à afficher toute une portion du volume en utilisant une accumulation de transparences. D'autre part, nous avons proposé de nouveaux algorithmes d'ombrage et de rendu par écorché, afin de faciliter la perception de la profondeur et de résoudre certaines occultations. Ces développements ont été guidés par un souci d'interactivité, ce qui a conduit à des implémentations réalisées directement sur la carte graphique. Cette approche présente par ailleurs l'avantage de ne reposer que sur des informations disponibles dans l'espace image, ignorant de fait la diversité des objets géométriques sous-jacents.

[1] En visualisation scientifique, de nombreux algorithmes d'ombrage et d'exploration par déformation du volume requièrent des données organisées selon une grille régulière, ce qui facilite grandement l'ensemble des traitements réalisés. Cette structure de données est typique des visualisations médicales, mais beaucoup moins fréquente dans les simulations numériques de type ingénieur.

Organisation du document

Dans le premier chapitre de ce manuscrit, nous nous intéressons plus en détail à la simulation numérique et à la visualisation scientifique telles qu'elles sont pratiquées à la R&D d'EDF. A partir des entretiens réalisés avec plusieurs ingénieurs, nous dégageons les problématiques principales relatives à l'exploration visuelle de gros volumes de données, en nous concentrant plus particulièrement sur les difficultés d'ordre perceptif.

Le second chapitre présente quelques éléments relatifs à la perception visuelle, et la manière dont ces connaissances sont déjà exploitées ou pourraient l'être dans le domaine de l'informatique graphique. Notamment, les algorithmes peuvent tirer profit des limites du système visuel, tant bas-niveau qu'attentionnelles, pour réduire la quantité d'information à afficher et ainsi accélérer les rendus. D'un autre côté, comprendre comment les différents indices de perception de la profondeur sont intégrés par le système visuel peut aider à produire des visualisations véhiculant plus efficacement le sens des relations spatiales entre les éléments observés.

Dans le troisième chapitre, nous essayons d'évaluer la précision avec laquelle le système visuel humain est capable de percevoir des relations de profondeur dans des rendus volumiques. Dans de tels rendus, les problèmes d'occultations de surfaces sont en partie résolus grâce à l'accumulation de transparences en profondeur. Cependant, les images produites apparaissent particulièrement ambiguës. Nous montrons que l'ajout d'informations dynamiques, par le mouvement de la scène, et d'une amplification perspective permettent en partie de résoudre les incertitudes relatives à la perception de la profondeur dans ces rendus transparents.

Nous présentons ensuite dans le quatrième chapitre un nouvel algorithme d'ombrage baptisé Eye-Dome Lighting. L'ombrage d'une scène demeure un problème difficile en informatique graphique, et il est délicat de trouver une méthode permettant d'obtenir interactivement un résultat convaincant, c'est-à-dire qui facilite la perception des formes et relations spatiales. L'approche non photoréaliste proposée génère des images plus claires sur ce plan. Etant implémentée directement sur la carte graphique, elle permet un calcul très rapide d'ombrage quelle que soit la nature géométrique des objets représentés, et est donc particulièrement adaptée à la visualisation scientifique.

Dans le cinquième chapitre, nous proposons une méthode interactive de génération d'images en écorché, elle aussi implémentée directement sur la carte graphique. Les rendus volumiques ne permettant pas de représenter clairement une scène complexe, il s'agit ici de découper selon le point de vue la scène de visualisation autour de structures sur lesquelles l'attention de l'utilisateur est focalisée, afin de gérer les problèmes d'occultation. Une découpe conique dans le volume facilite la perception de la région de focus dans le contexte environnant. Nous montrons comment cet algorithme peut être intégré à une architecture complète de rendu incluant notamment l'ombrage présenté précédemment.

Enfin, nous concluons ce document par une discussion des perspectives relatives au travail présenté, et plus généralement des orientations pertinentes concernant les modèles de représentations en visualisation scientifique.

CHAPITRE 1

La visualisation scientifique à EDF

Dans le cadre d'une entité comme EDF, la simulation de comportements physiques vise essentiellement à mieux appréhender les phénomènes complexes à l'œuvre dans des installations industrielles de géométrie bien définie, de manière à fournir des données aussi bien qualitatives que quantitatives fiables pour améliorer le processus de décision. Si l'on considère l'exemple d'une centrale électrique, l'enjeu est multiple. Il est d'abord utile de comprendre les causes de fatigue de certaines pièces en situation d'exploitation normale, afin de planifier au mieux les interventions de remplacement des éléments nécessitant un arrêt coûteux des installations. Pour des raisons de sécurité, il est également nécessaire de prévoir la réponse du système à certains scénarios particuliers, comme un choc thermique brutal, et de définir en conséquence les mesures à prendre pour éviter tout problème grave. Enfin, l'efficacité de nouvelles installations encore en projet doit pouvoir être explicitée et optimisée dans la limite des nombreuses contraintes physiques et architecturales. Pour cela, il faut disposer d'un modèle géométrique précis des installations sur lequel on puisse reproduire une partie ou la totalité des phénomènes physiques en jeu, qu'il s'agisse d'échanges thermiques, de mécanique des fluides ou des structures, voire de neutronique.

Longtemps, les maquettes à taille réduite ont constitué le seul moyen de réaliser de telles études. Par une reproduction exacte des phénomènes physiques en jeu, quoiqu'à une échelle différente, une telle approche permet d'obtenir des résultats relativement fiables. En contrepartie, les maquettes constituent un système fermé dans lequel seules quelques sondes très ponctuelles permettent de connaître l'état du système. Le positionnement de ces mesures doit être réalisé précisément de manière à obtenir les informations les plus pertinentes, en des points critiques de l'installation, tout en limitant leur impact sur le fonctionnement même de l'ensemble.

L'explosion de la puissance de calcul des ordinateurs permet désormais de réaliser de telles études au moyen de simulations numériques de plus en plus poussées. Les équations physiques approchées sont introduites dans des codes de calcul exécutés sur une modélisation discrétisée de la géométrie. L'état du système, représenté par l'ensemble des variables intégrées dans les équations, est alors connu sur l'ensemble du volume étudié et pour les différents pas de temps calculés. Il est ainsi possible de comprendre plus finement les causes des phénomènes observés, pour peu que ceux-ci repro-

duisent effectivement la réalité physique sans la déformer, ce qui doit permettre de faciliter la prise de décision résultant de l'étude. Cependant, en fonction de la physique mise en œuvre et de la complexité de la géométrie, il peut être nécessaire de recourir à une discrétisation très fine du volume, ce qui génère de très gros volumes de données. Actuellement, les maillages comportant des dizaines voire centaines de millions d'éléments se généralisent, et les résultats de simulations peuvent occuper plusieurs téraoctets en mémoire. La visualisation de tels ensembles, dont le but est de permettre l'analyse des résultats, pose alors de nombreux problèmes, tant matériels et logiciels que perceptifs.

Au sein de la R&D, le développement des moyens mis en œuvre pour la simulation numérique est actuellement structuré par un défi transverse aux différents cœurs de métier, baptisé *Simuler pour décider*. Ce projet regroupe les problématiques traditionnelles liées au calcul haute performance (*HPC*) et au développement des codes de calcul, mais comprend également un volet de visualisation scientifique. L'objectif clairement établi consiste à mettre en place les moyens de visualisation 3D interactive des résultats de calcul.

Afin de mieux comprendre les enjeux spécifiques à la visualisation, nous explicitons dans un premier temps le principe d'une simulation, depuis sa mise en place jusqu'à l'analyse des résultats (section 1). Nous nous intéressons ensuite à la définition des domaines et échelles, qui influencent fortement les modes de visualisation (section 2), ainsi qu'aux moyens matériels et logiciels qui définissent ce qui est pratiquement réalisable (section 3). Nous concluons par une analyse des pratiques et difficultés observées chez des ingénieurs dans le cadre de la visualisation de leurs résultats (section 4), afin de définir les priorités de notre travail (section 5).

1 De la préparation de la simulation aux conclusions de la visualisation

Le déroulement classique d'une étude est illustré sur la figure 1.1. Après avoir clairement établi les objectifs et résultats attendus, une modélisation 3D discrétisée du domaine physique est réalisée, si elle n'est pas déjà disponible. Celle-ci constitue le support de la simulation numérique, réalisée sur plusieurs pas de temps. Une fois les calculs convergés, les résultats sont analysés visuellement. Des problèmes liés au maillage ou aux codes de calcul peuvent apparaître à ce stade, ce qui nécessite de reprendre le travail en amont. Sinon, les résultats peuvent être exploités pour proposer des solutions aux problèmes posés. Il s'agit ici d'un fonctionnement en boucle ouverte, dans lequel l'ingénieur (mécanicien, neutronicien...) intervient à tous les niveaux. Il est en théorie possible de réaliser des simulations plus interactives (*computation steering*) dans lesquelles l'utilisateur n'attend pas passivement les résultats, mais les visualise au cours du calcul et interagit éventuellement en modifiant certains paramètres ou algorithmes. Cependant, dans le cadre d'études complexes pour lesquelles les tendances principales des résultats ne sont pas connues par avance, une telle approche est délicate à réaliser et les ingénieurs préfèrent définir fermement toutes les conditions de l'étude, quitte à relancer les calculs avec d'autres paramètres si le besoin s'en fait sentir.

1.1 Création du maillage

Dans le cas où la simulation porte sur une structure n'ayant fait l'objet d'aucune étude, il est nécessaire de créer le maillage volumique qui servira de support aux calculs. Cette étape délicate peut être réalisée à partir de plans ou d'un modèle CAO, soit par un traitement automatique (logiciels de maillage de type *Ideas* ou *Simail*), soit par un expert en 3D. Dans tous les cas, l'ingénieur responsable de l'étude joue un rôle important afin d'affiner le modèle en fonction de différentes contraintes. La

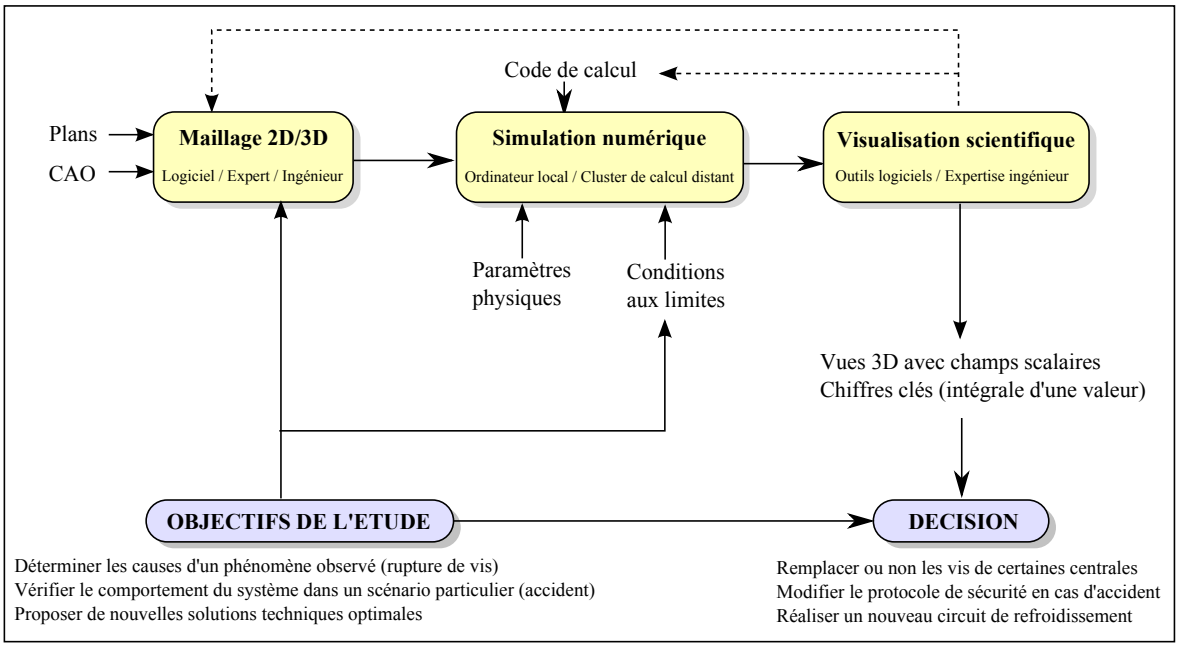


FIG. 1.1: Organigramme classique d'une étude menée à la R&D et comportant une boucle simulation numérique-visualisation scientifique.

figure 1.2 illustre par deux exemples concrets la complexité que peut présenter un maillage servant à une simulation numérique.

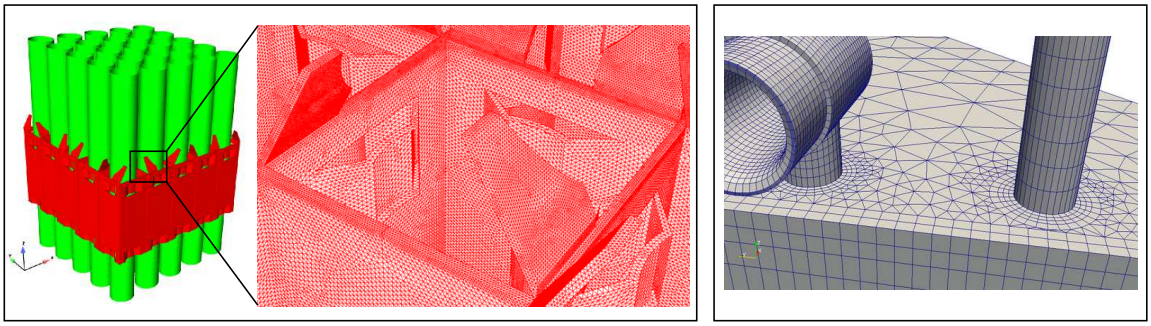
Le nombre d'éléments (mailles) employé définit la finesse de discrétisation du volume. En général, plus il est élevé, plus les résultats obtenus seront précis et fiables, mais plus les temps de calculs seront importants. La résolution du maillage est donc susceptible de varier d'un point à l'autre de l'espace, en fonction de la complexité des phénomènes physiques supposés à l'œuvre localement. Par exemple, pour l'étude de l'écoulement dans un tube simple, les mailles seront plus resserrées le long des parois qu'au centre du cylindre.

De plus, les codes de calculs eux-mêmes orientent la définition et le positionnement des mailles, notamment dans les cas de géométrie assez complexes. Leur stabilité peut en effet exiger l'utilisation de certaines primitives, mais également amener à la modification locale de la surface de bord d'éléments très anguleux. Ici, l'ingénieur apporte son expertise relative à la physique mais également au fonctionnement des codes de calcul. L'essentiel est d'obtenir une discrétisation spatiale qui garantisse ensuite une simulation la plus plausible possible.

L'élaboration de maillages 3D non réguliers nécessite une visualisation interactive posant de nombreuses difficultés, notamment de perception spatiale au sein de rendus denses de lignes. Dans le cas où la géométrie est utilisée pour la première fois, il est par ailleurs possible de réaliser un premier calcul simple dont les résultats sont connus ou attendus. Une visualisation rapide des résultats permet alors de vérifier que les calculs semblent converger raisonnablement sur la discrétisation établie.

1.2 Le temps de la simulation

En plus du maillage surfacique ou volumique, il faut fixer l'ensemble des paramètres de la simulation : constantes physiques des matériaux, conditions initiales, conditions aux limites pendant le temps de la simulation. Ces paramètres peuvent être connus précisément (viscosité du fluide utilisé), définis par les besoins de l'étude (étude d'un choc thermique), ou résulter des calculs réalisés sur une structure voisine. Dans ce dernier cas, les résultats peuvent être définis a priori comme entrée inva-



(a) (b)

FIG. 1.2: Exemples de maillages volumiques complexes. (a) La structure à mailler peut présenter une géométrie peu épaisse avec beaucoup d'angles, ce qui nécessite d'utiliser un maillage très fin. (b) Afin d'optimiser l'utilisation d'une ressource nécessairement limitée, à savoir le nombre d'éléments constituant le maillage, la taille des éléments de base peut varier d'un point à l'autre du volume, et plusieurs primitives différentes peuvent être utilisées pour s'adapter à la géométrie du problème (ici : grille non structurée d'hexaèdres et de pentaèdres).

riante du système, ou au contraire être également influencés par les calculs en cours. Par exemple, si l'on étudie le refroidissement produit par un système fluide, la structure solide entourant la cavité fluide doit aussi être étudiée, la température à l'interface des deux dépendant de l'évolution parallèle des deux systèmes. Il est donc nécessaire de réaliser deux simulations liées au cours du temps —on parle de *couplage de codes* —et parfois de visualiser conjointement les résultats obtenus.

Le temps requis par la simulation dépend de la complexité du problème (nombre de mailles), du nombre de pas de temps requis pour que les calculs convergent (discrétisation temporelle imposée par la physique étudiée, temps de convergence pour dépasser les régimes transitoires) ainsi que des ressources matérielles à disposition (ordinateur individuel, supercalculateur, grappe de PC). Dans la pratique, il n'est pas rare qu'une simulation dure plusieurs heures, voire plusieurs jours. Il est parfois possible d'introduire dans les calculs une variable dont l'évolution permet de contrôler la convergence de la simulation. Dans ce cas, une visualisation rapide des résultats intermédiaires peut permettre de connaître l'avancement des calculs, afin de ne mettre un terme à la simulation que quand celle-ci a atteint le stade souhaité.

1.3 Visualisation des résultats

Une fois la simulation aboutie, l'ingénieur analyse les résultats obtenus en vue de comprendre au mieux le phénomène étudié. C'est là qu'intervient la visualisation scientifique proprement dite, en proposant un ensemble de représentations 2D et 3D permettant d'expliciter visuellement les niveaux des différentes variables dans l'ensemble du volume, leurs variations au cours du temps, et éventuellement les incertitudes quant à leur valeur exacte.

Il s'agit là d'une *visualisation exploratoire*, au sens où l'ingénieur ne sait a priori pas par avance ce qu'il va découvrir, même si son expertise relative à l'étude permet de guider fortement l'analyse. La visualisation doit permettre de mettre en évidence certains points critiques ou des tendances générales (pressions localement plus importantes le long d'une paroi, existence de tourbillons), et si possible d'en déterminer les causes. Ceci requiert des outils souples et interactifs permettant de naviguer facilement dans l'espace de la visualisation, constitué de la géométrie 2D ou 3D et des champs scalaires, vectoriels ou tensoriels calculés.

Les structures utilisées pour représenter les données à l'écran dépendent de la nature des variables,

du domaine physique sur lequel elles reposent, et de l'étendue de la visualisation réalisée. Un champ de température ou de pression peut ainsi être vu sur la surface (*peau*) du maillage, sur un *plan de coupe* positionné autour d'une région intéressante ou au moyen d'une représentation transparente étendue sur l'ensemble de la 3D (*rendu volumique*). Il est aussi possible d'extraire des *isosurfaces*, qui correspondent aux régions de l'espace où la valeur de la variable est identique. Pour un champ vectoriel (vitesse du fluide par exemple), des glyphes (flèches, cônes...) permettent d'en visualiser l'orientation et l'amplitude locales, mais il est également possible d'extraire une information plus globale comme les lignes de courant, celles-ci étant caractérisées par leur tangence au champ en tout point. Ces exemples, illustrés sur la figure 1.3, n'épuisent pas l'ensemble des représentations existantes, mais ils en illustrent la diversité, sachant qu'une même visualisation peut nécessiter de recourir en même temps ou séquentiellement à différents outils.

Généralement, la visualisation comporte plusieurs écueils. Le premier consiste en une insuffisance de l'analyse qui pourrait faire passer à côté de phénomènes importants. Elle peut être liée à une limite pratique des outils à disposition, qui ne permettent pas de rendre compte de certains résultats clés, mais également à un manque d'interactivité qui freine et finalement handicape une analyse approfondie. A l'autre extrême, une visualisation très riche et détaillée peut donner l'illusion que les résultats obtenus sont exacts, alors que de nombreuses imprécisions résultent de la simulation elle-même (finesse de la discrétisation, connaissance des conditions aux limites, stabilité numérique). Si possible, les résultats dans une étude doivent être confrontés à la réalité (sondes dans les installations, simulations sur maquettes, observations de terrain).

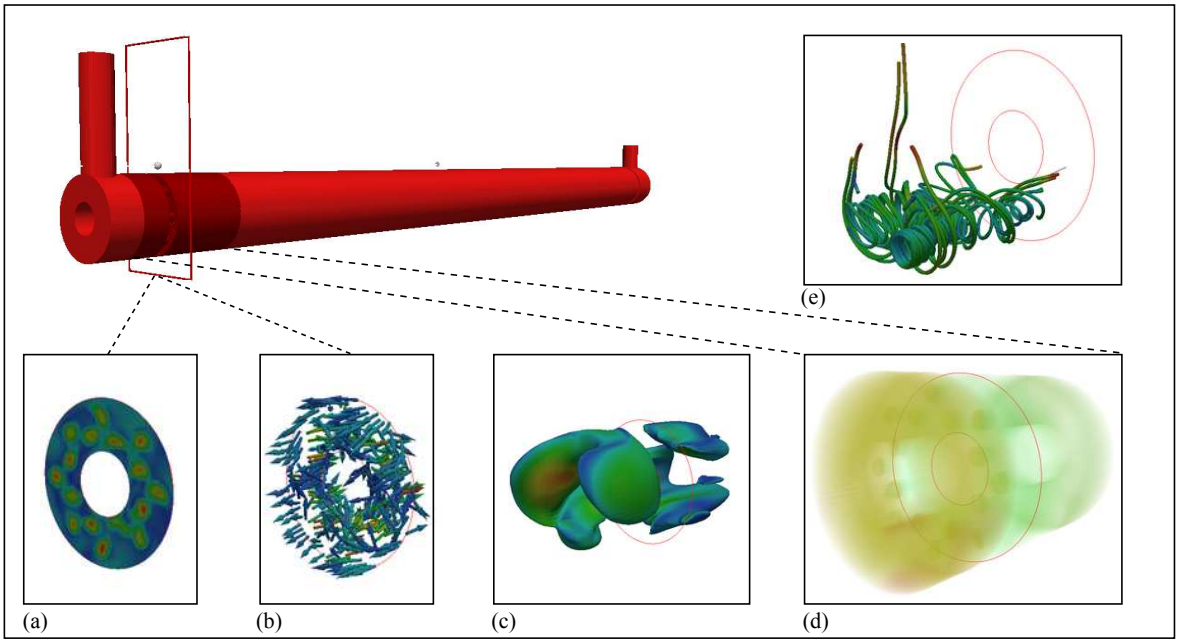


FIG. 1.3: Exemples de représentations tridimensionnelles utilisées en visualisation scientifique. Dans le cas présenté, l'écoulement d'un fluide de refroidissement dans un tube allongé avec des turbulences en entrée, les champs de vitesse (vectoriel), de pression et de viscosité turbulente (scalaires) du fluide ont été calculés. La norme des vitesses est ici visualisée sur un plan de coupe (a) ou sur une isosurface de viscosité (c). Les directions de l'écoulement peuvent être analysées au moyen de flèches sur une coupe (b) ou en traçant des lignes de courant représentées ensuite par des tubes épais (e). Un rendu par accumulation de transparences, appelé rendu volumique direct, permet de représenter l'ensemble du champ de pression sur un sous-volume (d).

1.4 La prise de décision

L'étape de visualisation doit permettre de détecter les principaux phénomènes physiques à l'œuvre dans l'étude, et si possible d'en déterminer les causes. Il faut ensuite présenter ces résultats, soit à des collègues disposant d'une expertise équivalente, soit à des interlocuteurs moins avertis, afin de répondre aux questions posées dans l'étude et de prendre les décisions qui s'imposent. Cette fois-ci, l'ingénieur connaît précisément les informations à faire ressortir. Il s'agit donc principalement de trouver les représentations les plus explicites, que ce soit une vidéo, des captures d'écran d'isosurfaces judicieusement positionnées, voire quelques courbes 2D (profils de température le long d'une paroi), ou un simple scalaire (intégrale de la pression le long d'une surface). Contrairement à la phase exploratoire, l'interactivité n'est pas absolument nécessaire ici, même si elle facilite évidemment la production de résultats.

1.5 Conclusion : place de la visualisation

Il est difficile d'estimer l'importance de la visualisation scientifique dans une étude d'un point de vue quantitatif. Les entretiens réalisés avec des ingénieurs de la R&D (voir section 4.1) situent à moins de 10% le temps passé à visualiser les résultats de calcul. Cependant, comme nous l'avons vu dans ce rapide tour d'horizon, elle est présente à différents moments, et sous des formes variées. D'abord au moment de l'élaboration du maillage, même s'il s'agit d'une forme plus classique de visualisation où l'objet observé est simplement une structure 3D dépourvue de champs numériques. Le but est alors de s'assurer que le maillage construit correspond bien au résultat souhaité, le problème principal est celui de la perception de la forme et de l'espace. Puis au cours de la simulation quelques résultats intermédiaires peuvent être analysés, afin de s'assurer de la convergence des calculs. A ce stade, l'exploration est généralement simple et rapide. La phase principale de visualisation, visant à comprendre en profondeur les résultats de la simulation, est la plus exigeante. Des représentations variées peuvent être extraites sur la base des champs physiques calculés, ces représentations étant généralement affichées sur un même écran. Pour faciliter ce travail, l'utilisateur doit pouvoir naviguer le plus interactivement possible dans cet ensemble multidimensionnel, l'exploration étant partie intégrante du processus de compréhension. Enfin, une dernière étape consiste à produire les images ou les chiffres les plus parlants, afin de faciliter la prise de décision ultérieure. Eventuellement, des techniques plus demandeuses en temps peuvent être mises en œuvre ici, ces images finales, statiques, devant convaincre d'elles-mêmes.

Les travaux présentés dans cette thèse, études et algorithmes de rendu, sont focalisés principalement sur la partie exploratoire de la visualisation.

2 Domaines et échelles

2.1 Domaines physiques

La visualisation réalisée dépend très fortement de la branche de la physique à laquelle la simulation est rattachée. Ainsi, un calcul de thermique comprendra essentiellement des champs scalaires, alors que les phénomènes de convection en jeu en mécanique des fluides nécessitent d'observer également des champs vectoriels (vitesse du fluide), voire tensoriels. Plus subtilement, les modèles de visualisation choisis par un ingénieur peuvent être fortement influencés par sa culture scientifique, c'est-à-dire par les représentations traditionnellement admises dans sa discipline, comme celles que l'on retrouve dans les manuels d'apprentissage. Par exemple, certains mécaniciens sont habitués à utiliser des développées de surface, particulièrement adaptées à la symétrie de révolution des structures

étudiées (figure 1.4). Il existe ainsi des codes visuels, qui rendent parfois l'utilisation de nouveaux outils difficile à faire accepter.

Comme mentionné précédemment, certaines études nécessitent de réaliser des simulations conjointes dans différents domaines physiques, les codes de calcul étant alors couplés. Le dépouillement des résultats peut alors nécessiter de visualiser les domaines en parallèle. Cette juxtaposition de données, notamment en cas de recouvrements dans le domaine physique (frontières communes des domaines par exemple) s'avère particulièrement délicate. Pour mieux comprendre et mettre en relation les différents phénomènes, faut-il représenter les résultats côte à côte, superposés sur une même image, ou séquentiellement, un domaine après l'autre ?

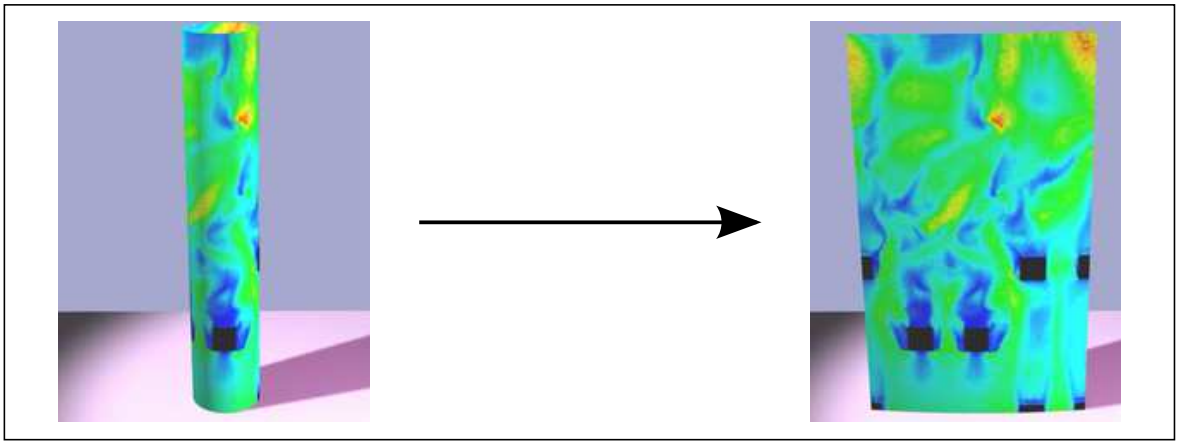


FIG. 1.4: Visualisation par développée de surface. Les types de représentations utilisées dépendent parfois de la culture scientifique de l'ingénieur, des pratiques en vigueur dans sa discipline. Par exemple, certains mécaniciens sont habitués à utiliser des développées de surface, celles-ci étant particulièrement adaptées à la symétrie de révolution des structures étudiées.

2.2 Echelles spatiales

Les phénomènes simulés peuvent présenter des échelles très diverses, de quelques microns pour l'analyse fine du comportement d'un matériau, à plusieurs centaines de mètres lorsqu'il s'agit de comprendre l'écoulement dans la réserve d'eau d'un barrage. Les simulations sont en général réalisées sur un seul ordre de grandeur, mais certains codes de calcul proposent de lier des calculs micro et macroscopiques. Cependant, pour l'essentiel, les visualisations effectuées demeurent pour l'instant concernées par une seule échelle.

En revanche, une visualisation peut faire apparaître des tendances très localisées , comme un extremum de température sur quelques mailles, dont l'explication est à rechercher sur l'ensemble du volume. Il faut alors pouvoir modifier le niveau de zoom facilement, et faire face à deux contraintes opposées : un gros plan sur une petit région permet de visualiser de nombreux détails, mais l'information environnante est perdue ; au contraire, une vue globale permet d'appréhender les tendances générales, mais risque de faire passer à côté de détails importants si ceux-ci ne sont pas mis en évidence (voir section 4.6).

2.3 Primitives géométriques

Comme nous l'avons mentionné dans la section 1.1, la simulation repose sur une discrétisation de l'espace. Un maillage est constitué de primitives 2D ou 3D élémentaires occupant l'ensemble du volume (voir figure 1.5). Le choix de ces éléments de base est en partie dicté par le code de calcul,

qui peut être plus ou moins stable ou rapide suivant les primitives sur lesquelles il est exécuté, ainsi que par les possibilités offertes par le logiciel de maillage. Le plus souvent, le domaine de simulation est constitué de tétraèdres ou d'hexaèdres, qui peuvent être positionnés selon une grille régulière ou de manière non structurée. Cependant, le maillage d'une CAO complexe, présentant par exemple des bords irréguliers, peut nécessiter l'emploi conjugué d'au moins deux primitives différentes (tétraèdres et pentaèdres).

La nature du maillage détermine fortement les types de posttraitement accessibles au moment de la visualisation, et surtout leur efficacité. En effet, de nombreux algorithmes sont développés et optimisés spécifiquement pour des grilles régulières (ensembles ordonnés d'hexaèdres), ce qui correspond au format typique des données médicales de type CT (*Computed Tomography*), mais généralement pas à celui rencontré à EDF, le maillage devant s'adapter à une forme particulière de géométrie (grilles de mélange, vis, conduits d'écoulement...). De plus, une visualisation peut également faire intervenir d'autres éléments, comme des ensembles de points particuliers sur lesquels certains calculs sont réalisés (points de Gauss). Si l'on cherche à proposer de nouvelles méthodes efficaces de visualisation applicables concrètement, il est nécessaire de tenir compte de cette diversité géométrique.

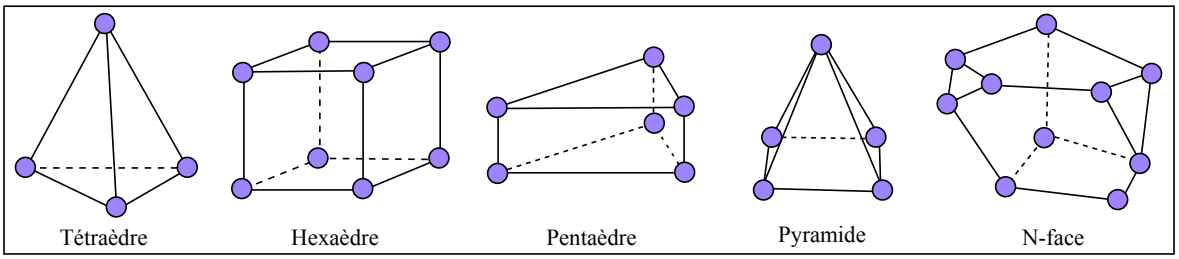


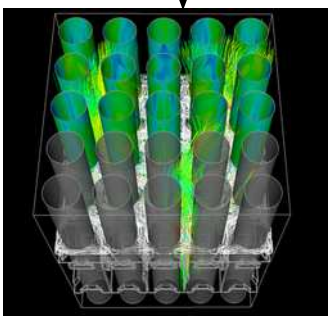
FIG. 1.5: Primitives géométriques 3D utilisées pour discrétiser le domaine d'étude. Le maillage peut comprendre un seul de ces éléments (souvent des tétraèdres ou des hexaèdres) ou plusieurs d'entre eux si la structure à mailler présente une géométrie complexe.

2.4 Grands volumes de données

La caractérisation la plus immédiate de la taille d'une simulation numérique consiste à regarder la place mémoire occupée par les résultats. Pour schématiser, celle-ci dépend de la taille du maillage, de la quantité de champs scalaires, vectoriels et tensoriels pris en compte, ainsi que du nombre de pas de temps enregistrés. Ces trois facteurs déterminent la puissance de calcul et les espaces mémoires nécessaires à la réalisation de la simulation. Ils sont donc ajustés en fonction des besoins propres à l'étude et du matériel à disposition, de manière à ce que la simulation prenne un temps raisonnable (de quelques heures à quelques jours, plus exceptionnellement quelques semaines), ce qui dépend également de la physique étudiée et de la complexité des systèmes d'équations sous-jacents.

Pour donner une idée des tailles de maillages des études les plus poussées à EDF, elles étaient début 2006 de l'ordre du million d'éléments en mécanique des structures, de la dizaine de millions en mécanique des fluides, et pouvaient monter jusque 250 millions pour des calculs de thermique. Deux ans après, ces chiffres ont augmenté d'environ un ordre de grandeur, avec une simulation réalisée avec un code parallélisé sur un milliard de tétraèdres. Si de tels volumes paraissent à première vue très importants, ils ne sont en fait qu'une limitation imposée par les contraintes de simulation. Ainsi, une grille régulière d'un milliard d'éléments ne contient "que" 1000 hexaèdres dans chaque direction de l'espace... En 2007, une simulation d'astrophysique a été menée au CEA sur 70 milliards de particules, 140 milliards de mailles, ce qui illustre les possibilités actuelles. En ce qui concerne le nombre de pas de temps, il peut varier de quelques dizaines, judicieusement espacés, à quelques milliers. Leur

équivalence réelle dépend principalement des temps caractéristiques de la physique sous-jacente. La figure 1.6 illustre sur une étude réalisée à EDF les objectifs à venir en termes de tailles de simulation et de visualisation. Le premier calcul, réalisé et visualisé en 2007, concernait une sous-partie très restreinte de la structure complète et n'était visualisé que sur un pas de temps montrant la solution convergée. Le but est d'étendre ce travail à tout l'ensemble réel, et de pouvoir comprendre l'évolution du phénomène (turbulences et fluctuations de température) au cours du temps.



AUJOURD'HUI	DEMAIN	A TERME
Grille 5*5	Grille 17*17	Structure complète grilles 17*17 10 grilles / assemblage 196 assemblages
100 millions de tétraèdres	220 millions d'hexaèdres	300 milliards d'hexaèdres
15 000 pas de temps	200 000 pas de temps	200 000 pas de temps
Analyse du résultat final	Analyse temporelle	Analyse temporelle
Données: 30 Gb RAM	Données: 300 Gb RAM / pas	Données: 500 Tb RAM / pas

FIG. 1.6: Exemple d'évolution des tailles de simulation à EDF. Un premier calcul de mécanique des fluides et de thermique a été réalisé en 2007 sur une grille de crayons de combustible de taille 5*5 modélisée par un maillage comportant 100 millions de tétraèdres. Le résultat a été visualisé sur un unique pas de temps. L'objectif est d'étendre la simulation à une grille complète (17*17), puis de considérer dans la totalité l'ensemble des grilles et assemblages présents dans le cœur de la centrale. La visualisation devra permettre d'analyser l'évolution du système au cours du temps.

En ce qui concerne la visualisation interactive de ces données, deux problématiques principales sont à dégager. La première se limite à l'exploration d'un unique pas de temps. La taille des données pose déjà de nombreuses difficultés, tant en ce qui concerne leur chargement que leur visualisation fluide. Par ailleurs, notamment pour des systèmes instationnaires ou dont l'évolution est l'objet même de l'étude, il peut être intéressant de comparer l'état du système à différents instants. Le chargement fluide de ces différents pas de temps et surtout la mise au point de transitions rapides et perceptivement claires entre les pas de temps successifs à point de vue constant posent ici de nouvelles difficultés. Dans nos travaux, nous nous limitons au premier aspect, et considérons seulement l'exploration spatiale et non temporelle de champs numériques de grande taille.

3 Moyens matériels et logiciels

D'un point de vue pratique, les outils à disposition constituent la principale limite quant à ce qui est possible ou non dans le cadre d'une étude, la limite basse étant logiquement la plus déterminante. En effet, si l'on peut réaliser un calcul de très grande taille mais que la visualisation se révèle impossible, il sera délicat d'en tirer quelque information utile. Cette situation, encore improbable il y a quelques années, est désormais courante, et explique et justifie pleinement l'essor de la visualisation scientifique comme domaine de recherche à part entière.

3.1 Codes de calcul

Pour l'essentiel, ils sont développés en interne, au sein de la R&D, au moins en partie par les ingénieurs qui les utilisent ou ont travaillé dans le domaine physique correspondant. Les principaux

sont :

- Code_Aster : logiciel de simulation en thermomécanique développé depuis plus de 20 ans, il intègre une grande variété de lois de comportement, différents éléments finis ainsi que des chargements variés. Si de nombreux ingénieurs l'utilisent en interne, il est également diffusé librement, sous licence GNU GPL.
- SYRTHES : code généraliste de thermique pour la résolution des problèmes de conduction et de rayonnement en milieu transparent. Principalement dévolu aux problèmes instationnaires, il nécessite un maillage composé de triangles en 2D, de tétraèdres en 3D.
- Code_Saturne : logiciel généraliste de mécanique des fluides, pour les écoulements incompressibles ou dilatables, avec ou sans turbulences. Les échanges de chaleur sont également gérés, ainsi que certaines physiques particulières (écoulements polyphasiques par exemple). Là encore différents types d'éléments finis peuvent être utilisés, et le code est librement accessible sous licence GNU GPL.
- Salomé : il ne s'agit pas à proprement parler d'un code de calcul, mais plutôt d'une plateforme liant modélisation, supervision des calculs et visualisation. C'est un projet Open Source codéveloppé notamment par EDF et le CEA.

L'internalisation du développement permet d'une part de bénéficier des connaissances-métiers spécifiques, mais aussi de faciliter les améliorations en fonction des besoins. C'est ainsi que SYRTHES a pu être parallélisé récemment pour bénéficier de la puissance de calcul disponible à EDF, ce qui a permis d'achever en 1h30 une simulation portant sur plus d'1 milliard de tétraèdres.

3.2 Calculateurs

Pour pouvoir effectuer des calculs aussi importants que ceux mentionnés plus haut, un ordinateur de bureau est bien évidemment insuffisant. Les plus grosses simulations sont réalisées sur des supercalculateurs, qui sont des machines possédant de nombreux processeurs, d'importantes mémoires pour stocker tous les résultats de calculs ainsi qu'une latence faible et une grande bande passante.

Jusque récemment, les gros calculs mis en place à EDF étaient envoyés au CCRT (Centre de Calcul Recherche et Technologie, au CEA/DAM) qui dispose d'une machine BULL baptisée 'Téra10' (près de 10000 processeurs, 52 Téraflops [1]). Désormais, EDF s'est équipé d'un IBM BlueGene/L (installé en 2007, 8192 processeurs, 18,6 Téraflops) puis d'un IBM BlueGene/P (installé en 2008, 32768 processeurs, 92,960 Téraflops), ce dernier ayant servi au calcul comportant plus de 1 milliard de tétraèdres. Alors que la puissance de calcul est longtemps restée assez concentrée dans certains centres, de telles machines sont aujourd'hui de plus en plus courantes. Si le 'Téra10' du CCRT occupait le 5ème rang mondial en 2006 en termes de puissance de calcul, il se retrouve en juin 2008 au 32ème rang, les BlueGene acquis par EDF figurant respectivement au 103ème et 13ème rangs[2].

Par ailleurs, l'évolution récente des cartes graphiques fournit un nouveau support de calculs hautes performances, avec plusieurs dizaines à plusieurs centaines d'unités de traitements en parallèle disponibles à faible coût. L'utilisation du GPGPU [3] pour des gros calculs commence à peine à être envisagée, par exemple au moyen de l'API CUDA de NVIDIA. Cependant, en raison de la mémoire très limitée disponible sur une carte graphique (au maximum voisine de 1 Go), cette solution convient a priori plus à des calculs sur un jeu de données de petite taille mais comportant de très nombreux pas de temps, ou pour des estimations statistiques (méthode de Monte-Carlo par exemple) [4].

[1] Un Téraflop correspond à 10^{12} opérations en virgule flottante par seconde

[2] source : *top500.org*. Le premier calculateur français est hébergé par l'IDRIS (Institut du développement et des ressources en informatique scientifique). Il occupe en juin 2008 le 9ème rang mondial, avec une puissance de 112 Téraflops.

[3] General-Purpose Computing on Graphics Processing Unit

[4] Le projet de calcul distribué Folding@Home, traitant du repliement des protéines, recensait récemment 1428 Téraflops

3.3 Postes de visualisation

Tant que les données issues des simulations sont de taille raisonnable, elles peuvent être visualisées sur une station de travail ordinaire, possédant typiquement de 1 à 4 Go de RAM et une simple carte graphique. Mais les simulations les plus poussées produisent des masses de données dépassant ces limites, nécessitant plusieurs dizaines de Go de RAM rien que pour pouvoir charger les informations en mémoire. Dans ce cas, des stations de travail comportant 64 Go de mémoire vive sont parfois utilisées. Il faut ensuite que le système graphique permette une visualisation relativement fluide de l'ensemble. Actuellement, l'affichage de dizaines de millions de polygones ne peut être réalisé au moyen d'une seule carte graphique. Il faut alors recourir à une *visualisation parallèle*, par exemple au moyen d'un cluster graphique. Celui-ci comprend plusieurs cartes traitant en parallèle une partie des données à afficher, les images générées par chacune des cartes étant composées pour produire le rendu final. En règle générale, un tel cluster est installé dans une pièce à part, l'utilisateur interagissant depuis son poste en envoyant ses commandes et en recevant en retour l'image calculée (architecture client/serveur). Le facteur bloquant est alors la bande passante, qui doit garantir une faible latence entre l'envoi et la réception d'informations. Des logiciels permettent également à plusieurs utilisateurs d'analyser ensemble le même jeu de données, chacun prenant par exemple la main à tour de rôle (on parle alors de *visualisation collaborative*).

Une solution de visualisation plus poussée consiste à utiliser un mur d'images comprenant plusieurs écrans juxtaposés. En 2008, un mur d'images de 23.9 millions de pixels (6528*3672) a été installé sur le site de Clamart de EDF R&D (figure 1.7). L'affichage est assuré par une grille de 4*4 projecteurs (résolutions 1920*1080) qui reçoivent le résultat de calculs graphiques réalisés sur un cluster composé de 64 nœuds graphiques. Une telle dalle peut être utilisée soit pour un affichage unique de grande taille, soit pour effectuer plusieurs explorations en juxtaposition, ce qui est utile par exemple pour les études multiphysiques. Ce dernier cas introduit un problème perceptif particulier : comment établir une relation visuelle simple entre des éléments présentés côte à côte ? Mentionnons également que le mur d'images intègre un système de vision stéréoscopique, ce qui peut éventuellement permettre de faciliter la perception de la profondeur dans certains cas.

La plupart des visualisations sont effectuées par un utilisateur seul devant son poste. Cependant, si un phénomène surprenant est rencontré, plusieurs experts peuvent se réunir devant l'écran pour essayer de comprendre ensemble ce qui se passe. Une visualisation à plusieurs peut aussi être réalisée dans le cas d'exploration de résultats de calculs multiphysiques, réunissant des experts de plusieurs domaines. Enfin, dans les cas de visualisation collaborative distante, deux utilisateurs peuvent être amenés à observer la même structure. Dans ce cas particulier, l'échange d'information ne peut se faire que par la voix et par un pointage de type souris, ce qui est susceptible de compliquer la transmission du raisonnement entre les deux communicants.

3.4 Logiciels de visualisation

Ils définissent les possibilités et limites effectives de la visualisation, suivant les fonctionnalités intégrées et la facilité à programmer des extensions. Un logiciel de visualisation permet en général d'effectuer la visualisation exploratoire (analyse interactive des données pour la compréhension du phénomène simulé), mais également de produire les captures d'écran, vidéos ou calculs utilisés lors de la présentation finale des résultats. Parfois, il est aussi utilisé pour afficher et vérifier le maillage avant de lancer la simulation.

Pour explorer des données de grande taille, les deux principaux logiciels utilisés à EDF sont *Ensight* et *Paraview*, tous deux permettant une visualisation parallèle. Seul le second, basé sur la

attribués à 12982 cartes graphiques, contre 202 Téraflops liés à 211 978 clients CPU sous Windows.

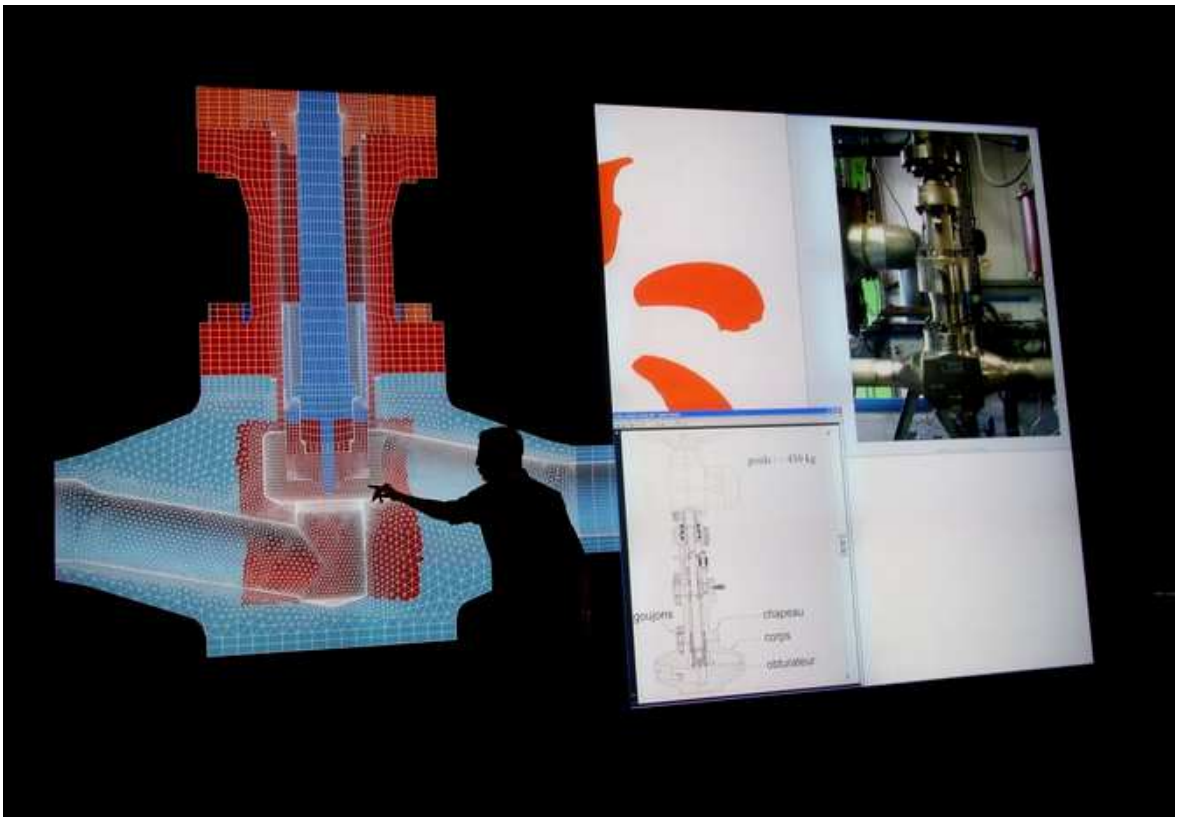


FIG. 1.7: Exemple d'utilisation du mur d'images installé sur le site EDF R&D de Clamart : présentation du maillage support du calcul thermomécanique.

bibliothèque open source VTK (*Visualization Toolkit* [SML06]), peut être utilisé pour développer de nouvelles fonctionnalités d'affichage, ce qui peut avoir son importance pour intégrer rapidement de nouveaux algorithmes.

En pratique, le temps consacré à l'apprentissage du logiciel de visualisation par l'ingénieur est nécessairement limité. Les fonctionnalités sont le plus souvent découvertes en fonction des besoins, et l'IHM [5] peut s'avérer prépondérante pour une utilisation optimale. De plus, toutes les fonctionnalités offertes par de tels outils ne sont pas nécessairement exploitées, certains métiers se cantonnant aux représentations traditionnellement utilisées dans leur domaine.

4 Pratiques et difficultés

Jusqu'ici, nous avons mentionné des aspects assez généraux de la simulation numérique et de la visualisation scientifique telle qu'elle est pratiquée au sein d'EDF R&D. Nous présentons maintenant une synthèse de problèmes pratiques relevés au cours de rencontres avec des ingénieurs en situation d'exploration visuelle de leurs résultats de calculs. Ces difficultés ne sont pas toutes d'ordre perceptif, elles peuvent simplement trouver leur origine dans l'ergonomie et l'efficacité du logiciel de visualisation, voire dans la complexité intrinsèque de la tâche d'exploration visant à mettre en évidence certains phénomènes.

[5] Interface Homme-Machine

4.1 Contexte des entrevues

Pour une bonne part, les remarques que nous formulons sur la visualisation scientifique résultent d'entretiens menés avec des ingénieur de la R&D d'EDF. Plusieurs rencontres ont été organisées, l'objectif principal étant l'observation des différentes pratiques de visualisation en situation d'exploration des résultats de simulation. Lorsque cela n'était pas directement possible, nous avons cherché malgré tout à comprendre la démarche que l'ingénieur avait suivie pour aboutir à ses conclusions, et les principales difficultés auxquelles il avait pu se heurter. Les cas de posttraitement dont nous avons pu discuter n'épuisent bien sûr pas l'ensemble de la visualisation scientifique telle qu'elle est pratiquée à EDF, mais elle donne déjà un bon aperçu de la « réalité du terrain ».

Une première série d'entretiens a été réalisée avec des ingénieurs spécialisés en mécanique des fluides et en thermique. Les études concernaient essentiellement des systèmes de refroidissement monophasiques modélisés par un maillage 3D (homogénéité ou non du refroidissement, présence de points chauds), mais également l'analyse des échanges de chaleur au niveau d'une toiture (maillage 2D, interfaces multiples) ainsi que la récupération de poussières et particules de fuel non brûlées dans la vapeur sortant d'une centrale. Le plus gros maillage comportait 4 millions de tétraèdres, et le plus petit quelques centaines d'éléments 2D. Pour les plus gros maillages, les ingénieurs soulignaient l'importance de la présence du numéricien au moment de leur élaboration, réalisée parfois par un prestataire extérieur ne possédant pas de connaissances particulières quant aux problématiques de simulation numérique. Les visualisations reposaient essentiellement sur des analyses de vitesses, de températures et de pression dans le fluide, observées sur des coupes ou à la surface du maillage. Des lignes de courant étaient parfois générées pour faciliter la compréhension de la nature de l'écoulement. Dans l'étude sur la récupération de poussières en sortie de centrale, des particules animées permettaient de mieux appréhender les trajectoires de ces déchets en fonction du mouvement de l'air induit par des ailettes dans le tube. Notons qu'une des études menées sur un refroidissement monophasique était menée en appui de simulations sur maquettes. La simulation numérique devait notamment permettre d'expliquer les causes des mesures relevées par les quelques sondes placées sur la maquette expérimentale, et aussi d'aider à optimiser le placement de ces sondes.

Nous avons également rencontré des ingénieurs spécialisés en thermohydromécanique. Une des études observées consistait à déterminer les conditions de rupture des éléments d'un robinet de centrale lors d'un choc thermique chaud-froid. Le profil de température aux frontières du domaine solide étudié est obtenu par une première simulation couplée fluide-thermique. La simulation menée ensuite permet de calculer en chaque point du volume l'évolution au cours du choc thermique du tenseur de contraintes, et de mesurer les déformations se produisant dans le matériau. Les maillages utilisés étaient de taille raisonnable (quelques dizaines de milliers de nœuds), mais la nécessité d'analyse temporelle et le logiciel de visualisation rendaient l'exploration délicate (essentiellement des tracés 2D).

Les visualisations correspondant aux cas précédents étaient réalisées par un ingénieur seul devant son poste de travail. Nous avons également participé à des essais de visualisation collaborative à distance. Dans ce cas, deux ingénieurs se trouvant dans des locaux différents explorent ensemble le même jeu de données, l'affichage étant identique sur les deux postes et les intervenants pouvant interagir à tour de rôle. Dans une telle configuration, l'ingénieur qui a analysé le cas cherche à retranscrire ce qu'il a compris du phénomène à son interlocuteur, et attend éventuellement de celui-ci des avis ou éclaircissements sur les raisons ou conséquences de ce qui ressort de la simulation.

4.2 Perception dans le volume

Souvent, les résultats à analyser ont une extension volumique, ce qui rend plus complexe leur compréhension. En effet, notre perception visuelle du monde repose usuellement sur des surfaces

pour lesquelles notre cerveau est généralement capable de déterminer une forme et par intégration, éventuellement, une étendue. Pour une simple raison d'occultation, nous ne nous trouvons jamais en situation de percevoir la structure complète d'un volume, sauf peut être si celui-ci est partiellement transparent. Si l'on souhaite par exemple accéder visuellement à la constitution interne d'un morceau de roche, le seul moyen d'y parvenir consistera alors à en réaliser des coupes. Cette pratique se retrouve logiquement en visualisation scientifique : afin de comprendre la structure 3D d'un champ, scalaire par exemple, souvent les ingénieurs vont soit réaliser une coupe dans une zone clé de l'espace, soit calculer une isosurface au voisinage d'une valeur particulière, ou simplement afficher le champ sur la peau du domaine. Pour éviter les occultations, le nombre de surfaces affichables simultanément est nécessairement limité, et comprendre la structure du champ demande alors de déplacer les points d'intérêts définissant ces surfaces, que ce soit directement dans le volume ou dans l'espace des données. La reconstruction du volume sous-jacent demande alors d'intégrer mentalement ces surfaces présentées successivement, cette tâche se révélant cognitivement très complexe. Si cette opération n'est pas toujours nécessaire, il faut malgré tout établir des relations spatiales entre le phénomène observé ponctuellement (coupe, isosurface) et le volume environnant (sous-structures caractéristiques du maillage, des vis de fixation par exemple). Certains ingénieurs ont ainsi fait ressortir le besoin de visualiser la peau du maillage par transparence, pour bien situer le phénomène observé.

4.3 Perception de la profondeur et de l'orientation

S'il est souvent difficile d'appréhender la structure spatiale dans une image très dense, ce problème se produit déjà lorsque le nombre d'objets représentés est relativement restreint. Nous l'avons par exemple observé alors qu'un ingénieur cherchait à comprendre l'écoulement du fluide dans une région sinueuse d'un système de refroidissement. Pour cela, il a positionné une coupe dans la région concernée, et le champ de vitesses du fluide était représenté par un ensemble de flèches portées par cette coupe. La peau du maillage était représentée en filaire afin de situer le mouvement du fluide relativement à la structure environnante. Pour comprendre effectivement le mouvement exact du fluide dans cette zone turbulente mais très localisée, l'ingénieur a alors été obligé de tourner autour de la coupe en modifiant le point de vue continûment, tant la représentation statique choisie, avec le maillage filaire superposé, était ambiguë. Dans d'autres cas, des utilisateurs ont fait ressortir l'avantage que présenterait la maîtrise de la direction d'éclairage, notamment afin que les arêtes du maillage soient bien visibles, fournissant ainsi un positionnement clair dans l'espace de l'objet visualisé.

L'expression la plus simple de cette difficulté est l'inversion de perception en profondeur, les éléments se situant en arrière plan étant vus en avant de la scène et réciproquement. A ce moment-là, une rotation de la scène dont l'axe n'est pas parallèle à la direction de vue induit une perception de rotation dans le sens opposé, ce qui ne fera qu'ajouter à la confusion de l'utilisateur.

4.4 Complexité de l'affichage

L'affichage simultané d'une grande quantité d'informations rend difficile la lecture et l'exploration visuelle de l'image. Les caractéristiques importantes que l'ingénieur cherche à analyser se retrouvent noyées au milieu d'un contexte trop saillant, et il devient difficile de focaliser son attention sur l'essentiel. Le problème est ici essentiellement 2D, provoqué par l'encombrement dans l'espace image. Dans ces situations, une solution consiste à représenter complètement les données dans une zone focalisée, et à adopter un autre rendu en dehors (simplification, transparence, grisé) de telle sorte que le contexte et les grandes tendances demeurent visibles mais clairement distincts. Une telle visualisation de type focus+contexte n'est pas toujours facile à mettre en place, et de fait est rarement implémentée dans les logiciels de visualisation, en dehors d'un dessin aux traits de la peau du

maillage (silhouettes, arêtes vives) introduisant le contexte géométrique.

4.5 Interactivité, latence

Le problème le plus direct lié à la visualisation de grands volumes de données réside dans l'ampleur des latences [6] au moment de la visualisation. Notamment, un taux de rafraîchissement assez bas —typiquement moins de cinq images par seconde —rend délicate la navigation dans le volume. L'utilisateur ayant alors tendance à trop zoomer d'un coup, et comme il ne peut voir les images entre son point de départ et d'arrivée il se retrouve très vite perdu dans l'espace des données. Une telle lenteur d'affichage limite également le recours au positionnement interactif des supports de visualisation, que ce soit les plans de coupe ou les sources utilisées pour calculer des lignes de courant.

Dans le cas d'une visualisation collaborative à distance, la latence de l'affichage perturbe fortement la transmission du message d'un interlocuteur à l'autre. Elle induit en effet une rupture involontaire entre le discours et l'image sur laquelle porte l'explication, et oblige parfois à des retours en arrière pour clarifier la situation. Pour que le raisonnement puisse être partagé facilement entre deux ingénieurs ne se trouvant pas face à face, l'interactivité de la visualisation est ainsi primordiale.

Une dernière observation, plutôt surprenante sur le coup, est assez révélatrice quant aux pratiques et besoins d'interactivité des utilisateurs de la visualisation. Une technique d'ombrage améliorant la perception de la forme (calcul d'ambient occlusion) a été présentée à un ingénieur. Si le rendu produit a été jugé intéressant, le temps nécessaire à la production de l'image finale —deux minutes, au moyen d'une accélération matérielle —s'est avéré beaucoup trop important, même s'il s'agit de produire des images claires en vue d'une présentation des résultats. Dans ce cas, la personne interrogée a précisé qu'elle préférait utiliser les outils à sa disposition, bien que limités mais utilisables interactivement, pour produire les vues qu'elle estimait les plus convaincantes. Si cette réponse particulière peut s'expliquer en partie par l'habitude liée à l'utilisation de certains outils, elle illustre malgré tout un point important : pour les numériciens, la production de résultats visuels clairs (courbes, images, vidéos...) fait partie intégrante du processus de visualisation, tant elle dépend de la connaissance acquise lors de la phase d'exploration, et elle requiert en conséquence un degré d'interactivité similaire pour être menée à bien.

4.6 Problèmes d'échelle

Si, en général, les simulations réalisées relèvent d'une même échelle physique (microscopique ou macroscopique), les phénomènes à observer peuvent comporter des tailles assez différentes dans l'affichage, comme l'illustre la figure 1.8. Dans cette étude, un écoulement vertical permet de refroidir la structure solide environnante. Si la norme des vitesses affichée sur l'ensemble de l'enveloppe du maillage permet de comprendre les tendances globales du mouvement du fluide, il faut ensuite s'intéresser à certaines zones beaucoup plus petites (trous pour vis de fixations). Mais quand on zoome sur un de ces éléments, le contexte global disparaît, ce qui rend plus difficile la localisation du phénomène observé en détail —perte de repères spatiaux —ainsi que sa mise en relation avec d'éventuelles causes plus distantes. Cette perte de contexte peut se révéler très préjudiciable à l'exploration, même si le passage d'un niveau de zoom à un autre peut être réalisé interactivement. Notons qu'elle trouve sa source dans la taille limitée de l'écran sur lequel la visualisation est d'ordinaire réalisée. L'emploi, par exemple, d'un mur d'images peut permettre de résoudre au moins en partie ce problème, puisqu'alors toute la structure peut être présentée sur la surface d'affichage, avec une résolution suffisante pour analyser finement des détails de la structure.

[6] au sens informatique du terme, c'est à dire le temps entre l'envoi d'une commande d'affichage et le retour visuel correspondant

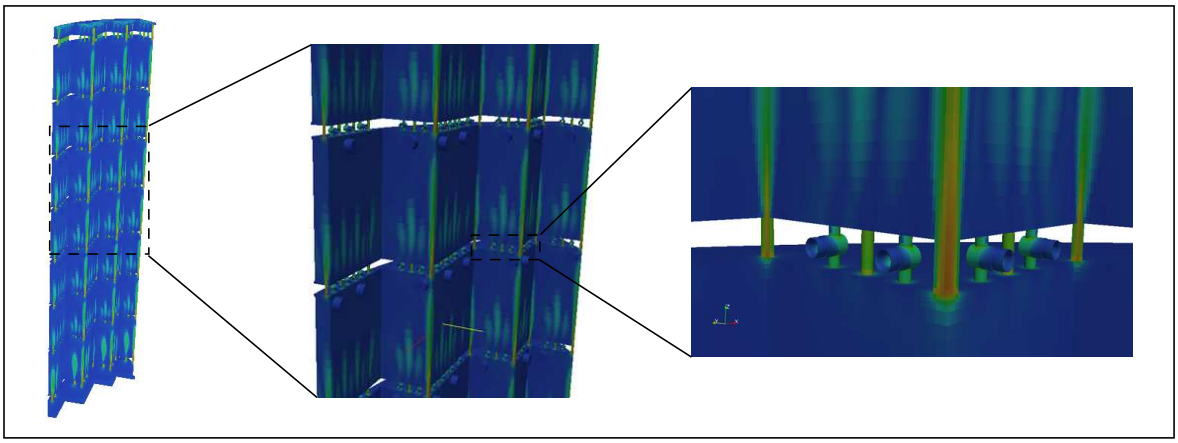


FIG. 1.8: Exemple de difficulté liée à un problème d'échelles spatiales en visualisation. Dans cette étude, on cherche à comprendre les raisons de l'échauffement se produisant au niveau de trous de vis (visibles sur l'image de droite). Pour cela, une simulation de l'écoulement du fluide est réalisée ici dans la structure entourante. L'étude de la nature de l'écoulement nécessite une analyse globale (gauche) ainsi que locale (droite). Dans le premier cas, on perd toute information au niveau des vis, alors qu'un zoom très poussé occulte complètement le contexte et rend plus difficile la localisation dans la structure.

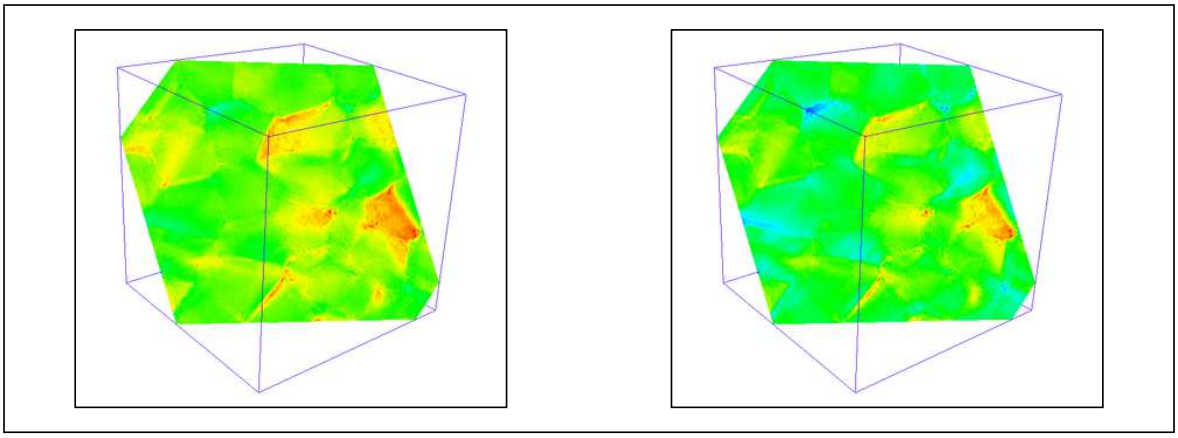


FIG. 1.9: Importance du choix de l'échelle de couleurs pour faire ressortir les extrema du champ. En fonction de l'étalement des données, des échelle linéaires ou logarithmiques peuvent s'avérer plus efficaces pour faire ressortir les principales informations.

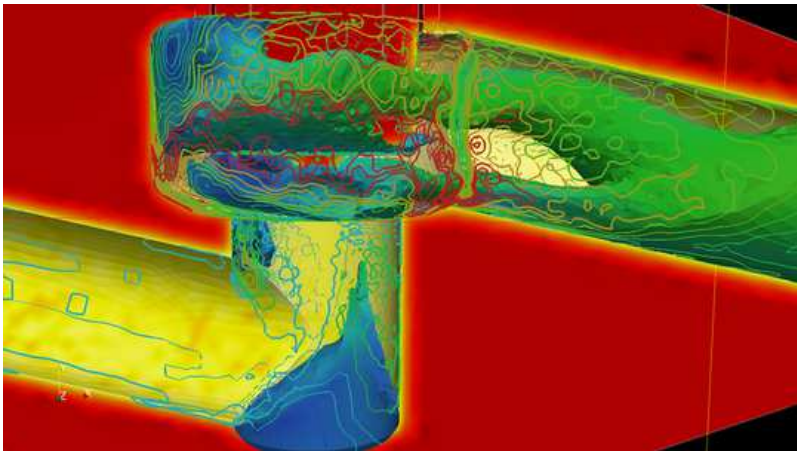
Une autre difficulté réside dans la définition des échelles d'observation des données. Les échelles de couleur, qui attribuent une couleur liée à la valeur du scalaire mesuré, en sont l'exemple le plus classique. Si elles sont mal définies, au regard de la distribution des valeurs prises par le scalaire observé, certains extrema peuvent ne pas apparaître clairement (voir figure 1.9). Une échelle logarithmique ou linéaire tronquée seront parfois plus efficaces pour faire ressortir clairement le phénomène simulé. L'utilisateur doit être en mesure de comprendre et modifier facilement cette carte de couleurs, y apportant sa connaissance de la physique sous-jacente (valeur clé de température à faire ressortir par exemple). D'autres échelles délicates à manipuler peuvent être à l'œuvre, comme la modification de taille de certains glyphes (flèches, point sprites, ellipsoïdes) en fonction d'une variable tierce, afin de visualiser conjointement deux mesures.

5 Conclusion et objectifs

Ce rapide tour d'horizon de la visualisation scientifique telle qu'elle est utilisée à EDF R&D nous a permis d'illustrer la diversité des pratiques en vigueur, ainsi que différentes problématiques liées spécifiquement aux grands volumes de données. Quel que soit le domaine, il ressort que l'on tend vers des données de plus en plus complexes dans leur structure et importantes en volume. Nous avons également vu que les maillages servant de support aux simulations peuvent être de nature très diverses, mélangeant parfois différentes primitives. Ceci a son importance, car de nombreux algorithmes de visualisation sont optimisés spécifiquement pour certaines structures de données, comme des grilles régulières omniprésentes en visualisation médicale. Si l'on souhaite proposer de nouveaux algorithmes de rendu réellement utilisables en mécanique des fluides et en thermique par exemple, cette diversité doit absolument être prise en compte.

A l'heure actuelle, deux grands problèmes concernant la visualisation des ces grands ensembles de données peuvent être dégagés. Le premier concerne la possibilité même d'afficher les données à l'écran, avec une navigation qui soit suffisamment fluide pour ne pas gêner l'exploration. Il faut donc disposer de la puissance de calcul graphique et de quantités de mémoire adéquates, et éventuellement envisager le recours à des structures de données optimisées pour fluidifier le pipeline de rendu. Mais même lorsque les résultats de calcul peuvent être affichés correctement à l'écran, la quantité d'objets projetés et l'extension dans le volume des phénomènes à observer rend difficile la compréhension spatiale de ce qui est représenté. En caricaturant, des visualisations mêlant différents rendus font parfois penser à un tableau de Piranesi dans lequel il est difficile de démêler d'un simple regard l'écheveau des pontons et escaliers arbitrairement déployés dans l'espace (voir figure 1.10). La visualisation scientifique, en tant que discipline, doit donc proposer des outils facilitant la perception des relations spatiales dans les scènes rendues, et si possible sans perte d'interactivité.

Un autre point important concerne la visualisation de l'évolution temporelle des systèmes physiques simulés, nécessitant l'affichage de résultats calculés sur plusieurs pas de temps. Outre le problème matériel lié au chargement fluide de l'ensemble des données, on peut s'interroger sur les représentations facilitant la perception et la compréhension des relations à la fois spatiales et temporelles entre les phénomènes simulés. Bien que ce champ de recherche constitue un défi incontestable pour la visualisation de demain, nous nous contenterons dans cette thèse d'aborder les problèmes perceptifs relatifs à des visualisations sur un seul pas de temps.



(a)

(b)

FIG. 1.10: Difficulté de perception de l'espace dans une scène de visualisation complexe. Lorsque plusieurs objets sont représentés en même temps (a, isosurface de température et isolignes de transfert thermique), il peut être tout aussi délicat d'appréhender leurs relations spatiales et orientations que de comprendre l'enchevêtrement des passerelles dans ce tableau de Piranesi (b, *Carceri d'invenzione, planche VII dite Le pont-levis).*

CHAPITRE 2

Perception visuelle et informatique graphique

> La perception n'est pas le constat d'une réalité objective, elle est la négociation d'une présence au monde.
>
> Derrick de Kerckhove

La perception visuelle est le processus par lequel nous construisons une représentation mentale du monde environnant suite à la stimulation de la rétine par des rayons lumineux. En tant que telle, elle est déterminée non seulement par la nature des capteurs recueillant cette information lumineuse, les yeux, mais aussi par l'ensemble des traitements réalisés par la suite dans le cerveau, ceux-ci pouvant être influencés par une expérience préalable, par exemple la connaissance de la courbure naturelle d'un objet, ou par d'autres informations sensorielles. Ainsi, comprendre la perception visuelle nécessite de s'intéresser aux différents traitements de l'information lumineuse, ce que traduit bien l'aphorisme courant : « La vision ne se résume pas à ce qui touche l'œil. » (*« There's more to vision than meets the eye »*). Par ailleurs, la perception n'est pas un phénomène purement passif, notre action peut avoir une influence sur ce que nous percevons. Par exemple notre mouvement, à travers le déplacement de l'image du monde sur notre rétine, nous apporte de l'information sur sa structure spatiale.

Depuis quelques années, les communautés de l'informatique graphique et de la psychophysique du système visuel humain (SVH) ont amorcé un rapprochement. Les grandes conférences de graphique et de visualisation intègrent de plus en plus de sessions et de cours dédiés à la perception visuelle [1], une conférence et un journal entièrement dédiés à la réunion des deux domaines ont été créés

[1] Par exemple, « Perceptual Issues for Visualization and Evaluation », par Mark Livingstone, IEEE Vis07 (principale conférence de visualisation scientifique)

(APGV[2], TAP[3]). Pour la communauté graphique, la compréhension du fonctionnement du système visuel et de ses limites peut aider à générer des images produisant effectivement l'effet souhaité, et ce le plus rapidement possible. On peut ainsi tirer parti des limites physiques du SVH pour ne pas présenter des détails dont on sait qu'ils ne seront finalement pas perçus, ce qui permet de réduire les temps de calculs. En outre, les connaissances relatives au système visuel humain peuvent guider les développements, en aidant à proposer et à paramétrer des algorithmes adaptés, mais elles permettent également d'éviter certaines impasses, par exemple de travailler sur des rendus transparents complexes présentant une confusion visuelle importante. En retour, les chercheurs en perception visuelle peuvent s'appuyer sur les rendus synthétiques pour mettre au point des stimuli relativement complexes utiles pour étudier certains aspects spécifiques de la vision.

Dans ce chapitre, nous présentons quelques aspects de la perception visuelle et de son étude, et montrons comment ils sont ou peuvent être exploités en informatique graphique et en visualisation scientifique. Nous ne prétendons pas rendre compte de façon détaillée de l'ensemble des connaissances potentiellement utiles en vision humaine, mais plutôt offrir un aperçu rapide pouvant guider certains travaux. Par ailleurs, les aspects de la perception visuelle chez l'homme spécifiques aux problèmes traités au cours de cette thèse, à savoir la perception de la transparence et de la forme par l'ombrage, seront abordés dans les chapitres correspondants (chapitres 3 et 4). Pour commencer, nous nous intéressons aux limites de ce qui peut être effectivement vu et perçu, une telle connaissance permettant de limiter la quantité d'information à visualiser et donc d'accélérer les temps de rendu (section 1). Comme nous l'avons vu dans le chapitre précédent, les visualisations réalisées à EDF peuvent présenter des structures spatiales assez complexes, il est donc nécessaire de s'interroger sur les indices visuels permettant de faciliter la perception de la profondeur et du volume, ainsi que sur le sens du mot interactivité d'un point de vue perceptif et cognitif (section 2). Par ailleurs, l'intérêt de l'utilisation de certains indices visuels ou de nouvelles méthodes de rendus peut parfois être évalué au moyen d'expériences de psychophysique dont nous discutons les avantages et limites en section 3, notamment en considérant la notion d'expertise. En conclusion, nous présentons les orientations des travaux réalisés pendant cette thèse, au regard des informations apportées au long de ce chapitre (section 4).

1 Rendre plus vite : Exploiter les limites du système visuel

En fonction du niveau de précision ou de réalisme désiré et de la quantité de données à afficher, le rendu par ordinateur peut nécessiter des temps de calcul importants, peu compatibles avec une exigence d'interactivité. Cependant, il n'est pas forcément nécessaire de tout représenter, car certaines informations ne seront pas perçues au moment de leur affichage, soit parce qu'elles dépassent la résolution même de ce que l'œil peut saisir (1.1), soit en raison du manque d'attention qui sera porté à ces informations au moment de la visualisation (1.2). La connaissance et la modélisation du système visuel permettent alors de concentrer les ressources computationnelles là où elles sont vraiment nécessaires.

1.1 Limites bas-niveau

1.1.1 Quelques éléments de vision

Dans l'œil (figure 2.1), les rayons lumineux sont réfractés par la cornée et le cristallin, la pupille déterminant la quantité de lumière entrante. Ils atteignent la surface arrière sur laquelle s'étend

[2] ACM SIGGRAPH Symposium on Applied Perception in Graphics and Visualization
[3] ACM Transactions on Applied Perception

la rétine, une couche intégrant deux familles de photorécepteurs, les cônes et bâtonnets. Les cônes (environ huit millions) sont de trois types, sensibles à des plages de longueurs d'onde, et donc de couleurs, différentes. Leur réponse est rapide, mais nécessite une quantité de lumière importante. Les bâtonnets (plus de cent millions) sont sensibles à la luminance, avec une réponse plus lente mais adaptée à des niveaux de luminance plus faibles (vision nocturne par exemple). La répartition des photorécepteurs n'est pas uniforme sur la rétine (figure 2.2 (a)). La densité des cônes est beaucoup plus importante au centre de la rétine, dans une région appelée la *fovéa* dont le rayon d'extension correspond approximativement à 2° autour de l'axe de la vision. Le pic de densité des bâtonnets est un peu plus éloigné, mais on observe également une décroissance de leur concentration à mesure que l'excentricité augmente. Cette répartition des photorécepteurs engendre une baisse de résolution du système visuel à mesure que l'excentricité augmente (vision périphérique), ce qui est schématisé par la figure 2.2 (b).

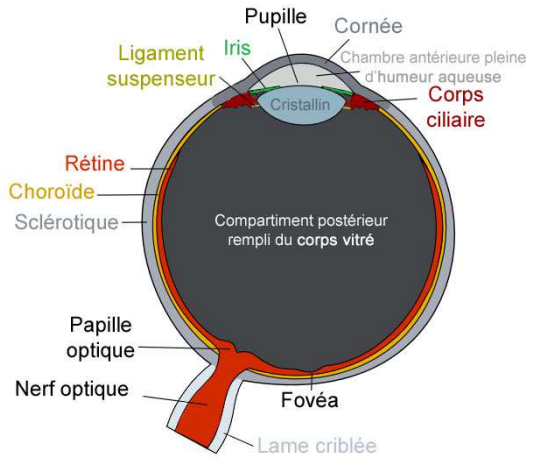


FIG. 2.1: Anatomie de l'œil.

La faible résolution de la vision périphérique est en partie compensée par une exploration du monde visuel réalisée au moyen de saccades et de fixations. Les saccades sont des mouvements très rapides des yeux (plusieurs centaines de degrés par seconde) qui surviennent en moyenne quatre à cinq fois par seconde, et qui sont suivis de pauses (fixations) pendant lesquelles différentes parties du monde visuel peuvent être projetées sur la fovéa. Au cours de ces saccades, une partie importante de l'information visuelle n'est pas disponible (*suppression saccadique*). La figure 2.4 montre un exemple des saccades et fixations réalisées au cours de l'exploration d'un tableau.

Un premier traitement de l'information recueillie par les photorécepteurs est réalisé dans la rétine, qui aboutit au niveau des cellules ganglionnaires à un codage des contrastes. Les champs récepteurs des cellules ganglionnaires présentent des tailles variées, codant globalement pour une plage étendue de fréquences spatiales. Pour chacune d'entre elles, il existe un niveau de contraste au-dessous duquel une image sera perçue comme uniforme. Pour une luminance moyenne donnée, la courbe représentant l'ensemble de ces seuils est appelée fonction de sensibilité au contraste (CSF, *Contrast Sensitivity Functions*, voir figure 2.3). En plus de cette limite existant pour une fréquence spatiale donnée, la superposition dans une image de patterns de fréquences différentes peut engendrer la disparition de certaines informations, ce qui est résumé sous le terme de masquage (*masking*).

Pour une introduction plus poussée à l'anatomie du système visuel et aux problèmes de sensibilité au contraste et de masquage, nous renvoyons par exemple à la lecture de l'article de James Ferwerda [Fer01].

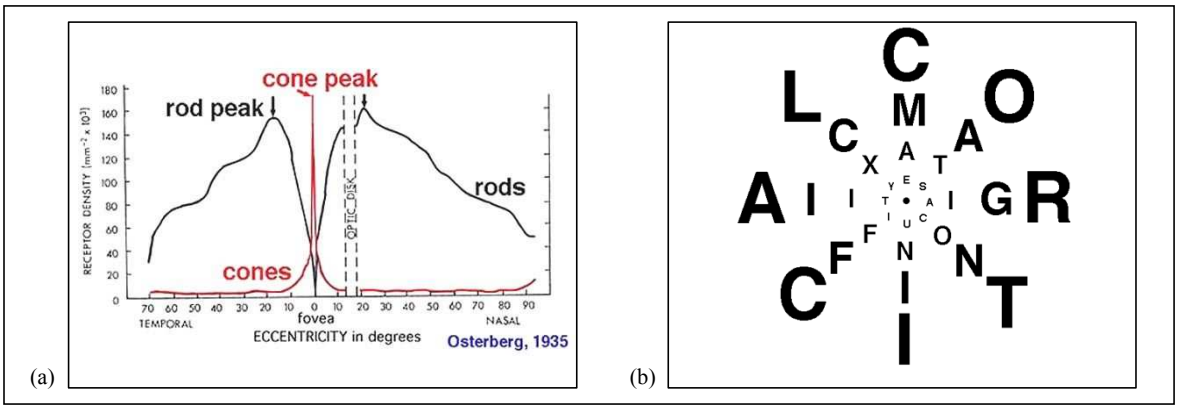


FIG. 2.2: (a) La répartition des capteurs rétiniens n'est pas uniforme mais décroît avec l'excentricité. Le pic central correspond à la fovéa. (b) Ce schéma réalisé par Anstis [Ans74] illustre la variation de l'acuité visuelle en fonction de l'excentricité : lorsque l'on fixe le centre, toutes les lettres présentent la même lisibilité.

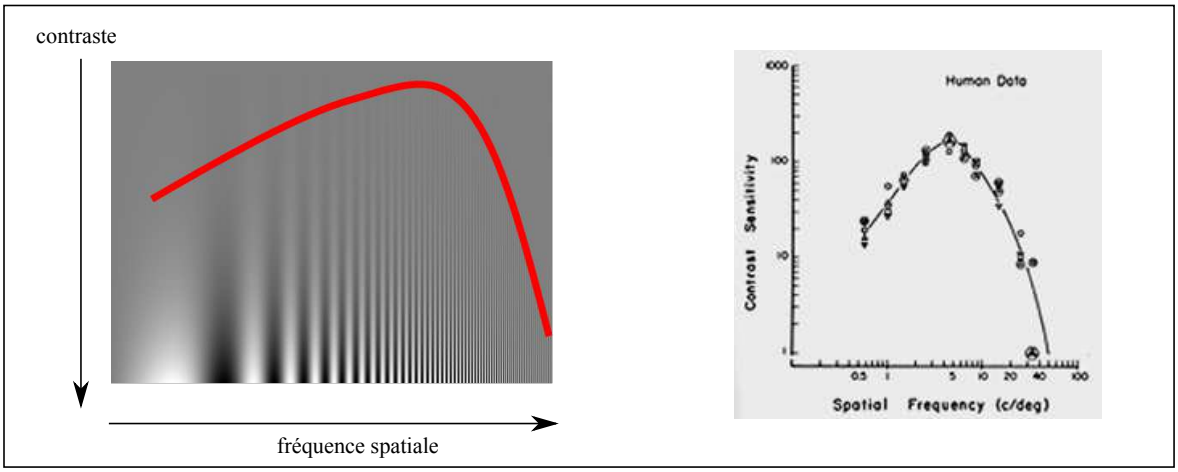


FIG. 2.3: Fonctions de sensibilité au contraste. Pour une luminance moyenne donnée, le seuil de détection d'une différence de luminances (image de gauche : courbe rouge) dépend de la fréquence spatiale du signal lumineux. Image de droite : mesure expérimentale de CSF, d'après DeValois et DeValois [DD90].

1.1.2 CSF et masquage en informatique graphique

Plusieurs auteurs ont proposé des méthodes d'accélération de rendu tenant compte des limites du système visuel humain relatives à la sensibilité au contraste et au phénomène de masquage. Pour cela, des métriques perceptives sont dérivées des fonctions psychophysiques mesurées expérimentalement, et sont ensuite appliquées pour choisir un niveau de simplification du rendu adapté. Une métrique couramment utilisée est le *Visual Differences Predictor* (VDP) de Daly [Dal93] qui repose en premier lieu sur les courbes de sensibilité au contraste.

Ces mesures sont utilisées par exemple pour le rendu en éclairage global, qui est une opération coûteuse. L'exécution de l'algorithme peut alors être arrêtée lorsque la différence entre deux itérations consécutives se situe en dessous du seuil de perception indiqué par la métrique, sur la base des informations contenues dans l'image. C'est l'approche proposée par exemple par [BM98, RPG99]. Yee et al. [YPG01] ont par ailleurs adapté cette métrique dans le cas des animations, en intégrant une dimension temporelle, l'idée étant de limiter les calculs dans des régions dont le mouvement limite la

FIG. 2.4: Exemple d'exploration d'un tableau par saccades oculaires (*Jeune fille en vert*, tableau de Tamara de Lempicka, vers 1927).

quantité d'informations effectivement perceptibles.

L'autre application courante concerne la sélection d'un niveau de détails (LOD, *Level Of Detail*) pour l'affichage de surfaces maillées. En effet, il est possible pour un maillage donné de construire différentes résolutions du même objet, la résolution la plus fine comportant plus de triangles et donc apportant plus de détails visuels, mais nécessitant en retour plus de ressources graphiques pour être affichée. La notion de LOD a été initialement définie par Clark [Cla76], des représentations globalement simplifiées étant utilisées pour les objets plus petits ou plus distants, et étendue par la suite à des représentations progressives de simplification [Hop96]. Funkhouser et Séquin [FS93] ont les premiers suggéré d'utiliser les limites du SVH pour choisir des LOD adaptés pour chaque objet, en tenant compte notamment de leur vitesse et de leur taille. Reddy [Red01] propose un système dans lequel un polygone affiché à l'écran est divisé en sous-unités de façon récursive jusqu'à atteindre le seuil de perceptibilité défini par une métrique d'acuité intégrant sensibilité au contraste, vitesse et excentricité. Luebke et Hallen [LH01] proposent également une méthode de niveaux de détails dépendant de la vue et qui se base sur une métrique perceptive précalculée pour les différents blocs du maillage, les contrastes étant calculés en pire cas pour un ombrage plat de type Gouraud. Ils ajoutent aussi une prise en compte de la préservation des silhouettes. Cette méthode est étendue dans le cas des maillages texturés et éclairés par Williams et al. [WLC+03]. L'inconvénient de ces méthodes est

qu'elles nécessitent de réaliser de nombreux calculs pour sélectionner le niveau de détail approprié, ce qui peut nuire en retour au temps de rendu effectif, et sont souvent beaucoup trop conservatives. Pour résoudre ces problèmes, Drettakis et al. [DBD+07] proposent ainsi une implémentation GPU fonctionnant au niveau de l'image produite, et non pour chaque objet indépendamment, ce qui permet en plus de tenir compte des effets de masquage induits par les interactions entre objets, notamment en raison des ombres projetées. Dans un registre un peu différent, Lavoué [Lav07] propose une mesure de rugosité d'un maillage permettant par la suite de guider sa simplification par des considérations de masquage.

1.1.3 Affichages fonction de l'excentricité

Les GCD (*Gaze Contingent Displays*, affichages dépendant du regard) exploitent la baisse d'acuité en périphérie du champ visuel. Pour cela, ils proposent en général de limiter la quantité d'information à afficher en périphérie, ce qui permet de réduire la quantité de calculs à réaliser dans le cas de rendus de scènes 3D, ou la taille des données à transmettre sur un canal si l'on dégrade une image. Pour cela, ces méthodes nécessitent généralement de recourir à un capteur de regard, appareil permettant de déterminer la position du regard de l'utilisateur sur l'écran avec une précision importante (généralement inférieure à 1° à l'heure actuelle), même si les premières implémentations correspondent à des simulateurs de vols avec casque de réalité virtuelle (dans ce cas, la position du regard est relativement stable au centre des écrans du casque). Pour une vision complète de tels systèmes, nous renvoyons le lecteur aux revues rédigées par [RLMS03, DCM04].

Concernant le rendu d'objets 3D, Reddy [Red01] et Luebke et Hallen [LH01] ajoutent à leur métrique perceptive pour un rendu en multirésolution une prise en compte de l'excentricité, celle-ci déterminant en partie le nombre de triangles effectivement rendus. Plus tôt, Levoy [LW90] avait proposé d'accélérer les méthodes de rendu volumique par lancer de rayons en diminuant la densité des rayons dans l'espace image en fonction de la distance au point de fixation du regard. Pour la visualisation de surfaces triangulées, Murphy et Duchowski [MD07] proposent une méthode hybride. Un lancer de rayon basé sur une structure accélératrice permet de rendre le modèle avec une précision décroissante en fonction de l'angle au centre d'intérêt, alors que des modèles simplifiés de l'objet permettent de préserver les silhouettes. Une étude expérimentale de recherche visuelle réalisée avec cet algorithme met en évidence l'intérêt de cette préservation de contours, et montre aussi que la dégradation d'informations en périphérie facilite la recherche. Ce résultat a priori étonnant peut s'expliquer par le fait que la simplification d'informations trop riches en périphérie limite la distraction du sujet, ce qui constitue un avantage potentiel de ce type de rendu. Lorsqu'il s'agit de dégrader une information purement 2D, Watson et al. [WWHW97] ont montré que la complexité de l'information visuelle (sa résolution) pouvait être réduite d'un facteur deux sans que cela ne gêne la capacité à rechercher une cible dans l'affichage.

Une approche de type GCD peut aussi être utilisée pour simplifier d'autres types de calculs graphiques, qu'il s'agisse par exemple d'utiliser des modèles simplifiés de calculs de collisions entre objets se trouvant en périphérie du champ de vision [OD01] ou de générer une profondeur de champ suivant un modèle exact dans la zone fovéale et approché en périphérie [MvL00, Vil06].

Dans de tels systèmes, une difficulté pratique consiste à rendre non détectable le changement de résolution réalisé lorsque le regard se déplace d'un point à l'autre de l'image. Dans le cas contraire, un mouvement apparent risque d'être perçu par l'utilisateur et d'attirer son attention, ce qui constituerait un défaut du système. Dans une première étude, McConkie et Loschky [ML02] utilisent deux versions d'une même image, une en pleine résolution et une avec du flou. Lors de l'initiation d'une saccade, l'image en pleine résolution est remplacée par sa version dégradée, et réaffichée peu après la fin de la saccade. Dans cette configuration, le changement d'image n'est pas perçu si ce délai de réaffichage

ne dépasse pas 6 ms. Cependant, en utilisant un affichage de type GCD, pour lequel le brouillage de l'image est réalisé progressivement en fonction de l'excentricité, les mêmes auteurs [LM05] trouvent que la mise à jour de l'affichage peut être réalisée jusqu'à 60 ms après la fin de la saccade sans que cette opération ne soit perçue par l'utilisateur. Pour la simplification d'objets 3D, le problème est plus délicat, car le changement de résolution opéré en périphérie est nécessairement discontinu, ce qui peut entraîner un mouvement apparent facilement détectable, surtout si les contours d'un objet ne sont pas ou mal préservés [MD01].

(a) (b)

FIG. 2.5: Exemples d'affichages dépendant du regard (GCD, *Gaze Contingent Displays*). (a) Tiré de Loschky et McConkie [LM05]. Le niveau de flou dans l'image varie radialement à partir d'un point situé en haut à droite de l'image. (b) Le maillage en pleine résolution (1,5 millions de triangles) est rendu seulement là où l'utilisateur regarde, ce qui est matérialisé par l'anneau bleu, et une version simplifiée (15000 triangles) est affichée en périphérie.

1.2 Le problème de l'attention

1.2.1 Introduction à l'attention visuelle

Si la répartition des photorécepteurs dans la rétine et leur sensibilité à la lumière limitent naturellement la quantité d'information effectivement transmise aux aires visuelles, l'attention visuelle —dans ce qui suit on parlera simplement d'attention —constitue un goulot d'étranglement réduisant drastiquement l'information effectivement accessible. L'attention[4] opère une sélection des objets sur lesquels des traitements de plus haut niveau vont être appliqués, le reste de l'information étant en partie ignoré. Elle est symboliquement représentée comme un faisceau de lumière se projetant sur un sous-ensemble de l'information sensorielle, même si une analogie plus exacte serait celle de plusieurs faisceaux opérant en parallèle [5]. En plus d'une focalisation sur certaines régions de l'espace visuel, l'attention possède une résolution spatiale propre [Cav04] indépendante de l'acuité visuelle définie par la répartition des photorécepteurs sur la rétine, la présence de distracteurs autour d'une

[4]La définition la plus classique de l'attention est celle de William James [Jam90] : « Everyone knows what attention is. It is the taking possession by the mind in clear and vivid form, of one out of what seem several simultaneously possible objects or trains of thought... It implies withdrawal from some things in order to deal effectively with others. »

[5]comme le révèlent les expériences de MOT (Multiple Object Tracking), dans lesquelles jusqu'à quatre cibles en mouvement au milieu de distracteurs peuvent être correctement suivies pendant plusieurs secondes (voir par exemple [SP00])

cible située dans le champ visuel périphérique rendant plus difficile voire impossible l'identification de cette cible [6].

L'attention peut être focalisée sur la région fovéale de l'image rétinienne —on parle d'*attention ouverte* —mais également sur certaines parties du champ visuel périphérique —*attention couverte*[7]. Elle peut être attirée par certaines caractéristiques de la stimulation visuelle extraites au cours de traitements préattentifs, par exemple un point plus lumineux ou en mouvement par rapport à l'ensemble de l'image rétinienne (contrôle *bottom-up*), ou dirigée consciemment, en fonction d'un but particulier (contrôle *top-down*).

1.2.2 Exploration visuelle

Dans l'introduction rapide sur le système visuel (section 1.1.1 p.36), nous avons mentionné que nous explorions le monde visuel au moyen de saccades oculaires, des mouvements très rapides des yeux se produisant en moyenne quatre à cinq fois par seconde. Entre autres, les saccades permettent de déplacer la fovéa sur des portions du monde sur lesquelles notre attention se focalise, ce qui permet de bénéficier de la résolution maximale de la rétine. Cette exploration par saccades a été mise en évidence dans les années 1960 par le psychologue russe Alfred Yarbus [Yar67] qui a enregistré les fixations oculaires de sujets regardant des tableaux de scènes naturelles. Une des observations principales relevées par Yarbus concerne le guidage des saccades par la tâche à accomplir. Ainsi, si l'on enregistre les fixations réalisées par une personne en train de regarder une peinture, les trajectoires oculaires relevées dépendent fortement de la question posée au sujet (voir figure 2.6). Récemment, Lipps et Pelz [LP04] ont reproduit l'expérience de Yarbus, sur des périodes d'observation plus courtes et en utilisant un capteur de regard beaucoup moins invasif. Ils trouvent eux aussi un guidage des mouvements des yeux relatifs à la tâche demandée, mais avec une différenciation moins marquée, certains sujets fixant même toujours les mêmes points de l'image.

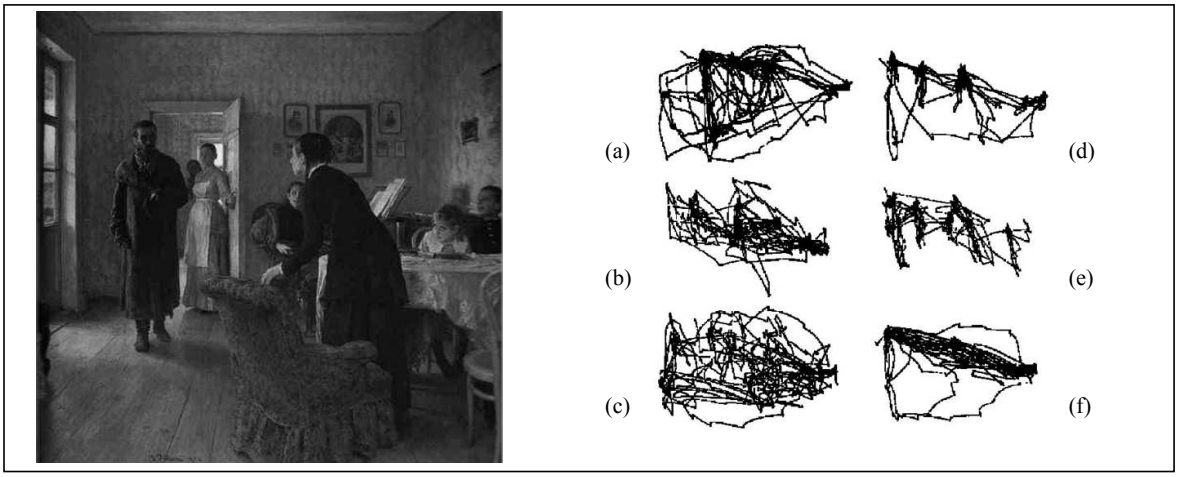


FIG. 2.6: L'exploration visuelle d'un tableau dépend de la tâche à réaliser. Dans cette expérience réalisée par Yarbus [Yar67], les sujets regardent un tableau du peintre russe Ilya Repin (*They did not Expect Him*, 1884-1888), et le trajet de leur regard est enregistré alors que différentes questions leur sont posées. Les explorations reportées sur l'image de droite correspondent aux situations suivantes : (a) Exploration libre. (b) Quelle était l'activité précédente des personnage ? (c) Souvenez-vous des objets. (d) Quel est l'âge des protagonistes ? (e) Souvenez-vous des vêtements des personnages. (f) Combien de temps le visiteur a-t-il été absent ?

[6] effet dit de *crowding*

[7] en Anglais, respectivement, *overt attention* et *covert attention*

1.2.3 Cécité attentionnelle

Si la fonction première de l'attention est de sélectionner l'information à traiter prioritairement, qu'advient-il des objets ignorés volontairement ou non ? Sont-ils malgré tout accessibles à l'expérience consciente ? Reprenant d'anciennes observations, Simons et Chabris [SC99] ont montré qu'il existe une cécité attentionnelle [8] importante, qui se manifeste par l'incapacité à percevoir des objets d'une scène auxquelles on ne prête pas directement attention, même si ceux-ci occupent une portion importante de l'image ou se trouvent dans la région vers laquelle le regard et implicitement l'attention sont momentanément focalisés. Dans une des expériences mises en place, deux équipes, les blancs et les noirs, se passent chacune un ballon, tout en bougeant dans la pièce. Il est demandé au sujet de compter les passes que se font les joueurs blancs, mais pendant cette séquence un gorille passe lentement en arrière plan, croisant même la course du ballon. Dans la moitié des cas, l'intrusion du gorille passe complètement inaperçue. Du point de vue du codage de l'information visuelle, cette cécité ne signifie pas nécessairement que les détails de l'image soient perdus, mais plutôt que si ces détails existent certains mécanismes empêchent d'y accéder [SR05]. En informatique graphique, Cater et al. [CCL02] proposent d'exploiter cette cécité pour accélérer certains rendus, en relation avec le guidage des saccades par la tâche. Ils montrent ainsi que dans une scène synthétique de bureau pour laquelle on demande à des sujets de compter les crayons se trouvant dans des pots, il est possible de restreindre la qualité de l'anticrénelage au voisinage de ces objets sans que cela ne soit perçu par les participants réalisant la tâche de comptage, alors que la différence apparaît clairement lors d'une observation libre des images.

Une autre aspect important de l'attention et de la mémoire visuelle à court terme concerne la cécité au changement [9], qui correspond à l'incapacité à détecter un changement brusque, ou au contraire très lent, s'opérant dans la scène (apparition, disparition ou déplacement d'un élément, changement d'échelle). Une telle cécité peut se produire dans différentes conditions. McConkie et Curry [MC96] ont observé que, dans un stimulus visuel pas trop simpliste, des changements de taille ou l'apparition ou disparition d'un élément de la scène ont une probabilité importante de ne pas être perçus s'ils sont réalisés pendant une saccade. Rensink et al. [ROJ97] utilisent un protocole de *flicker*, dans lequel une image vierge est présentée pendant 80 ms entre deux images dans lesquelles un élément de taille importante est déplacé. Ils observent que plusieurs répétitions d'un tel affichage sont en général nécessaires avant que le déplacement ne puisse être détecté. Dans l'expérience d'O'Reagan et al. [ORC99], l'image vierge est remplacée par la version modifiée de l'image initiale dans laquelle de petites taches sont introduites, mais ne recouvrent pas la zone de changement. Après 80 ms, les taches sont enlevées, mais les sujets ne remarquent que rarement le changement intervenu.

En théorie, la cécité au changement pourrait être utilisée pour masquer les changements brusques survenant dans l'affichage de scènes synthétiques lors d'un changement de LOD notamment. O'Sullivan [ODH03] proposent ainsi de changer les niveaux de résolution dans un jeu vidéo pendant les saccades, et Schumacher et al. [SAH04] montrent que, dans un dispositif de réalité virtuelle à écran large, un déplacement de la scène n'est pas ou peu perçu s'il est effectué pendant une saccade. L'utilisation effective de *flickers* ou de distracteurs (taches de [ORC99]) paraît plus délicate à mettre en œuvre, même si elle ne nécessite pas de recourir à un capteur de regard, car elle risque non seulement de masquer les changements réalisés, mais également de distraire l'attention de la personne regardant l'affichage. Dans tous les cas, ces résultats doivent alerter sur un point important : il n'est pas toujours suffisant d'afficher une information de manière apparemment claire pour que celle-ci soit effectivement perceptible, et les phénomènes de scintillement d'un affichage peuvent entraîner une relative cécité attentionnelle. Cela dépend fortement de la tâche réalisée par le sujet regardant, et de

[8] *inattentional blindness*
[9] *change blindness*

l'attention qu'il prête à l'information présentée.

1.2.4 Saillance de l'information visuelle

Dans une image, certaines régions sont plus susceptibles d'attirer l'attention et le regard que d'autres, par exemple lorsqu'elles sont plus claires ou plus contrastées que leur voisinage. Une carte de saillance (*saliency map*) est une représentation codant les positions et saillances potentielles de différentes régions de l'espace. L'établissement d'une telle carte peut permettre de prédire une partie des saccades qui seront réalisées par un sujet regardant librement l'image correspondante, et éventuellement de simplifier l'image ou le rendu dans les régions les moins saillantes.

Itti et al. [IKN98] proposent un modèle direct de calcul de cartes de saillance pour une image statique inspiré de la théorie de l'intégration de caractéristiques de Treisman et Gelade [TG80] (*feature integration theory*). Ce modèle computationnel repose sur le calcul à différentes échelles spatiales de variations d'intensité lumineuse, de couleurs et d'orientation. Les cartes résultantes sont combinées et normalisées d'abord entre échelles puis entre caractéristiques pour donner la carte de saillance complète. Celle-ci est utilisée pour générer une trajectoire oculaire, la succession des saccades étant déterminée par un mécanisme de *winner-take-all* et d'inhibition des zones déjà visitées. Pour rendre compte des mouvements des yeux effectivement relevés lors de la visualisation de scènes complexes, Henderson et al. [HWH99] suggèrent un modèle plus complexe : une première carte est construite à partir d'indices visuels de bas-niveau (luminance, contraste, texture, couleur), avec de plus un biais de proximité favorisant l'exploration de régions voisines. Au fur et à mesure de l'exploration de l'image, la carte de saillance encode également des aspects plus haut-niveau, comme la sémantique associée aux différentes régions ainsi que des effets de mémoire. Dans un tel modèle, les mouvements des yeux sont donc guidés en premier lieu par des indices purement visuels, ensuite par des informations plutôt cognitives, une fois que chaque région a été explorée. Une autre définition des cartes de saillance, baptisée *Contextual guidance model* [TOCH06], fusionne dans un cadre bayésien des informations globales de contexte et des caractéristiques bas-niveau. Tous ces modèles sont limités à l'exploration visuelle d'images statiques. Dans une étude expérimentale enregistrant le regard de sujets jouant à un jeu vidéo, Peters et Itti [PI08] montrent que dans ce cas dynamique une heuristique basée sur les mouvements continus des objets et leur apparition soudaine explique l'essentiel des saccades réalisées.

Les cartes de saillance d'une image peuvent être utilisées pour simplifier et accélérer le rendu dans les zones moins susceptibles d'attirer l'attention. Par exemple, Sundstedt et al. [SDL+05] utilisent une carte calculée sur une version simplifiée de l'image finale pour diminuer les ressources de calcul d'éclairage global dans les régions peu importantes. Une étude expérimentale montre que les différences avec une image rendue complètement passent inaperçues. DeCarlo et Santella [DS02] utilisent une mesure directe de saillance, à savoir les fixations enregistrées pour différents sujets sur une image, pour guider un algorithme de stylisation et d'abstraction de photographies préservant plus d'informations dans les zones où le regard s'est le plus attardé.

Dans tout ce qui précède, la saillance est définie et calculée sur une image 2D, ce qui correspond dans le cas de visualisations 3D à un point de vue unique. Lee et al. [LVJ05] transposent l'idée du modèle de [IKN98] à des maillages 3D, la carte de saillance étant calculée sur l'ensemble de la surface par accumulation de différences de courbures entre un point et son voisinage à différentes échelles spatiales. Le résultat est utilisé pour guider un algorithme de simplification de maillage (QSlim, [GH97]). Howlett et al. [HHO05] calculent la carte de saillance du modèle 3D en accumulant les fixations de plusieurs sujets, et simplifient le maillage sur la base de cette information similairement à [LVJ05]. Ils montrent par ailleurs qu'à niveau de simplification équivalent, il est plus facile de reconnaître les maillages simplifiés au moyen de cette mesure perceptive que sans.

Inversement, l'idée de saillance peut être utilisée par un rendu pour dégrader une partie de l'in-

formation visuelle de manière à attirer le regard, et avec lui l'attention ouverte, vers des régions particulières de l'image, tout en préservant une représentation suffisamment informative du contexte environnant. Cela est utilisé notamment dans des représentations de type Focus+Contexte (F+C), par exemple en appliquant un flou de profondeur [KMH+02] ou une simplification stylisée [CDF+06] à certaines parties de l'image. Dans le cas du rendu volumique, Kim et Varshney [KV06] proposent une fonction inverse permettant d'augmenter la saillance visuelle de certains voxels en modulant leur couleur et luminance sur la base d'informations choisies par l'utilisateur ou dépendant de certaines caractéristiques du champ scalaire sous-jacent. L'objectif est de faciliter l'exploration et la compréhension des données visualisées en attirant l'attention sur les valeurs ou sous-ensembles d'intérêt.

1.2.5 Limites des modèles de saillance : le regard de l'expert

Comme nous l'avons vu, l'exploration visuelle d'une image peut être guidée en partie par des caractéristiques préattentives, comme des extrema locaux de contraste, de couleur et d'orientation, ainsi que par le but poursuivi par le sujet regardant. Cependant, un autre facteur peut avoir son importance : il s'agit de l'expertise du sujet vis-à-vis du type d'image observé, voire de son habitude simplement à étudier et analyser des images, ce qui caractérise manifestement le public à qui s'adresse notre travail, à savoir des physiciens entraînés à explorer visuellement des résultats de simulation numérique. Par exemple, dans le cas du jeu d'échecs, il a été observé que les premières fixations réalisées par un joueur expérimenté dans une position de milieu de jeu correspondent à des positions clés, et les saccades sont effectuées entre des pièces ayant un lien direct entre elles ([Get96], p.33).

Vogt et Magnussen [VM07] ont mené une étude dans laquelle ils mesurent les fixations oculaires d'artistes (enseignants ou étudiants confirmés d'une école d'art) et de personnes non entraînées (« homme de la rue » ne possédant pas d'expertise picturale) observant des photos plus ou moins esthétiques, ainsi que des représentations abstraites. Au cours d'une première session, les seize images sont vues librement pendant 40 s chacune, alors que dans la deuxième session les participants ont pour instruction de mémoriser le plus d'éléments possibles. Les auteurs ont observé des trajets oculaires très différents pour les deux groupes, notamment lors de l'exploration de représentations abstraites, et des variations opposées d'une session à l'autre. Ainsi, les participants non entraînés présentent en moyenne plus de fixation que les artistes lorsqu'ils voient les images pour la première fois, et ce nombre décline lors de la deuxième présentation, ce qui est attribué à un effet de mémoire déjà observé expérimentalement (*Eye-Movement-Based Memory Effect*, [AC99], [HWH99]). Au contraire, le nombre de fixations mesurées pour les artistes augmente au cours de la deuxième session, pour dépasser celui des non experts.

Si les modèles de simplification reposant sur des cartes de saillance ou sur la tâche à réaliser peuvent être appliqués pour des visualisations courantes destinées à l'homme de la rue, il n'est pas certain qu'ils puissent être appliqués avec succès et sans perte importante d'information dans le cas de visualisations techniques réalisées par un expert. Anticiper l'exploration visuelle d'un expert est une tâche plus complexe, qui demanderait notamment de pouvoir caractériser la manière dont cette expertise influe sur les autres traitements perceptifs et cognitifs.

2 Rendre mieux : Perception de la forme et de la profondeur, interactivité

2.1 La perception de l'espace

2.1.1 Forme, profondeur et volume

La perception de l'espace est par nature un problème difficile à résoudre pour le système visuel, dans la mesure où elle correspond à la projection d'un monde tridimensionnel sur un espace essentiellement 2D, la rétine. Par perception de l'espace, nous entendons l'ensemble des caractéristiques qui définissent la spatialisation des objets perçus : leur distance à l'observateur et leurs positions relatives, leur forme définie par l'orientation locale de leur surface, et leur volume, c'est-à-dire leur extension en profondeur. Dans le cas de la visualisation de scènes 3D générées par ordinateur, la tâche est rendue encore plus complexe par une projection préalable sur un support 2D, l'écran, qui limite la quantité d'informations disponibles pour construire une représentation mentale de l'espace observé. Certains afficheurs volumétriques commencent à être développés, mais leur résolution et leur plus-value demeurent limitées [10].

Pour attribuer des caractéristiques spatiales aux objets projetés sur la rétine, le SVH peut potentiellement s'appuyer sur un nombre important d'indices visuels et visuomoteurs, leur combinaison pouvant faciliter la résolution d'ambiguïtés inhérentes à ce problème inverse. Des listes plus ou moins exhaustives de ces différents indices et de leur importance relative, ainsi que des discussions poussées sur l'extraction des informations de surface et de profondeur par le SVH peuvent être trouvées dans la littérature [CV95, Pal99][FA03, Tod04].

Dans ce qui suit, nous proposons un aperçu succinct des indices rencontrés dans des images statiques ou induits par les caractéristiques du système visuel, le mouvement ainsi que les connaissances projetées par l'observateur [11]. Nous discutons la manière dont ces différentes informations peuvent se compléter ou interférer, et nous abordons le cas problématique de la visualisation dans le volume.

2.1.2 Indices picturaux

Les indices picturaux correspondent à ceux que l'on retrouve dans une image statique, telle un tableau, une photographie ou une image de synthèse. Ils ont été étudiés et détaillés aussi bien par les artistes que les neuroscientifiques, et correspondent à une perception monoculaire, même s'ils participent bien sûr à la construction d'un percept 3D lors d'une observation avec les deux yeux. Ces indices sont résumés sur la figure 2.7.

Perspective linéaire La perspective linéaire représente l'une des principales découvertes artistiques de la Renaissance, les peintres ayant alors cherché à définir mathématiquement les règles de projection de points de l'espace 3D sur un plan 2D. Notons que ces transformations, comme la matrice de projection perspective classiquement utilisée en infographie, ne constituent qu'une approximation de la projection effective du monde sur la rétine.

La perspective linéaire induit notamment la convergence dans l'image des lignes parallèles dans l'espace 3D, la réduction de la taille des objets en fonction de leur éloignement et des gradients de texture. Ces informations sont parmi les indices les plus puissants pour déterminer les distances et

[10] Par exemple, le *Percepta Display* de la société *Actuality Systems* affichant des données dans un volume au moyen d'un disque en rotation rapide [GB06].

[11] Il ne s'agit pas de proposer une taxonomie précise, celle-ci variant d'un auteur à l'autre. Par exemple, [Bor42] parle d'indices primaires (convergence, accommodation, disparité rétinienne) et d'indices secondaires (occultation, taille, perspective, ombrage, parallaxe de mouvement...).

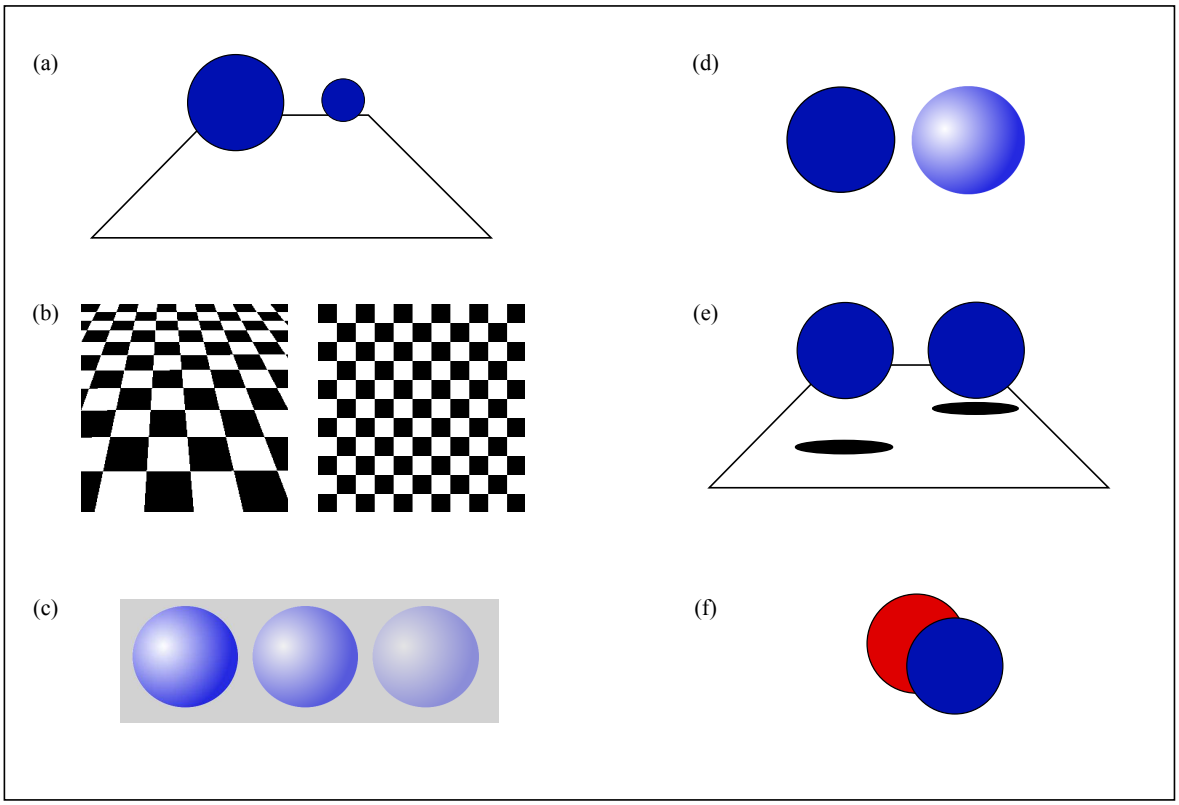


FIG. 2.7: Indices picturaux de perception de la forme et de la profondeur. (a) Variation de la taille d'un objet en fonction de sa distance à l'œil. (b) Gradient de texture. (c) Perspective atmosphérique. (d) Ombrage de surface. (e) Ombres portées. (f) Occultation.

tailles relatives, ainsi que les orientations de surfaces planes. En revanche, elles ne permettent pas d'accéder à la taille absolue d'un objet, ce qui est le cas possible avec une projection orthographique.

En image de synthèse, l'utilisation de la perspective est des plus courantes, car elle est totalement intégrée dans la chaîne de rendu. Cependant, il faut souligner que la perception correcte d'une scène rendue par perspective linéaire est soumise à certaines hypothèses, notamment la régularité des éléments perçus. Si ces hypothèses ne sont pas vérifiées, l'interprétation de la scène peut être erronée, surtout en l'absence d'autres indices. Par exemple, dans un affichage de visualisation scientifique, l'emploi de points dont la taille varie suivant la valeur scalaire du champ sous-jacent peut localement contredire l'hypothèse de variation de la taille avec la distance, et faire percevoir comme distants des points en réalité proches, mais de petite taille. A contrario, deux éléments de même taille dans l'image mais perçus à des distances différentes vont apparaître comme de taille réelle différente.

Perspective aérienne Parfois également appelée perspective atmosphérique, elle correspond à une diminution des contrastes et à un bleuissement des objets les plus distants, résultant du trajet des rayons lumineux dans l'atmosphère. Elle a été décrite en détail par Léonard de Vinci, même si elle était pratiquée dans les ateliers de la Renaissance italienne avant lui ([Ara03], p. 303). Si son effet est normalement visible seulement pour des distances importantes (plus de cent mètres), un effet de brouillard mimant ses caractéristiques principales est en général implémenté efficacement et facilement paramétrable dans les librairies de programmation graphique (par exemple *GL_FOG* dans l'API OpenGL®).

Ombrage L'éclairage d'une scène produit des gradients de luminance qui dépendent notamment de la réflectance des surfaces et de leur orientation relative à la source de lumière et à l'observateur [KvD03]. Le système visuel humain est capable d'utiliser cette information d'ombrage pour reconstruire localement la géométrie de la surface, notamment son orientation et sa courbure. Cependant, l'ambiguïté du bas-relief démontre théoriquement que dans le cas d'un éclairage directionnel cette inférence ne peut être réalisée qu'à une transformation affine près, une même image pouvant correspondre à une infinité de couples surface/direction d'éclairage. En mesurant le relief perçu par des sujets observant un objet photographié, Battu et al. [BKK07] remarquent que la variabilité des réponses obtenues peut effectivement être réduite si l'on considère le relief pictural [12] comme défini à une transformation affine près. Par ailleurs, la perception de la forme par l'ombrage est influencée par certaines hypothèses du système visuel, celui-ci supposant par exemple un éclairage venant du haut et de la gauche (voir section 3.1.1).

En plus de l'ombrage, la présence d'un objet entre la source de lumière principale et une surface produit sur cette dernière une ombre portée. Celle-ci peut être utilisée pour déterminer la distance et la taille de cet objet [KKMB96], la direction de l'éclairage devant également être connue pour lever les ambiguïtés potentielles.

Les problèmes liés à l'extraction d'informations spatiales par l'ombrage ainsi que les algorithmes d'ombrage existant en image de synthèse seront abordés plus en détails dans le chapitre 4.

Occultation Le recouvrement partiel d'un objet par un autre dans l'espace image constitue un des indices les plus forts pour inférer l'ordre relatif en profondeur de deux objets, tout en posant une limite forte à ce qui peut être effectivement vu depuis un point de vue fixe en l'absence de transparence. Cette occultation, ou interposition, se traduit dans l'image par des terminaisons en T dont la détection est assurée par des cellules spécialisées du système visuel. Comme pour la projection perspective, l'occultation est implémentée en standard sur les cartes graphiques, au moyen du *z-buffer*.

2.1.3 Indices oculomoteurs et binoculaires

L'accommodation correspond à l'ajustement de la puissance optique de l'œil, ce qui permet d'assurer une mise au point sur l'objet focalisé. Elle peut fournir directement une estimation de profondeur, mais celle-ci est relativement faible et se limite aux objets proches. En vision monoculaire, les objets situés en avant et en arrière du point de focalisation apparaissent flous, ce qui définit un effet de profondeur de champ analogue à celui obtenu avec un appareil photo. D'un point de vue image de synthèse, cet effet de flou de profondeur est relativement coûteux à obtenir, même si des algorithmes approchés utilisant l'accélération de la carte graphique ont été proposés récemment pour le réaliser de manière interactive [KLO06, KS07].

En vision binoculaire, chaque œil fournit une image légèrement différente du monde, la distance interoculaire induisant un petit effet de parallaxe. Lorsque les deux yeux sont focalisés sur le même point de l'espace, l'orientation relative des deux yeux —la vergence —permet d'évaluer directement la distance de ce point. Les deux images gauche et droite sont normalement fusionnées, l'ensemble des points pouvant être fusionnés étant défini par une surface passant par le point de focalisation, l'horoptère. Les points de l'espace ne se situant pas à proximité de cette surface ne peuvent être correctement fusionnés, et la disparité dans les deux images rétiniennes donne normalement lieu à une perception tridimensionnelle [Jul71] (vision stéréoscopique).

Il existe différents systèmes de réalité virtuelle permettant de produire une vision stéréoscopique, par exemple en utilisant des lunettes polarisées ou à ouverture alternée. Un tel système est coûteux en

[12] Le relief pictural correspond à la forme d'une surface dans l'espace pictural, c'est-à-dire l'espace de l'image.

temps de calcul, car il nécessite de produire deux rendus pour la scène, une image par œil. Par ailleurs, il peut entraîner une fatigue importante à l'utilisation, surtout s'il est mal calibré. Une des principales raisons de cette fatigue provient du conflit entre les indices de vergence et d'accommodation. En effet, en conditions normales, la première contrôle la seconde, adaptant la puissance optique en fonction de l'objet focalisé. Mais un système de réalité virtuelle sera généralement constitué d'un écran situé à une distance fixe de l'observateur, sur laquelle l'accommodation doit être réalisée alors que la vergence peut correspondre à des points virtuellement situés devant ou derrière cette surface.

2.1.4 Indices cinématiques

Les indices décrits jusqu'à présent correspondent à la visualisation d'images essentiellement statiques. Lorsqu'un objet est en mouvement relativement à l'observateur, le flux optique généré sur la rétine par son déplacement peut être utilisé pour construire une perception tridimensionnelle. C'est ce que l'on appelle l'effet cinétique de profondeur (KDE, *Kinetic Depth Effect* [WO53]), ou extraction de structure par le mouvement (SfM, *Structure from Motion* [Ull79]), qui se produit même pour des stimuli ne contenant aucun autre indice visuel de profondeur, comme la projection sur un écran de la silhouette d'un cube [WO53] ou de points attachés à la surface d'un cylindre [Gre61]. Généralement, la perception est guidée par une hypothèse de rigidité de l'objet vu en mouvement. Dans le cas d'un objet en rotation, le stimulus peut présenter une bistabilité, le sens de rotation semblant alterner, ce qui induit immédiatement une inversion de la perception en profondeur. On peut notamment observer ce phénomène lorsque l'on fait tourner à la souris une scène rendue par ordinateur, ou lors de mouvements importants de caméra au cinéma ou dans une scène synthétique.

L'extraction de forme peut également être réalisée lorsque le sujet lui-même est en mouvement. Ce déplacement du point de vue de l'observateur introduit un déplacement de l'image du monde sur la rétine, appelé parallaxe de mouvement. Ce mouvement actif s'accompagne de signaux extrarétiniens qui contribuent grandement à lever l'ambiguïté de la perception de profondeur, en faisant intervenir l'hypothèse de stationnarité des objets perçus [WPLD01]. La parallaxe de mouvement constitue un indice de profondeur très important, notamment pour les animaux ne possédant pas de vision stéréoscopique [13]. D'un point de vue pratique, la réalisation de parallaxe de mouvement sur un écran d'ordinateur requiert l'utilisation d'un capteur de position de la tête, la projection utilisée pour le rendu variant alors avec le mouvement de la tête de l'observateur.

2.1.5 Indices cognitifs

La plupart des indices picturaux, voire des indices cinématiques, ne peuvent suffire à extraire une perception tridimensionnelle non ambiguë de l'image du monde projetée sur la rétine. Nous avons vu à plusieurs reprises que le système visuel humain utilise des hypothèses sur ce qui est perçu : régularité des lignes pour la perspective, utilisation de la taille d'un objet familier pour calibrer la profondeur et la taille perçues des objets environnants, hypothèse de rigidité et stationnarité pour l'extraction de la forme par le mouvement... Si ces hypothèses sont mises à mal, il peut résulter une perception erronée, voire aberrante. Ainsi, dans la chambre d'Ames, deux personnes identiques paraissent être de tailles très différentes, car elles se trouvent à deux extrémités différentes d'une pièce trapézoïdale qui paraît en réalité cubique, donc régulière, du point de vue de l'observateur.

Les a priori utilisés par le système visuel humain pour construire une perception 3D sont de nature très diverses. Aux hypothèses citées ci-dessus, on peut ajouter la préférence pour la perception

[13] Les rats, par exemple, effectuent des mouvements de tête pour déterminer la distance aux objets de leur environnement.

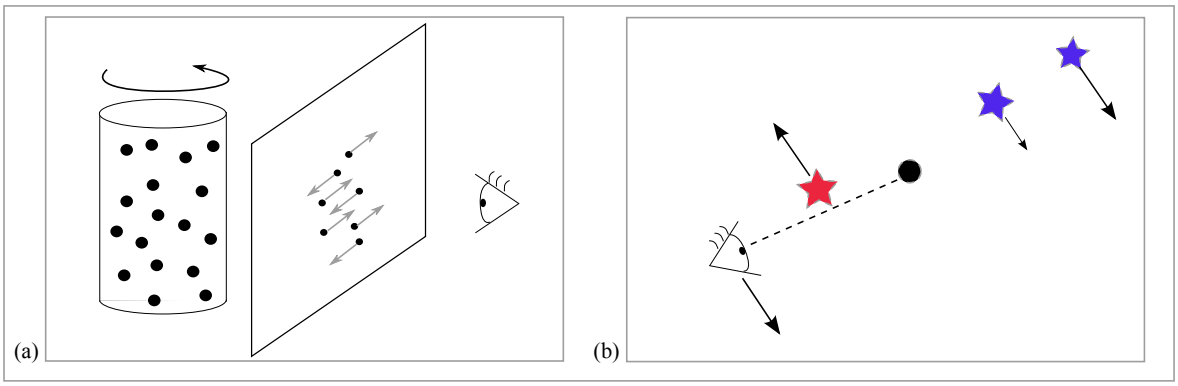


FIG. 2.8: Indices cinématiques de perception de la profondeur. (a) Le mouvement d'un nuage de points projetés sur un écran 2D engendre une perception tridimensionnelle par effet cinétique de profondeur (KDE). (b) Le mouvement de l'observateur fixant un point de la scène fournit une indication importante sur les relations de profondeur entre les différents éléments visuels (parallaxe de mouvement). Notamment, les objets les plus proches se déplacent dans une direction opposée au mouvement, les plus lointains dans la même direction, et les vitesses de déplacement des projections rétiniennes varient avec la distance.

de surfaces globalement convexes, qui se manifeste notamment dans l'illusion du masque [14], ainsi que l'influence de la connaissance a priori de la forme d'un objet [SP96a]. Cette dernière peut même dominer en l'absence d'indices stéréoscopiques contradictoires, comme l'ont montré Bülthoff et al. [BBS98] en utilisant quelques points dont l'animation 2D correspond à un mouvement biologique.

2.1.6 Combinaison d'indices multiples

Les informations spatiales peuvent donc être extraites au moyen de différents indices visuels picturaux, dynamiques et stéréoscopiques. Si un seul de ces indices est la plupart du temps insuffisant pour lever les ambiguïtés inhérentes à la perception tridimensionnelle, la combinaison de plusieurs d'entres eux peut faciliter l'émergence d'une perception tridimensionnelle véridique, c'est à dire rendant compte exactement des caractéristiques spatiales du monde observé. Plusieurs études montrent que cette combinaison peut être réalisée de manière quasi optimale, en minimisant la variance des profondeurs estimées [HWLB04, KS03]. Les modèles d'intégration proposés prennent également en compte les indices ordinaux, c'est-à-dire n'offrant qu'une information relative de profondeur, comme la configuration spatiale résultant de l'occultation [BPP05], ainsi que les a priori sur la forme des objets, comme le biais de convexité [AGE04].

Si l'accumulation d'indices visuels facilite la perception de la profondeur et des relations spatiales entre objets, deux indices peuvent se retrouver en conflit. Cela peut aboutir à une perception faussée ou aberrante, bien qu'apparemment robuste. Par exemple, dans le cas d'un masque en rotation, le biais de convexité induit un changement de la rotation perçue, alors même que cette perception d'inversion de profondeur impliquerait un changement soudain de la direction d'éclairage, ainsi que des incohérences de la surface sur les bords de l'objet. En visualisation, le conflit le plus courant résulte de l'observation d'une scène tridimensionnelle sur un écran 2D, l'utilisation d'écrans stéréoscopiques étant peu répandue. Dans ce cas, les indices picturaux présents dans l'image contredisent les indices binoculaires, qui signalent un objet plat, ainsi que les informations de mouvement propre, le mouvement de l'observateur devant son écran induisant un effet de parallaxe de mouvement non compatible

[14] *Hollow-face illusion* : un masque creux apparaît le plus souvent comme convexe, même lorsqu'il est vu de l'arrière. Lorsque masque est mis en rotation, le sens de rotation paraît s'inverser lorsque la face creuse fait face à l'observateur.

avec la structure spatiale de la scène. On est dans le cas très classique de l'observation de tableaux, pour lequel les artistes de la Renaissance suggéraient qu'il était nécessaire de regarder l'œuvre avec un seul œil positionné exactement au point de projection. A l'écart de ce point, la projection de l'image sur la rétine est déformée. Cependant, tourner autour d'un tableau ne nous donne pas l'impression d'une déformation du contenu de l'image, nous éprouvons une relative constance de l'espace pictural. Vishwanath et al. [VGB05] ont montré que celle-ci est plus importante si l'observation est réalisée avec les deux yeux plutôt qu'avec un seul œil, et si de plus le cadre de l'image est visible, ce qui suggère que l'information d'inclinaison de la surface de l'image est utilisée pour compenser la déformation de l'image quand on regarde celle-ci de biais. Concernant le conflit induit par la vision binoculaire d'une surface plane, il n'y a pas destruction de la perception en profondeur dans l'image, cependant le relief pictural effectivement perçu est moindre que celui résultant d'une observation avec un seul œil [Koe98].

Dans le cas de la visualisation scientifique, le nombre d'indices effectivement utilisables est soumis à deux contraintes, à savoir le temps nécessaire à leur production, qui interfère potentiellement avec l'interactivité de l'exploration, ainsi que le masquage potentiel des informations clés. L'idéal serait de pouvoir sélectionner les indices les plus informatifs, tout en s'assurant que ceux-ci ne masquent pas les représentations spécifiquement associées aux valeurs des données. En effet, la perception de l'espace ne constitue qu'un aspect de la visualisation, et l'utilisation par exemple de textures pour améliorer la perception de la courbure d'une surface peut réduire la saillance des informations relatives aux valeurs du champ affichées par une carte de couleur. Inversement, l'affichage de données peut perturber la perception de l'espace : une carte de couleur induit un texturage irrégulier, la modulation de la taille de points par la valeur d'un scalaire génère une fausse information perspective... La sélection des indices les plus pertinents est un exercice délicat, qui dépend aussi de la tâche à réaliser, et de la complexité de l'affichage. En utilisant un rendu synthétique de deux objets en suspension dans une pièce virtuelle, Wanger et al. [WFG92] montrent par exemple que perspective et ombres portées facilitent la détermination de la position relative des objets en profondeur, alors que la perspective, contrairement au mouvement de la caméra, réduit la capacité à comparer l'orientation et la taille de deux objets identiques occupant des positions différentes. Dans une autre étude, Hubona et al. [HWSB99] montrent que l'apport des ombres portées pour positionner des objets diminue lorsque deux sources de lumière sont utilisées, au lieu d'une. En effet, les ombres portées apportent une information utilisable par le système visuel seulement si celui-ci est capable de résoudre le problème de correspondance, à savoir attribuer correctement une ombre à l'objet qui en est la source. Dans un rendu comportant de très nombreux éléments, cette tâche devient trop complexe, surtout si les ombres ne sont pas projetées localement. Dans le choix des indices visuels pertinents pour rendre compte de la configuration spatiale d'une scène de visualisation, il convient donc toujours de considérer la nature des objets visualisés, ainsi que la tâche à accomplir par l'opérateur. Par exemple, si celui-ci doit avant tout comparer l'extension de phénomènes distants, le recours à une projection perspective doit parfois être évité.

2.1.7 Perception du volume

Fondamentalement, la capacité du système visuel humain à construire une perception tridimensionnelle repose sur l'extraction d'informations relatives à des surfaces et à des objets solides avec lesquels nous pouvons interagir. Comme le soulignent Fleming et Anderson [FA03], notre perception de la profondeur doit vraisemblablement refléter la structure du monde dans lequel nous évoluons, et pour lequel les entités accessibles sont des objets et des surfaces.

Cependant, les entités auxquelles la visualisation scientifique est consacrée ont souvent une nature profondément différente. Les phénomènes simulés ont une extension volumique, les champs scalaires,

vectoriels ou tensoriels des variables étudiées sont calculés en tout point de l'espace, et pas seulement à la frontière des domaines. Analyser de tels objets demande non seulement de percevoir des relations spatiales, mais également de comprendre l'extension et la variation d'un phénomène en profondeur. D'un point de vue pratique, les modes de représentations disponibles pour extraire de l'information en visualisation demeurent pour l'essentiel linéiques ou surfaciques. En raison de l'occultation induite par de tels rendus solides, il peut alors être nécessaire de déplacer les valeurs d'intérêt. Par exemple, une isosurface de température est limitée à une valeur unique ; accéder à une plage de valeur dans une région de l'espace nécessite alors d'afficher successivement les différentes surfaces, et de reconstruire mentalement le volume correspondant, ce qui semble cognitivement très difficile, voire impossible dans de nombreux cas.

Pour contourner le problème de l'occultation, de nouvelles solutions reposant sur l'utilisation de la transparence ont été proposées. L'idée consiste à rendre visible les surfaces normalement cachées en appliquant une certaine transparence aux objets normalement occultants. Le rendu volumique direct (*Direct Volume Rendering*, DVR) cherche à étendre ce principe à un volume continu, représenté comme un objet partiellement transparent dans lequel la lumière est en partie absorbée, si bien que l'intensité finale d'un pixel dans l'image produite correspond non plus à un élément de surface, mais à un segment s'étendant en profondeur. De telles visualisations ont été proposées pour visualiser des champs scalaires [DCH88, Sab88], ainsi que pour étendre une technique surfacique de visualisation de champs de vecteurs, les LIC (*Line Integral Convolution*, [CL93]) à un volume 3D [IG97]. Nous reviendrons plus avant sur les problèmes perceptifs relatifs à ces rendus volumiques dans le chapitre 3. Dans le cas où il s'agit seulement de visualiser deux surfaces superposées, Interrante et al. [IFP97] proposent un rendu de la surface occultante par transparence, en superposant une texture opaque constituée de traits épars orientés localement en fonction de la courbure de la surface. Ils mènent par ailleurs une expérience qui montre qu'une telle approche, comparée à une transparence complète, permet de mieux rendre compte de la forme de la surface occultante, ainsi que de sa distance à la surface arrière. Les paramètres des textures facilitant une telle visualisation ont été étudiés expérimentalement plus en détail par [HBW06, BH07], avec pour objectif de déterminer les types de textures perceptivement optimales pour visualiser de multiples surfaces. Notons cependant que, dans de tels rendus, l'utilisation de textures limite l'information affichage d'une part, et génère une occultation partielle, si bien que l'approche peut difficilement être étendue à plus de deux surfaces.

Dans tous les cas, la perception des volumes pose des questions nouvelles en termes de perception visuelle, et constitue un défi important en informatique graphique. Si le système visuel humain est capable de percevoir la spatialité du monde qui nous entoure sur la base de différents indices, peut-il appréhender une information volumique ? Quelles sont ses limites en ce domaine, et quels indices visuels, quels types de représentations peuvent éventuellement faciliter cette tâche ?

2.2 L'interactivité en visualisation scientifique

2.3 Interactivité et temps réel

Les termes d'interactivité et de temps réel sont utilisés très couramment en informatique et en réalité virtuelle, mais la réalité qu'ils recouvrent peut varier fortement d'un contexte à l'autre, suivant la finalité du système auquel ils sont appliqués.

L'interaction désigne la capacité d'un système à réagir aux commandes de l'utilisateur, à fournir une réponse aux différentes actions effectuées par celui-ci. L'interactivité effective d'un système va dépendre du temps intervenant entre la commande de l'utilisateur et la réponse du système, à savoir la latence. Dès lors, dans le cas d'un système 3D, on peut considérer de façon stricte qu'une « interaction en temps réel est obtenue si l'utilisateur ne perçoit pas de décalage temporel (latence) entre son action

sur l'environnement virtuel et la réponse sensorielle de ce dernier » [15].

En pratique, interactivité et temps réel sont définis avec plus de souplesse, leur signification dépend de la tâche à réaliser et de la dynamique des informations produites par le système. Lorsqu'il s'agit d'extraire certaines informations d'un grand ensemble de données pour une visualisation statique, un retour obtenu en quelques secondes peut être considéré comme interactif [Sai94]. En réalité virtuelle, le rafraîchissement de l'affichage doit être au minimum de dix images par seconde pour maintenir un sentiment de présence nécessaire à l'immersion, et l'environnement doit afficher les données correspondant à une sonde manipulée par l'utilisateur en moins d'un tiers de seconde pour que celui-ci ait un réel sentiment d'exploration [Bry05]. Pour les jeux vidéos, on parle généralement de 3D interactive au delà de 30, voire 60 images par seconde, cette vitesse répondant à l'impératif d'interagir avec des objets bougeant rapidement, et à la volonté d'éviter toute fatigue entraînée par un affichage apparaissant saccadé.

En ce qui concerne la visualisation d'objets 3D sur un écran, nous avons vu précédemment qu'il est possible d'obtenir une perception tridimensionnelle au moyen du déplacement 2D des objets projetés (KDE). Dès lors, il peut être intéressant de s'appuyer sur les conditions spatio-temporelles nécessaires à la perception d'un KDE pour définir un rendu interactif.

2.4 Notions sur la perception du mouvement

Dans les expériences initiales étudiant l'effet cinétique de profondeur (KDE), la visualisation proposée était continue. A l'époque, le dispositif était constitué d'un objet réel en rotation et dont une ombre était projetée sur un écran. L'effet de 3D était alors obtenu en observant cette ombre, dont le mouvement sur l'écran était parfaitement continu. Cependant, en informatique, comme au cinéma, l'illusion de mouvement naît de la succession d'images statiques présentées rapidement, le déplacement des éléments visuels d'une image à l'autre produisant un mouvement apparent.

Il est possible d'obtenir cette illusion de mouvement avec des stimuli très simples, comme un point unique présenté successivement à droite ou à gauche de l'affichage. Cependant, suivant l'amplitude des variations spatiales entre deux images et le temps qui les sépare, le mouvement perçu aura des caractéristiques différentes. L'observateur peut avoir l'impression d'un mouvement continu, ou avoir l'impression que l'objet s'est déplacé, comme téléporté, d'un endroit à l'autre de l'affichage. Dans le cas où plusieurs objets sont présents, la modification du temps écoulé entre deux images peut affecter le type de mouvement perçu. Par exemple, dans l'effet Ternus (discuté dans [SSH01]), la présentation d'un point successivement à gauche et à droite de deux autres points donne l'impression que ce point bouge successivement à gauche et à droite de deux éléments statiques, ou au contraire que les trois points bougent ensemble vers la gauche et vers la droite, suivant que le blanc séparant les deux images est inférieur ou supérieur à 40-50 ms (voir figure 2.9).

La littérature relative à la perception du mouvement apparent chez l'homme est très riche, et il ne saurait être question ici de la passer entièrement en revue. Nous présentons simplement quelques résultats expérimentaux relatifs aux conditions spatiales et temporelles de détection d'un mouvement cohérent 2D et 3D dans une succession d'images.

Les travaux menés par Braddick [Bra74] ont mis en évidence l'existence de deux mécanismes de perception du mouvement opérant respectivement à courte portée (short range), c'est-à-dire pour de faibles intervalles spatio-temporels, et à longue portée (long range). Le premier correspond plutôt à un traitement de flux optique par le système visuel, alors que le second relève de la détection de mouvement par la mise en correspondance de caractéristiques entre deux images.

[15] site internet de l'Association Française de Réalité Virtuelle, Augmentée, Mixte et d'Interaction 3D (http ://www.afrv.fr/)

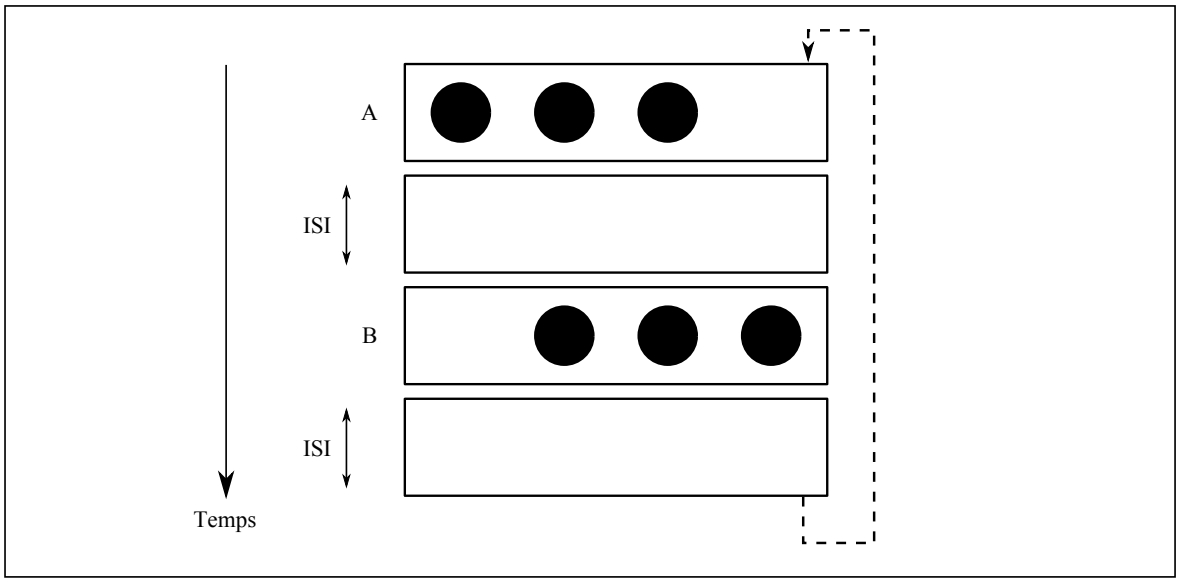


FIG. 2.9: Dans l'illusion de mouvement de Ternus, les images A et B sont présentées en alternance, et séparées par un blanc dont la durée détermine le type de mouvement perçu. Pour un ISI (*interstimulus interval*) court, un point semble osciller de gauche à droite, alors qu'un ISI plus long donne l'illusion d'un mouvement d'ensemble, les trois points semblant se déplacer tous les trois de gauche à droite. Des ISI intermédiaires (autour de 40-50 ms) donnent lieu à une perception bistable, le mouvement ressenti alternant au cours du temps entre les deux possibilités.

L'expérience de Braddick repose sur un nuage de points dont une partie seulement ont un mouvement 2D entre deux images compatible avec le déplacement horizontal ou vertical d'un rectangle. Les deux images sont présentées pendant 75 ms, avec un ISI (Inter-Stimulus Interval) blanc de 10 ms. La capacité à détecter le rectangle se détériore très vite pour des déplacements supérieurs à 15-20' d'arc. Pour des petits déplacements, l'augmentation de la durée du blanc entre les images à 40 et 80 ms dégrade également la capacité à détecter un mouvement cohérent. Baker et Braddick [BB85] confirment que la perception d'un mouvement à courte portée dans un nuage de points nécessite que les deux images apparaissent dans un intervalle de temps inférieur à 40 ms, la vitesse du mouvement ne constituant pas le facteur prépondérant. L'utilisation d'un stimulus dont l'intensité varie continûment selon une sinusoïde, mais pour lequel la composante fondamentale a été retirée, permet à Georgeson et Harris [GH90] de discriminer facilement les mécanismes à courte et longue portées, les directions de mouvement perçues étant alors opposées suivant le mécanisme prévalent. Dans cette configuration, ils trouvent également un ISI voisin de 40 ms séparant les deux types de traitement du mouvement.

Dans les travaux précédents, le mouvement apparent étudié est bidimensionnel, au sens où il n'induit pas d'impression de profondeur dans l'affichage. Dick et al. [DUS91] utilisent un stimulus constitué de points dont une partie appartiennent à un cylindre en rotation, le reste bougeant aléatoirement. Dans leur expérience, quatre images présentées 32 ms chacune se succèdent. Ils observent que la détection de la présence d'un cylindre en rotation est effectivement corrélée au nombre de points présentant un mouvement caractérisé par une extension spatiale à courte portée, à savoir moins de 20' d'arc. Pour des amplitudes plus importantes, il est nécessaire d'augmenter le nombre d'images pour susciter une extraction de structure par le mouvement (SfM). Par ailleurs, pour obtenir un effet cinétique de profondeur à partir d'un nuage de points, il n'est pas nécessaire que tous les points se déplacent continûment : dans un affichage pour lequel les points ont une durée de vie de

seulement deux images avant d'être réinitialisés aléatoirement sur la surface en mouvement, la perception d'une forme 3D persiste, malgré le scintillement et le bruit introduits dans l'affichage [SD94].

2.5 Interactivité requise pour les algorithmes de visualisation scientifique

La perception du mouvement est un phénomène complexe qui ne saurait être réduit aux effets de courte et longue portées décrits précédemment. Elle est également fortement influencée par l'intensité lumineuse moyenne et le contraste de l'image. La perception d'un mouvement tridimensionnel sur la base de déplacements bidimensionnels —effet cinétique de profondeur —est réalisée très rapidement pour des stimuli élémentaires comme des nuages de points pour une plage de valeurs correspondant à un mouvement à courte portée, mais elle peut exister au-delà, notamment si d'autres indices visuels de profondeur sont présents dans l'affichage.

En tout état de cause, il est donc délicat de donner une définition très précise et universelle de l'interactivité nécessaire en visualisation scientifique pour faciliter la perception de l'espace lors de changements continus du point de vue contrôlés par exemple par l'utilisateur au moyen de la souris. Il semble néanmoins que les algorithmes de rendu mis en œuvre doivent permettre de générer un minimum de 15 à 20 images par seconde.

3 Place de la perception en visualisation scientifique : études utilisateurs, principes et limites

Le but principal de la visualisation scientifique est d'offrir des outils efficaces pour l'exploration visuelle de données complexes. Les techniques de représentations proposées doivent d'une part transmettre effectivement l'information utile, d'autre part en faciliter l'extraction par une exploration visuelle. Comme le formulent justement Gaither et al. [GGLE04], cela nécessite entre autres de tenir compte des propriétés intrinsèques de la perception visuelle chez l'homme, ce qui exige une connaissance de la biologie du système visuel, mais aussi des principes relatifs à la théorie de la vision et à la psychologie de la perception.

Les expériences de psychophysique [16] peuvent ainsi aider à déterminer des caractéristiques essentielles du système visuel humain, et produire des résultats directement transposables. Par exemple, l'étude de la perception des couleurs permet de développer des échelles de couleur optimisant la quantité d'information transmise, mélangeant couleurs et intensité en tenant compte des propriétés du SVH [KHI+03]. Les travaux relatifs à l'acuité du SVH et au problème de l'attention visuelle (voir section 1) peuvent aider à développer des métriques de simplification destinées à accélérer le rendu d'ensembles de données complexes sans perte d'information utile.

Mais si les expériences de psychophysique ou études utilisateur peuvent orienter les développements en visualisation, elles se heurtent à certaines limites. La conscience de ces limites est une étape nécessaire, notamment pour recueillir le plus d'informations de telles expériences et éviter d'en tirer des généralisations trop hâtives.

3.1 Première limite : la complexité du stimulus

La première limite des expériences menées en perception découle de la qualité des stimuli utilisés. Pour comprendre un aspect particulier de la vision humaine, il est nécessaire de restreindre le stimulus à la dimension pertinente, afin de limiter les effets potentiels d'autres indices visuels. Mais il s'agit

[16] la psychophysique est la discipline qui vise à déterminer les relations entre des stimuli physiques et les états subjectifs qu'ils induisent (perception).

d'un exercice délicat, comme le souligne Koenderink [KvD03] à propos de la perception de la forme par l'ombrage : on ne peut être certain a priori que les résultats obtenus puissent être étendus à une situation plus normale ou « écologique » de vision, qu'il s'agisse de notre perception visuelle dans le monde réel ou, pour ce qui nous intéresse plus particulièrement en visualisation scientifique, de notre perception relative à des affichages synthétiques complexes. La difficulté est en fait double. Il faut d'une part s'assurer que le stimulus reproduise effectivement les caractéristiques visuelles que l'on souhaite étudier, ce qui est délicat si l'on étudie le système visuel en soi au moyen de stimuli générés par ordinateur, un peu moins quand on s'intéresse à la perception induite par un rendu par ordinateur. D'autre part, le nombre de conditions que l'on peut explorer au cours d'une expérience est nécessairement restreint, car l'augmentation du nombre de dimensions dans le stimulus accroît de manière exponentielle le nombre d'essais à réaliser, et rend plus délicate l'analyse des résultats.

Pour illustrer cette idée, prenons deux exemples concrets relatifs à la visualisation. Dans une étude récente, Weigle et David [WB08] ont cherché à comparer les apports respectifs de l'ombrage et de la perspective linéaire pour la discrimination en profondeur de lignes de champ intriquées représentées sous forme de tubes blancs. Leurs résultats expérimentaux, dans la configuration particulière de l'expérience proposée, montrent qu'un éclairage global plutôt que local apporte autant d'information qu'une projection perspective plutôt qu'orthographique pour réaliser la tâche demandée. Cette expérience apporte un résultat intéressant, mais dont la généralisation soulève plusieurs questions : dans une visualisation plus complexe où les tailles d'objets varient, qu'apporte réellement la perspective ? Et si on ajoute une information supplémentaire sur les lignes sous la forme d'une carte de couleur et d'une texture, comment cela module-t-il l'information apportée par le modèle d'éclairage ? Par ailleurs, des études expérimentales ont mis en évidence que certains types de textures facilitent la perception de la forme d'une surface [TR90]. Cependant, l'utilisation de ces résultats en situation concrète de visualisation se heurte par exemple au problème de masquage de l'information pertinente que ces textures sont susceptibles d'introduire [KHI+03], ce qui est difficile à déterminer a priori.

Malgré tout, ces deux expériences apportent une information véritablement utile qui peut guider le développement de nouvelles techniques de visualisation. Seulement, il ne faut pas oublier leurs limites potentielles, afin d'éviter de tomber dans une impasse.

3.2 Deuxième limite : l'expertise du sujet

La perception n'est pas un processus passif et figé, mais susceptible d'évoluer en fonction des expériences vécues, des informations que notre cerveau doit traiter de manière récurrente. Suivant notre activité principale, nous sommes capables d'acquérir une forme particulière d'expertise perceptive, qui nous rend plus à même d'extraire certaines informations présentes dans notre champ visuel, voire influe sur notre manière d'explorer visuellement une image (section 1.2.5, p.45). Par exemple, un joueur d'échecs distingue très rapidement les groupes de pièces importantes dans une configuration de milieu de jeu et mémorise mieux les pièces qu'un novice seulement s'il existe une cohérence relative au jeu dans leur organisation [Get96]. Lorsqu'il s'agit de compter rapidement le nombre d'objets présents sur un écran, l'organisation de ces objets sous la forme d'un tableau 2D améliore significativement plus les performances de contrôleurs aériens que celles de sujets non experts [AM08]. Par ailleurs, Haller et Radue [HR05] ont observé que les radiologues sont plus rapides que des non experts pour déterminer si une modification a été introduite dans une image, que cette dernière appartienne à leur domaine de compétence ou non. Par ailleurs, ils relèvent que les zones du cerveau activées lors de la réalisation de cette tâche diffèrent dans les deux groupes.

La question de savoir ce qui détermine l'expertise perceptive et la façon dont elle peut être codée dans le cerveau est encore très ouverte. L'expertise perceptive résulte très vraisemblablement de processus de traitement bas-niveau de l'information visuelle, ainsi que de la modulation de ces signaux

par des informations cognitives plus complexes, comme la mémoire ou la capacité à catégoriser certaines formes. Plusieurs études ont ainsi mis en évidence l'existence d'un apprentissage perceptif (*perceptual learning*). Celui-ci est issu directement de l'exposition répétée à un même type de données, et permettrait par exemple de réduire le bruit perceptif pour certains types d'objets ou de stimulations visuelles. Cette adaptation peut être réalisée sur la dimension du stimulus sur laquelle l'attention se focalise, mais également sur des modalités différentes. Par exemple, si l'on demande à des sujets de reconnaître une lettre affichée au centre de l'écran tout en leur présentant des points en mouvement en périphérie, on constate qu'ils sont plus à même de discriminer un mouvement cohérent dans un nuage de points après l'expérience qu'avant, alors que la tâche de reconnaissance était sans rapport [WNS01]. Seitz et Watanabe [SW05] proposent un modèle d'apprentissage perceptif rendant compte de ce phénomène, en suggérant que l'attention portée à la réalisation d'une tâche modifie la capacité de notre système perceptif à apprendre, à se modifier pour répondre à l'ensemble des stimulations présentées à ce moment-là.

Palmeri et al. [PWG04] décrivent quelques-unes des propriétés inhérentes à l'expertise perceptive. Parmi celles-ci, on note chez les experts une plus grande rapidité d'exécution, la possibilité de réaliser une autre tâche en parallèle sans être distraits, une capacité à généraliser, de manière limitée, leurs connaissances, mais également une plus grande difficulté à ignorer des parties d'images non pertinentes pour la tâche d'appariement à réaliser. Concernant la catégorisation, les experts sont capables d'extraire et de s'appuyer sur les dimensions les plus importantes caractérisant un objet, celles-ci pouvant à la fois être visuellement explicites (couleur, taille, forme) ou relever de connaissances plus abstraites (présence d'une poignée, de pieds). L'expertise perceptive pourrait ainsi relever d'une meilleure orientation de l'attention vers les dimensions pertinentes d'un objet. Elle pourrait également être caractérisée par la constitution de nombreux blocs perceptifs (*perceptual chunks*) regroupant certaines caractéristiques liées au domaine concerné, et rendues familières par la pratique [GLC+01].

Pour en revenir aux expériences menées en perception visuelle, l'expertise perceptive pose notamment le problème de l'influence des participants sur les résultats obtenus. Leur sélection n'aura que peu d'importance si l'on étudie des propriétés très basiques du système visuel, comme l'acuité en vision fovéale ou périphérique. Par contre, l'étude de la plus-value d'un rendu relativement complexe peut donner des résultats assez différents selon que l'on sonde une population quelconque ou les utilisateurs finaux à qui cette technique est destinée, en raison de l'expertise de ces derniers. Cependant, la mise en place d'expériences concrètes avec des praticiens est délicate à réaliser, surtout dans des domaines comme la simulation en mécanique des fluides et en thermique où les problèmes posés sont très variés, au contraire des cliniciens par exemple qui forment une communauté importante concentrée sur des problématiques et des objets de visualisation relativement similaires (angiographies, radiographies des poumons, IRM...).

3.3 Troisième limite : la nature de la tâche expérimentale

Lorsque qu'un ingénieur explore visuellement un ensemble de données résultant d'une simulation numérique, les phénomènes principaux qui vont émerger ne sont pas parfaitement définis. La visualisation, à travers des modes de rendu appropriés, doit lui servir à extraire cette information, mais il s'appuie également sur un ensemble de connaissances a priori relatives à la physique sous-jacente, au problème spécifique étudié et aux hypothèses formalisées dans ce cadre, voire sur des résultats obtenus par d'autres moyens (expériences sur des maquettes, observations de terrain). Cela complique notablement la mise au point d'une tâche expérimentale permettant par exemple de vérifier quantitativement, voire même qualitativement, l'apport d'un nouvel algorithme, une tâche trop simple risquant de ne pas mettre correctement en valeur les apports et faiblesses effectives de celui-ci.

Parfois, il sera plus utile d'obtenir des retours informels d'utilisateurs finaux, de discuter avec eux pour essayer de cerner les avantages potentiels de telle ou telle technique. Kosara et al. [KHI+03] citent ainsi un exemple d'étude sur une nouvelle technique de réalité virtuelle destinée à l'archéologie. Les auteurs de la technique ne réussissant pas à mettre au point un protocole permettant de valider l'intérêt de leur technique, ils ont simplement procédé à un visionnage de la technique auprès de spécialistes du domaine. L'échange qui s'en est suivi a finalement permis de faire émerger de nouvelles idées et de valider l'intérêt de l'approche proposée.

Souvent, la mise au point de nouveaux algorithmes de visualisation est un jeu en double aveugle où deux expertises bien distinctes se répondent. Le chercheur en informatique graphique possède une compétence et/ou une pratique limitées dans le domaine physique pour lequel il propose de nouvelles solutions. Le praticien, quant à lui, est habitué à un nombre restreint de modes de visualisation, de par sa formation technique mais également en raison des implémentations effectivement disponibles dans les logiciels qu'il utilise. Il n'a donc en général qu'une vague idée de ce qui est réalisable effectivement, n'ayant pas le temps pratiquement de s'intéresser aux conférences de visualisation [vW06]. Dans tous les cas, il est donc au moins aussi utile d'établir des liens importants avec les praticiens que de produire et évaluer a posteriori un nouvel algorithme de visualisation afin d'essayer de mieux appréhender leurs problématiques principales, de cerner les difficultés auxquelles ils sont confrontés, voire quand c'est possible de les faire participer au processus créatif.

3.4 Conclusion

Les limites que nous venons de souligner ne remettent pas en cause l'importance pour la visualisation scientifique des travaux menés en perception visuelle. Des expériences de psychophysique, même simples, peuvent contribuer à définir les possibilités relatives à l'utilisation de certains indices visuels, et mettre en évidence les difficultés d'origine perceptive sous-jacentes à certains modes de représentation (perception de la transparence, problème de l'attention...). En cela, elles constituent une source importante d'inspiration pour de nouvelles visualisations. Elles permettent d'éviter les impasses et de guider les futurs travaux, en aidant à formuler rigoureusement les problèmes posés.

Par ailleurs, dans ce qui précède, il ne s'agit en aucun cas de soumettre la visualisation scientifique à un cahier des charges restrictif et défini uniquement par les besoins des utilisateurs finaux. Mais si certaines disciplines peuvent, et doivent, se détacher de l'utilisabilité des recherches qu'elles produisent, ce qui est le but de toute recherche fondamentale, la visualisation scientifique a un statut plus partagé. C'est une science appliquée, et les questions et pratiques auxquelles elle doit répondre sont ancrées dans la réalité de pratiques quotidiennes. Ces dernières, si elles sont bien comprises, peuvent constituer un terreau fertile favorisant l'émergence de nouvelles idées et de nouvelles problématiques à explorer.

Pour terminer, notons que si les recherches menées jusqu'à présent en sciences cognitives ont permis de lever le voile sur de nombreux aspects de la perception visuelle, le système visuel humain demeure un système très complexe dont le fonctionnement nous échappe encore grandement. Pour aider à mieux le comprendre, Cavanagh [Cav05] suggère de recourir à l'artiste en tant que neuroscientifique averti. En effet, les artistes, de par leur pratique, ont une connaissance implicite de la perception, ils savent comment utiliser au mieux nos sens, ou au contraire comment les tromper. De même, les techniques issues de la peinture ou de l'illustration technique peuvent constituer une source d'inspiration importante pour créer de nouvelles visualisations perceptivement efficaces. De nombreux travaux ont été réalisés en ce sens, présentant des techniques de visualisation qualifiées de non photoréalistes (NPR, *Non-Photorealistic Rendering*), par exemple dans le cas du rendu de données volumiques scalaires [RE01, SLM02, SE03, BKR+05]. Keefe et al. [KKVL05] ont même mené une expérimentation visant à déterminer comment des artistes peuvent participer au processus d'élaboration de nouvelles

visualisations applicables à des cas d'exploration, et pas simplement d'illustration de résultats. Les auteurs vont jusqu'à suggérer que des artistes seront bientôt embauchés pour travailler à plein temps sur des problématiques de visualisation scientifique. Si cette conclusion semble un peu exagérée, il n'en reste pas moins que, tout autant que les connaissances directes du système visuel issues de recherches en neurosciences et d'expériences avec des sujets humains, les représentations dérivées du monde de l'art peuvent fournir des idées utiles et efficaces pour développer de nouvelles visualisations, en ce qu'elles relèvent souvent d'une connaissance implicite du fonctionnement de la perception visuelle chez l'homme.

4 Perspectives, orientation du travail

Dans le domaine de la visualisation scientifique, des logiciels informatiques servent à générer des images représentant des données physiques tridimensionnelles résultant de simulations numériques ou des mesures reconstituées à partir de divers capteurs. Ces images sont analysées visuellement par un spécialiste du domaine concerné afin d'extraire une information difficilement accessible par un autre moyen. Cette exploration met en jeu d'une part les capacités et limites du système visuel, qui déterminent l'information effectivement perceptible par l'utilisateur final, d'autre part l'expertise de ce dernier qui influence la manière dont l'exploration est menée et l'interprétation des indices visuels présents dans l'image. Les logiciels doivent idéalement permettre de produire des rendus adaptés aux demandes de l'utilisateur, en fonction de la tâche à réaliser et de la nature des données à analyser, et contenir des indices facilitant la perception des formes des objets représentés et des relations spatiales entre différents éléments. Ces rendus doivent être interactifs, l'utilisateur doit pouvoir modifier rapidement les données au cours de l'exploration et obtenir un mouvement fluide lorsque le point de vue est modifié, afin d'une part de faciliter le positionnement de la caméra et d'autre part, si possible, de fournir une information de profondeur à l'occasion de ce mouvement. Pour schématiser, la visualisation scientifique peut être vue comme un triangle d'acteurs (figure 2.10) dont les interactions doivent être prises en compte pour obtenir une visualisation efficace, qui remplisse pleinement ses objectifs : l'architecture matérielle et logicielle de rendu, la vision humaine (perception) et l'expertise de l'utilisateur (cognition).

Comme nous l'avons vu dans le chapitre précédent, l'explosion des capacités de calcul engendre des données de plus en plus volumineuses. Celles-ci occupent typiquement plusieurs gigaoctets de mémoire physique, qui correspondent à l'enregistrement de nombreuses variables sur de multiples pas de temps pour des millions, voire des milliards de primitives géométriques. L'exploration visuelle de ces résultats pose alors deux problèmes fondamentaux : la capacité à réaliser des rendus interactifs, de sorte que l'utilisateur puisse effectivement interagir avec les représentations proposées et en modifier les paramètres en fonction des besoins de son analyse, et la possibilité de s'appuyer sur des indices visuels efficaces facilitant l'appréhension de la profondeur et du volume dans les scènes complexes visualisées.

Pour atteindre l'objectif de rendu interactif, la solution la plus simple requiert de déployer des environnements matériels et logiciels adaptés aux tailles de données sur lesquelles l'utilisateur est amené à travailler, clusters graphiques et logiciels de visualisation parallèles par exemple. Une autre approche pour accélérer les rendus, éventuellement complémentaire de la première, consiste à s'appuyer sur des représentations partiellement simplifiées qui limitent le nombre de primitives à afficher. Les limites de la perception visuelle chez l'homme, tant physiologiques que cognitives, peuvent être utilisées pour guider ces simplifications afin qu'elles ne conduisent pas à une dégradation de l'information effectivement accessible au cours de la visualisation. Cependant, une telle démarche nécessite de prendre en compte deux éléments importants : l'expertise de l'utilisateur, qui peut influer sur le

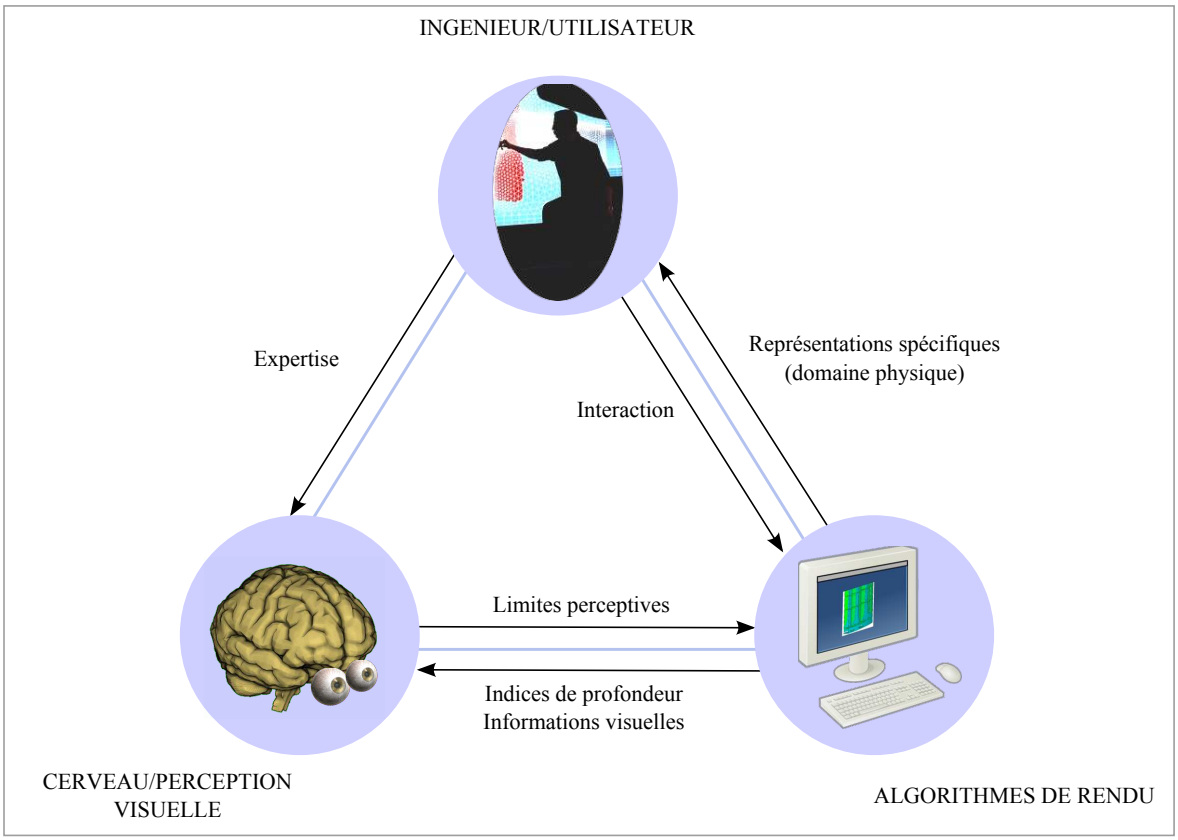


FIG. 2.10: Pour schématiser, les découvertes possibles au cours d'une visualisation scientifique dépendent des caractéristiques (performances et limites) des trois acteurs en présence, et de leurs interactions. Le matériel, les logiciels et les algorithmes déterminent les modes d'exploration et les indices de profondeur effectivement utilisables, ainsi que le niveau d'interactivité accessible. Cette information sera filtrée par le système visuel pour générer une perception guidée par l'expertise de l'utilisateur. Ce dernier sélectionne par ailleurs les modes de représentation les plus adaptés, en fonction de son expérience et de sa connaissance particulière du problème traité.

processus d'exploration visuelle et sur la perception même de certains éléments, et l'impératif d'interactivité dans la manipulation des données, les structures géométriques éventuellement créées en cours de visualisation ne pouvant faire l'objet de prétraitements trop longs.

En ce qui concerne la perception de l'espace dans les scènes complexes de visualisation, il est nécessaire de sélectionner et de rendre utilisables en pratique les indices visuels les plus pertinents au regard des conditions de visualisation. Ceux-ci doivent clarifier les images présentées, tout en ne nécessitant pas de temps de calculs trop importants. Par ailleurs, comme nous l'avons souligné dans la section 2.1.7 p.51, la visualisation scientifique est confrontée à un problème non naturel de perception d'objets volumiques, et non simplement surfaciques, la compréhension de l'extension en profondeur de certains phénomènes et la résolution des occultations dans des affichages denses se révélant particulièrement délicates. Dans les chapitres suivants, nous présentons des travaux menés dans cette direction. Nous proposons en premier lieu une étude perceptive des rendus dits volumiques visant à définir les limites de l'utilisation de la transparence pour représenter plusieurs couches de données (chapitre 3). Nous proposons ensuite deux nouveaux algorithmes de rendu implémentés directement sur la carte graphique et en espace image, afin d'accélérer le rendu indépendamment de la taille des données sous-jacentes. Le premier concerne une nouvelle méthode d'éclairage baptisée EyeDome

Lighting (chapitre 4), alors que le deuxième vise à régler le problème de l'occultation de données sur lesquelles l'attention est portée en creusant interactivement le contexte environnant en fonction du point de vue (chapitre 5).

CHAPITRE 3

Evaluation des rendus volumiques

Réaliser la représentation de l'irreprésentable,
voir l'invisible, toucher et percevoir l'impalpable.

NOVALIS

La représentation de données tridimensionnelles denses pose un défi important en visualisation scientifique. Idéalement, il faudrait pouvoir accéder en un coup d'œil à l'ensemble des phénomènes qui s'étendent dans le volume représenté, et être capable d'appréhender simplement l'organisation spatiale et les variations des champs physiques issus de la simulation numérique ou de l'acquisition réalisée. Cependant, de telles représentations ne sont pas naturelles, le monde qui nous entoure étant constitué principalement de surfaces opaques qui marquent les limites d'un volume et empêchent le système visuel humain d'accéder à leur intérieur ou aux objets plus distants qui se trouvent occultés. En visualisation, les représentations les plus classiques de champs scalaires reposent sur l'extraction de surfaces caractéristiques, que ce soit des plans de coupe ou des surfaces d'isovaleurs. On parle alors de *rendu volumique indirect*, l'exploration du volume nécessitant de calculer au préalable certaines primitives géométriques. Dans ce cas, la quantité d'information présente sur une image est nécessairement limitée par les nombreuses occultations introduites par ces surfaces, et il est alors nécessaire de reconstruire mentalement le volume d'intérêt à partir d'une succession d'images, ce qui représente une tâche cognitive très complexe.

Bien avant le développement des techniques informatiques de visualisation assistée par ordinateur, plusieurs solutions reposant sur l'interaction d'un rayonnement avec le milieu traversé ont été développées pour accéder à une information volumique non visible autrement. Ainsi, la technique d'imagerie de Schlieren utilise une perturbation d'une source de lumière collimatée afin de capturer sur une photographie les variations volumiques de densité d'un fluide, ce qui permet par exemple de voir la structure d'une onde de choc. Dans le domaine médical, l'imagerie par rayons X produit une impression photographique qui reflète la structure de l'ensemble du volume traversé, permettant de visualiser en une image les tissus et les os autrement non accessibles à l'observation. S'inspirant de

FIG. 3.1: Exemples naturels d'objets transparents ou translucides. Plusieurs vitres superposées reflétant des objets environnants (a), une brume à densité variable (b) ou un bloc de glace non homogène (c) présentent des structures se superposant visuellement dans la profondeur. Comprendre l'organisation spatiale ou les successions de densité en profondeur dans ces images est une tâche complexe.

ces techniques, il a été suggéré en visualisation scientifique de considérer le champ scalaire à analyser comme un volume de particules dotées de propriétés d'absorption et d'émission de la lumière, comme des gouttes d'eau dans un nuage. Cela a entraîné le développement des méthodes de *rendu volumique direct* (DVR, *Direct Volume Rendering*), qui reposent sur le principe d'une accumulation de transparences en profondeur [Sab88, DCH88, UK88].

Si de nombreuses implémentations et améliorations de ces techniques ont été proposées depuis la fin des années 80, il est curieux de constater qu'elles semblent peu utilisées en dehors de la communauté médicale, comme nous avons pu l'observer. En dehors de l'ignorance potentielle des utilisateurs finaux concernant leur existence, cela peut être dû à l'ambiguïté des images générées, la couleur de chaque pixel dépendant d'un continuum théorique de points tridimensionnels semi-transparents, ainsi qu'à la difficulté, même pour un utilisateur expert, du réglage de la fonction de transfert (FT) qui détermine la qualité du rendu final. Les images produites peuvent paraître jolies, mais en l'absence de structures caractéristiques dans les données il peut être plus que délicat de percevoir les relations spatiales dans le volume projeté, rendant la visualisation apparemment plus imprécise qu'avec d'autres techniques. Au quotidien, les situations dans lesquelles notre système visuel est confronté à de tels objets transparents ne sont d'ailleurs pas usuelles. On peut parfois observer les objets reflétés par plusieurs vitres superposées, regarder l'intérieur d'un bloc de glace non homogène ou voir un paysage à travers un brouillard discontinu (figure 3.1). Si l'on perçoit la nature transparente en soi de ces objets, leur extension spatiale nous paraît difficilement accessible. Cependant, lorsque Kaufman et Mueller [KM05] présentent une revue détaillée des techniques de DVR, aucune des près de trois cents références ne concerne une évaluation perceptive des rendus à base de transparence ou une validation de ces méthodes sur la base d'études utilisateurs.

Dans ce chapitre, nous nous intéressons à la question de la perception par le SVH des relations spatiales entre objets rendus par des méthodes volumiques, en nous restreignant au cas achromatique. Nous nous limitons à la perception de la profondeur, qui constitue un aspect important des rendus transparents par ailleurs facilement quantifiable, et ne traitons pas d'autres problèmes perceptifs tels que la reconnaissance de forme ou l'appariement. L'objectif ne saurait être de formuler des conclusions définitives sur ce sujet, mais plutôt d'orienter les futurs travaux dans ce domaine, ainsi que de

poser la question de la mise en pratique des techniques de DVR au sein des outils de visualisation d'EDF R&D. La charge computationnelle normalement requise par les rendus volumiques impose souvent des représentations statiques, alors que le mouvement peut fournir des informations utiles à la perception en profondeur (KDE, section 2.1.4 p.49). De plus, les ambiguïtés de profondeur peuvent être partiellement levées lorsqu'une perspective linéaire est préférée à une projection orthographique. Nous avons donc mené trois expériences visant à étudier la perception de la profondeur par un sujet humain dans un contexte statique ou dynamique, l'influence de la projection étant examinée dans ce dernier cas. Notre hypothèse de travail est que les rendus volumiques, même s'ils réalisent une accumulation de transparence ordonnée en profondeur, ne permettent que difficilement de comprendre les relations spatiales, et que les ambiguïtés inhérentes à ces affichages peuvent être partiellement levées par l'apport de mouvement et de déformations perspectives. Après une revue des techniques de rendu volumique (section 1) et des travaux relatifs à la perception de la transparence chez l'homme (section 2), nous présentons une première expérience dans laquelle nous étudions la capacité de sujets humains à discriminer en profondeur deux objets statiques rendus en DVR (section 3), l'accumulation de transparences réalisé en projection orthographique constituant ici le seul indice de profondeur. La deuxième expérience réalisée se concentre sur la perception de la profondeur dans un volume semi-transparent en rotation (section 4), et l'importance du mode de projection est étudiée dans une troisième expérience (section 5). Nous terminons par une discussion générale des résultats obtenus, et proposons quelques pistes pour améliorer la visualisation de données volumiques (section 6).

1 Les techniques de rendu volumique direct

Dans cette section, nous présentons un aperçu rapide du rendu volumique en informatique graphique, et des évaluations et améliorations de cette technique qui ont été proposées. Une lecture plus complète concernant le DVR et son implémentation avec accélération sur la carte graphique peut être trouvée dans [KM05, Pfi05].

1.1 Principe

L'idée d'utiliser des rendus semi-transparents pour visualiser des données scalaires trouve son origine en informatique graphique dans les travaux sur le rendu de nuages [KH84, Max95], objets pour lesquels il est nécessaire de calculer le trajet de la lumière et ses interactions avec le milieu traversé. Il a été proposé de transposer ces techniques de lancer de rayon pour produire une image volumique d'un champ scalaire tridimensionnel, en considérant chaque élément volumique élémentaire du champ comme une particule dotée de propriétés d'absorption et d'émission de lumière, la dispersion des rayons lumineux dans le volume n'étant pas prise en compte afin de simplifier les calculs [Sab88, DCH88, UK88].

Indépendamment de la technique effectivement implémentée, le DVR nécessite pour chaque pixel d'évaluer l'*intégrale de rendu volumique*. La couleur C d'un pixel p est calculée en sommant les contributions des voxels rencontrés par le rayon $x(\lambda)$ lancé depuis p, λ représentant la distance à l'œil de l'observateur (voir figure 3.2). Ce rayon traverse le volume le long d'un segment de longueur D, et la couleur de p est donnée par :

$$\mathsf{C} = \int_0^D C(x(\lambda))e^{-\int_0^\lambda \tau(x(\lambda')d\lambda')}d\lambda. \tag{3.1}$$

$C(x)$ représente la couleur d'émission du voxel x, $\tau(x)$ son coefficient d'extinction linéique. En pratique, les calculs sont réalisés en effectuant une somme de Riemann sur le segment $[O : D]$ découpé

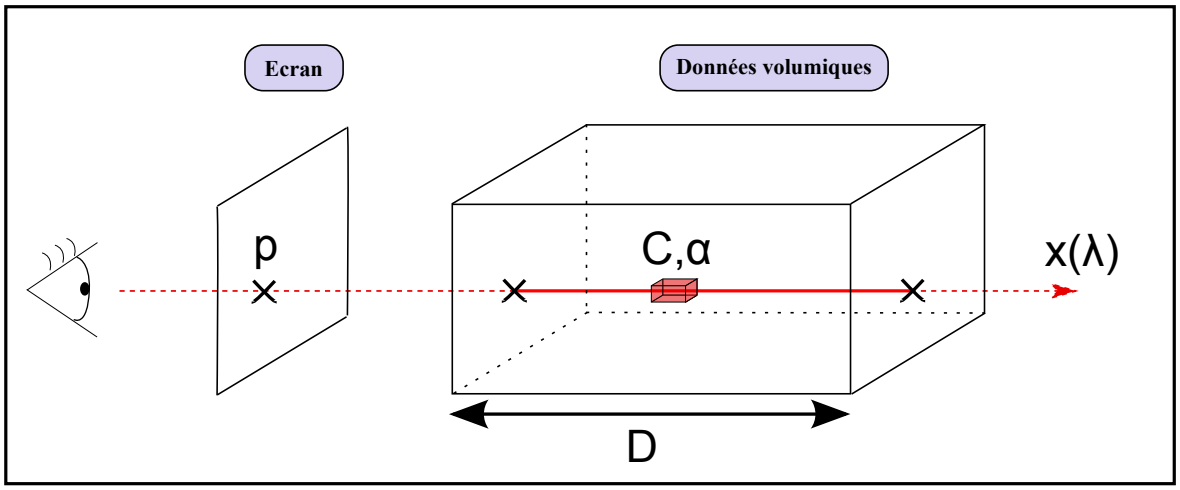


FIG. 3.2: Principe du rendu volumique par lancer de rayon.

en n intervalles

$$\mathsf{C} \approx \sum_{i=0}^{n} \tilde{C}_i \prod_{j=0}^{i-1} (1-\alpha_j) \tag{3.2}$$

Dans ce cas, α_i représente l'opacité du $i^{ème}$ voxel le long du rayon[1], donnée par

$$\alpha_i \approx 1 - e^{-\tau(x(i\frac{D}{n}))d}, \tag{3.3}$$

et $\tilde{C}_i$ sa couleur associée. Pour simplifier, cette somme est normalement calculée de l'arrière vers l'avant du volume, mais il existe également une formulation de cette solution approchée destinée à un calcul en sens inverse. Dans tous les cas, l'ordre des voxels est primordial, et le DVR nécessite donc de trier en profondeur les éléments du volume, ce qui peut s'avérer coûteux en temps de calcul, surtout si les données ne sont pas organisées selon une grille régulière.

Par ailleurs, pour faire face au problème d'échantillonnage spatial inhérent au rendu volumique, Engel et al. [EKE01] ont proposé une technique appelée *préintégration* qui améliore la qualité visuelle par un précalcul par morceaux de l'intégrale du rendu volumique, les résultats étant enregistrés dans une lookup table.

1.2 Choix de la fonction de transfert

Pour appliquer ce modèle optique simplifié au rendu d'un champ scalaire tridimensionnel, il est nécessaire de définir une fonction de transfert (FT)

$$\phi : s \rightarrow (C, \alpha) \tag{3.4}$$

qui associe au scalaire $s(x)$ porté par le voxel x une couleur C et une opacité α : c'est l'étape dite de *classification*. La couleur associée du voxel correspond au produit des deux : $\tilde{C} = \alpha C$. Dans le cas d'un rendu volumique achromatique, la couleur C peut être remplacée par une simple luminance l. Le rendu volumique étant de fait discret, il existe deux schémas d'interpolation sur les données : dans le cas de la *postclassification*, l'interpolation est réalisée sur le champ scalaire avant application de la FT, alors qu'en *préclassification* les couleurs et opacités sont calculées avant d'être interpolées.

Le choix de cette fonction de transfert est critique, car elle détermine la qualité du rendu final et la mise en valeur des régions importantes, mais le nombre de degrés de libertés qu'elle peut comporter

[1] La transparence d'un voxel i est définie par $1 - \alpha_i$.

rend la tâche très difficile. Il peut s'avérer nécessaire d'automatiser au moins en partie cette étape (voir [PLS+00] pour une discussion à ce sujet). Des formulations complexes ont été proposées, dans lesquelles les FT sont multidimensionnelles, prenant en compte non seulement le champ scalaire mais aussi son gradient ou une autre variable, voire reposent sur un calcul de courbure [KWTM03]. Les fonctions linéaires peuvent être remplacées par des gaussiennes centrées sur certaines régions dans l'espace des données [KPI+03]. Cependant, plus le nombre de paramètres à manipuler est important, plus le réglage de la FT par l'utilisateur devient difficile. En pratique, les FT choisies demeurent généralement linéaires par morceau.

1.3 Rendus volumiques non ordonnés

Un problème important du DVR concerne la nécessité d'intégrer pour chaque pixel les valeurs du champ suivant leur ordre en profondeur. Si cela s'avère relativement simple dans le cas d'un lancer de rayon réalisé sur une grille régulière cubique, il n'en est pas de même lorsque l'objet à représenter est constitué d'un arrangement irrégulier de primitives, par exemple de tétraèdres, qui doivent alors être correctement triées préalablement au rendu.

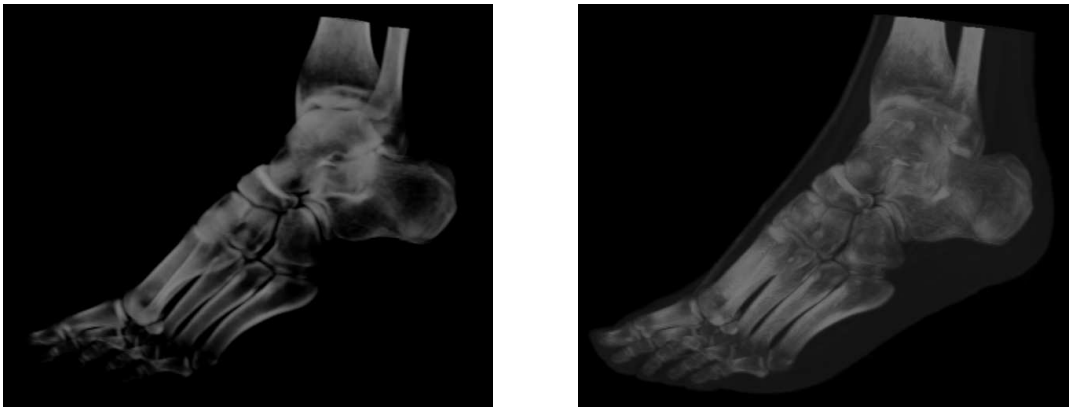
Deux techniques permettent de s'affranchir de l'ordre des données. Le MIP (*Maximum Intensity Projection*) consiste à afficher uniquement la valeur maximale projetée sur chaque pixel. A contrario, un rendu de type X-RAY accumule l'ensemble des intensités lumineuses le long d'un rayon, mais sans tenir compte de l'ordre. Cette technique aussi appelée DRR (*Digitally Reconstructed Radiograph*) vise à reproduire les images obtenues par rayons X. Elle est essentiellement utilisée sur des ensembles de données reconstruits par tomographie axiale calculée (CT, *Computed Tomography*).

Par nature, MIP et X-RAY sont ambigus quant à la perception de la profondeur dans les données affichées dans la mesure où, en projection orthographique, une vue arrière et avant du volume donneront deux images miroir l'une de l'autre. De plus, dans le cas du MIP, un changement de point de vue continu peut entraîner une modification importante de l'affichage, certains maxima apparaissant ou disparaissant. Pour éviter cet effet et augmenter la stabilité visuelle des structures représentées, il est possible de ne considérer que le maximum le plus proche du point de vue dépassant un certain seuil [SSN+98].

Dans ce qui suit, le terme DVR désignera uniquement un rendu volumique ordonné, c'est-à-dire basé sur le modèle optique décrit dans la section 1.1. Les techniques de MIP et X-RAY constituent également des rendus volumiques directs, mais nous y ferons référence par leur nom spécifique, comme c'est souvent le cas dans la littérature. La figure 3.3 montre des rendus obtenus respectivement par DVR et MIP.

1.4 Améliorations du rendu

En dehors des travaux visant à améliorer l'approximation de l'intégrale de rendu volumique, de nombreuses techniques ont été développées afin de faciliter la perception des formes dans les visualisations de rendu volumique. Pour faciliter la perception des surfaces au sein du volume, Levoy [Lev88] introduit ainsi un calcul d'éclairage dans le processus de rendu. Suivant les idées de Saito [Sai94], Rheingans et Ebert [RE01] intègrent au DVR des techniques de rendu non photoréaliste, comme la mise en valeur de silhouettes et le tone shading dans un cadre qu'ils appellent illustration volumique (*volume illustration*). Mora et Ebert [ME04] suggèrent d'améliorer MIP et X-RAY en s'appuyant sur les gradients du champ scalaire dans le volume, et ils proposent une visualisation stéréoscopique de ces techniques. Pour améliorer la perception de la profondeur relative des structures visibles, Bruckner et Gröller [BG07] introduisent un calcul de halos. Certaines techniques permettent de mettre en valeur des régions de l'espace choisies par l'utilisateur, en calculant une mesure d'impor-



DVR MIP

FIG. 3.3: Exemple de rendu volumique direct ordonné et de MIP.

tance [VG05] ou de saillance [KV06] sur le champ scalaire, celle-ci étant intégrée dans la fonction de transfert. L'équation physique de rendu peut également être transformée localement en fonction des données représentées : Marchesin et al. [MDM07] modifient l'opacité pour chaque pixel en fonction du nombre de voxels contribuant à la couleur finale le long du rayon correspondant, tandis que Kraus [Kra05] propose une représentation basée sur les variations dans les données le long d'un rayon qui garantit une indépendance du rendu aux échelles spatiales.

1.5 Evaluation

Pour commencer, notons que le rendu volumique est majoritairement développé à des fins de visualisation médicale, même si diverses applications de cette technique sont proposées dans d'autres disciplines, par exemple en mécanique des fluides. Par conséquent, les retours d'expérience quant aux apports du DVR proviennent essentiellement de la communauté médicale[2]. Lorsque plusieurs cliniciens sont interrogés sur l'intérêt du DVR [MZH+05], ils soulignent que les rendus volumiques sont de plus en plus utilisés dans leurs pratiques quotidiennes, même si les difficultés inhérentes à la mise en œuvre de l'ensemble de la chaîne d'acquisition/rendu et à la définition des fonctions de transfert pour des données d'IRM par exemple freinent ce développement. Cependant, l'un d'entre eux souligne aussi la difficulté à discriminer dans les images produites des objets d'intensités lumineuses proches, les contrastes de luminance fournissant une information limitée. En dehors du monde médical, Hibbard [Hib00] note qu'en dépit de l'implémentation du DVR dans le logiciel de visualisation Vis5D, les météorologues qu'il a rencontrés n'utilisent que rarement cette technique, préférant recourir aux isosurfaces et plans de coupe.

En général, les diverses techniques de rendu volumique sont essentiellement évaluées sur la base de la qualité apparente des images produites, le facteur esthétique [GMS+07] ou la précision du rendu [MHB+00, WQW+06] dominant la capacité à effectivement percevoir des structures dont l'agencement spatial est non ambigu dans l'image. Par exemple, Giesen et al. [GMS+07] ont réalisé une étude à grande échelle visant à comparer différents jeux de paramètres, le critère de sélection étant la qualité perçue des images générées dans chaque cas. Cependant, il existe très peu d'études évaluant quantitativement les apports et limites du rendu volumique au niveau perceptif. Kersten et al. [KSTE06] comparent l'influence de la perspective atmosphérique et de la stéréo sur la perception de la profon-

[2] Ainsi, les utilisateurs de visualisation scientifique d'EDF R&D rencontrés au cours de ce travail n'avaient pas connaissance du DVR, notamment car cette technique n'était pas implémentée dans les logiciels à leur disposition.

deur dans des rendus X-RAY, et Ropinski et al. [RSH06] évaluent l'apport de techniques non photoréalistes améliorant le rendu MIP pour l'angiographie[3]. A notre connaissance, aucune investigation n'a été réalisée quant à la perception de la profondeur dans des rendus volumiques ordonnés.

2 Perception de la transparence

2.1 Conditions d'occurence

Pour le système visuel, la transparence est un phénomène complexe à appréhender. En effet, il doit séparer une unique luminance perçue en plusieurs composantes. Ce processus, appelé *scission*, est une forme de segmentation perceptive de la profondeur en plusieurs surfaces non pas voisines mais superposées [FA03]. On peut alors s'interroger sur les distributions de luminances dans un stimulus susceptibles de provoquer une illusion de transparence, sur la capacité du système visuel à évaluer correctement les caractéristiques des surfaces transparentes, ainsi que sur les facteurs déterminant l'ordre en profondeur effectivement perçu par le SVH.

Les premiers travaux importants sur la perception de la transparence ont été menés par Metelli [Met74]. Celui-ci a proposé un modèle additif de génération de transparence qui repose sur un disque comportant un secteur ouvert de surface relative α. Lorsque cet épiscotistère, de réflectance t, est en rotation rapide devant un fond partagé en deux bandes de réflectances respectives a et b, il est en fait perçu comme une surface homogène transparente recouvrant partiellement le fond (figure 3.4). Dans l'image, cette transparente est scindée en deux demi-disques de réflectance p et q vérifiant

$$p = \alpha a + (1-\alpha)t \tag{3.5}$$

$$q = \alpha b + (1-\alpha)t. \tag{3.6}$$

Inversement, en ne considérant que l'image finale, Metelli suggère que le degré de transparence perçu dans le stimulus est déterminé par le rapport

$$\alpha = \frac{p-q}{a-b}, \tag{3.7}$$

à condition que les deux différences soient de même signe. Il est intéressant de noter que le modèle de l'épiscotistère correspond aux équations de l'*alpha-blending* utilisé en informatique graphique [PD84]. Pour mieux rendre compte des données expérimentales, d'autres modèles ont été proposés. Gerbino et al. [GSTdW90] remplacent la réflectance par la luminance dans le modèle de Metelli, tandis que Singh et Anderson [SA02, Sin04] montrent qu'un modèle de perception de transparence achromatique reposant sur le contraste de Michelson [4] donne de meilleurs résultats. Si de nombreuses études ont cherché à définir les distributions de luminance générant un percept de transparence dans des rendus achromatiques [KK01, KvDPR08], les composantes de couleurs peuvent être ajoutées aux équations précédentes, et les conditions d'existence de transparence chromatique et achromatique étudiées en conséquence (voir par exemple [FSM06]).

En dehors de toute notion quantitative, l'existence d'une perception de transparence dans une image comme celle de la figure 3.4 est soumise à deux contraintes essentielles : la surface transparente doit engendrer une réduction des contrastes ($|p-q| < |a-b|$), et respecter la polarité des contrastes le long des contours ($a-b$ et $p-q$ de même signe). Anderson [And97] exprime cela en une règle simple : « Quand deux contours alignés subissent une variation discontinue de l'amplitude

[3] Technique d'imagerie médicale permettant de visualiser les vaisseaux sanguins non accessibles par radiographie standard.

[4] Le contraste de Michelson entre deux surfaces de luminance a et b vaut $\frac{|a-b|}{a+b}$.

du contraste, mais préservent la polarité du contraste, la région la moins contrastée est décomposée en deux surfaces causales ».

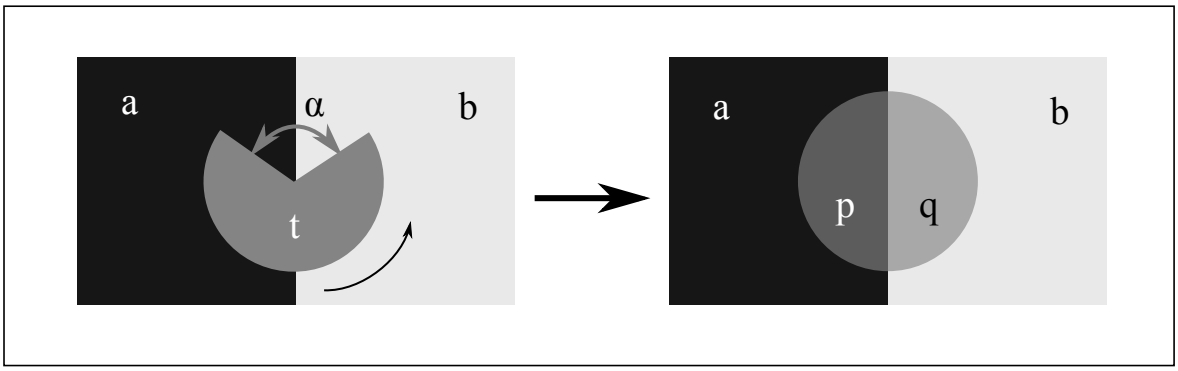


FIG. 3.4: Stimulus original utilisé par Metelli pour estimer la transparence perçue. Pour son modèle, Metelli s'inspire d'un disque présentant un secteur ouvert de surface relative α (image de gauche). La rotation rapide de cet épiscotistère induit un mélange visuel (loi de Talbot) entre sa réflectance et celles des deux couleurs de fond (image de droite). Le taux de transparence perçue dans l'image est déterminé par les relations de luminance dans les régions voisines de l'image.

2.2 Importance des jonctions

Une condition essentielle à l'existence d'une perception de transparence est la présence dans l'image de jonctions de contours [And97]. Un contour est marqué par une discontinuité de luminance entre régions adjacentes. Les jonctions marquent les croisements entre les contours des différents objets présents dans l'image. Lorsqu'elles sont occultées, la probabilité de percevoir de la transparence est nettement réduite [KK02].

Suivant les configurations, différents types de jonctions peuvent être rencontrés (voir figure 3.5). Les jonctions en X et en Ψ sont entourées de quatre régions, avec deux paires de contours colinéaires dans le premier cas et seulement une pour les jonctions en Ψ. Dans les jonctions en T, seulement trois régions apparaissent.

Les jonctions en X sont les plus susceptibles de susciter une perception de transparence. Les jonctions en T correspondent normalement à des situations d'occultation, mais peuvent également révéler une certaine transparence, comme l'ont montré [And97] et [Mas06]. Le cas des jonctions en Ψ, étudié par Singh et Hoffman [SH98] avec des stimuli proches de ceux de la figure 3.5 (droite), est plus intéressant. En effet, Singh et Hoffman relèvent que plus l'angle de la jonction s'éloigne de celui d'une jonction en X, plus les deux éléments centraux ont tendance à apparaître comme deux parties opaques distinctes. Ils expliquent ce résultat par deux principes d'organisation de la perception. Le principe de généricité stipule qu'une interprétation non stable par une légère perturbation du stimulus a peu de chance d'être réalisée, ce qui correspond ici à l'alignement de l'angle de l'objet « avant » avec le contour vertical central. La règle des minima, quant à elle, correspond à la séparation réalisée par le SVH d'un objet en plusieurs parties : la saillance des différents éléments est d'autant plus importante que la courbure au niveau de leur liaison est importante, ce qui correspond ici à une augmentation de l'angle de la jonction en Ψ. Ces observations montrent que la perception de transparence est fortement liée à la perception et la reconnaissance d'objets et de parties, ce sur quoi nous reviendrons dans la section suivante.

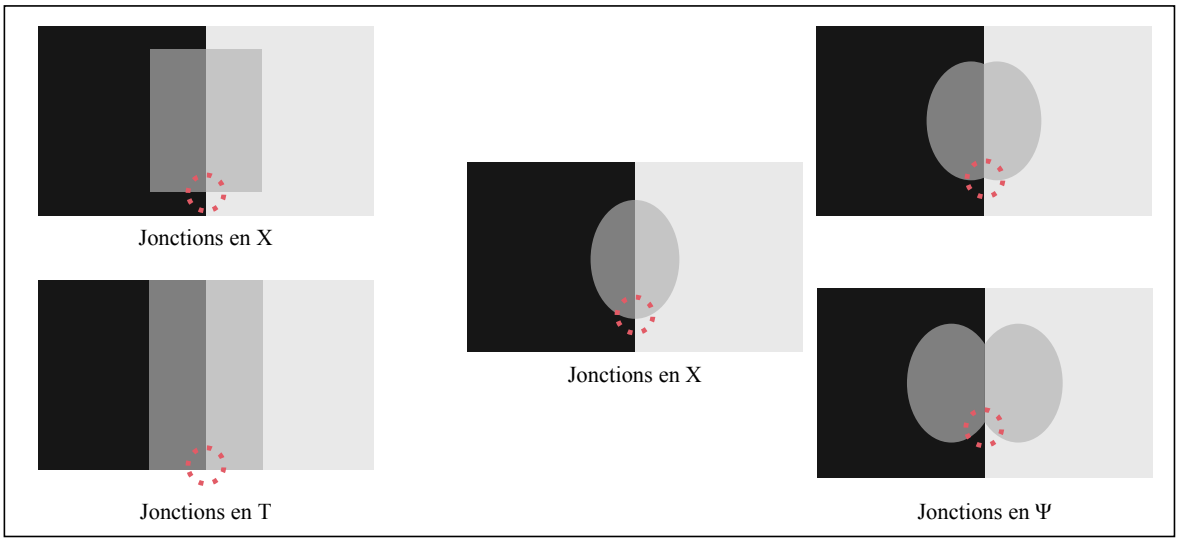


FIG. 3.5: Transparence phénoménologique : importance des jonctions de contours. Les jonctions en X sont les plus à même de générer une impression de transparence, même si une telle perception peut être observée également avec des jonctions en T ou en Ψ.

2.3 Transparence et ordre en profondeur

La perception de la transparence d'un objet est intrinsèquement liée à la notion de séparation en profondeur, puisqu'elle implique que l'on voit une surface à travers une autre.

Lorsque deux surfaces se recouvrent partiellement dans l'image, Adelson et Anandan [AA90] ont montré que la disposition des inversions de contrastes au niveau des jonctions en X détermine d'une part l'existence d'une perception de transparence, et d'autre part impose l'ordre en profondeur des différentes surfaces (figure 3.6), ce que l'on peut retrouver en traçant une ligne brisée allant de la région la plus claire à la plus sombre. Une forme en « Z » induit une perception bistable, chacune des deux surfaces pouvant être vue alternativement comme transparente et au-dessus de l'autre. Une forme en « C » produit une séparation claire, une seule surface pouvant être vue comme non opaque et plus proche de l'œil, tandis qu'un croisement dans le tracé est incompatible avec une perception de transparence.

Ce stimulus a été repris par Kersten et al. [KBSK92] dans une expérience de perception de structure par le mouvement (SfM) : les deux rectangles sont positionnés à deux profondeurs légèrement différentes, celui du bas étant plus proche de l'observateur, et mis en rotation autour d'un axe vertical placé à une distance intermédiaire entre les deux. Lorsque les deux surfaces sont opaques, le mouvement de rotation en accord avec la structure 3D du stimulus est correctement perçu par les observateurs. Mais quand la luminance de la région d'intersection dans l'espace image est modifiée de manière à faire apparaître la surface distance comme occultant la plus proche, la rotation se transforme en une impression de glissement des deux surfaces l'une sur l'autre. Et si les deux surfaces peuvent apparaître toutes deux comme transparentes, Kersten et al. relèvent une perception bistable, le changement de mouvement 3D perçu s'accompagnant instantanément d'une modification des propriétés des deux surfaces.

Afin de mieux comprendre l'interaction entre perception de la transparence et organisation spatiale, Koenderink et al. [KvDPR08] travaillent sur une variante du stimulus de la figure 3.4, le disque central étant remplacé par un carré. Mais au lieu de s'intéresser seulement au niveau de transparence perçu, ils étudient la configuration spatiale vue par les sujets, le nombre de possibilités étant plus important qu'un simple carré central transparent ou une mosaïque de surfaces opaques : par exemple,

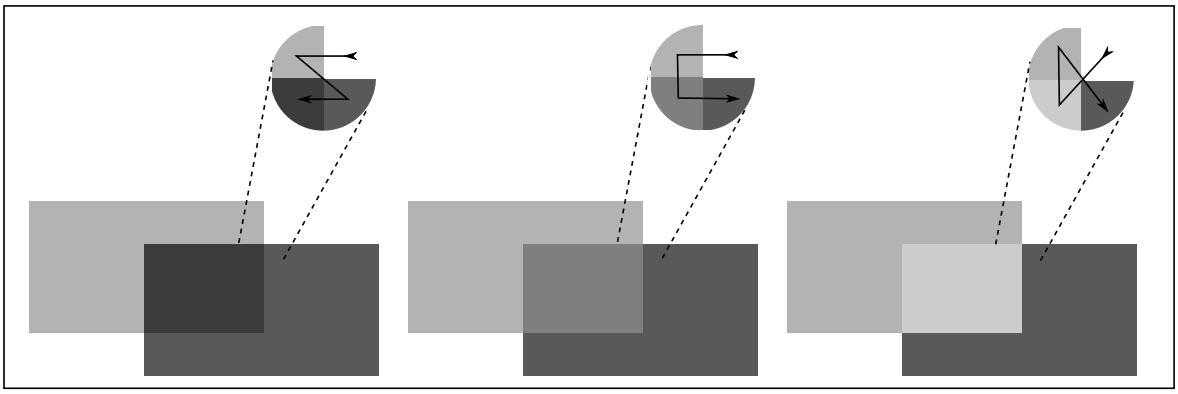


FIG. 3.6: L'ordre en profondeur perçu pour deux objets transparents se recouvrant partiellement est déterminé par les inversions de contrastes au niveau d'une jonction en croix [AA90, FA03]. Ces inversions peuvent être représentées par une ligne allant de la région la plus claire à la plus foncée. Si ce trajet a la forme d'un « *Z* » (gauche), les deux surfaces peuvent être vues comme se trouvant au-dessus (perception bistable), alors que s'il a la forme d'un « *C* » l'ordre en profondeur est clairement déterminé (centre). Une forme en « α » est incompatible avec une perception de transparence (droite).

en modifiant les luminances des différentes régions, il est possible de faire apparaître une bande verticale transparente à droite ou à gauche, ou une surface transparente comportant un trou central. Leurs résultats sont globalement en accord avec les prédictions réalisées à partir du modèle théorique de Metelli et de principes d'organisation de la perception tirés de lois de la Gestalt, tels que la simplicité des formes perçues ou la stabilité de la perception en cas de légère modification du stimulus. Cependant, ils observent des déviations nécessitant d'autres hypothèses, la présence d'un trou dans une surface transparente étant notamment moins souvent perçu qu'attendu.

Si la distribution des luminances dans le stimulus peut servir d'ancrage à la perception de transparence en soi et guider l'organisation de la perception en profondeur des différentes surfaces, il ne faut pas perdre de vue que d'autres facteurs peuvent intervenir à ce niveau. Comme le soulignent Fleming et Anderson [FA03], il s'agit d'un jeu à double sens : une scission de la profondeur peut engendrer une modification des luminances apparentes, mais l'inverse est aussi vrai : une modification des luminances perçues peut altérer la perception de la superposition en profondeur des surfaces observées. Ainsi, la reconnaissance d'objets familiers ou la préférence pour des surfaces convexes ou simplement non creusées sont des critères susceptibles de modifier fortement notre perception en profondeur dans des rendus transparents.

3 Expérience 1 : perception statique en DVR

Cette première expérience vise à étudier la capacité du système visuel humain à estimer correctement l'arrangement spatial d'objets volumiques rendus par DVR, le point de vue et les éléments de la scène demeurant fixes pendant la présentation. Les participants doivent déterminer, parmi trois choix possibles, l'ordre en profondeur de deux cylindres semi-transparents. Le stimulus sélectionné est volontairement simple, de sorte que la superposition ordonnée de transparences constitue l'unique indice de profondeur dans l'image.

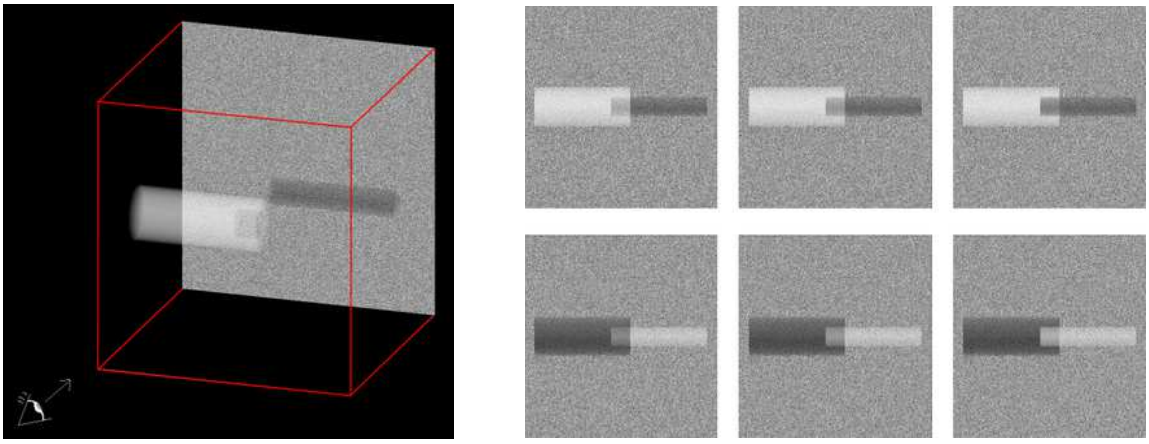


FIG. 3.7: Stimulus mis en place pour étudier la perception de la profondeur dans une scène transparente statique. Les deux cylindres pleins sont disposés soit à deux profondeurs distinctes, soit au même niveau, le plus gros des deux étant évidé à son extrémité pour s'emboîter avec le plus petit. (a) Vue de 3/4, le gros cylindre étant devant le petit. (b) Les six stimuli principaux utilisés au cours de l'expérience, le gros cylindre étant alternativement clair (ligne du haut) ou foncé (ligne du bas), et devant le petit (colonne de gauche), derrière lui (droite) ou au même niveau (milieu).

3.1 Dispositif expérimental

3.1.1 Description du stimulus

La scène est composée de deux cylindres de même longueur mais dont les rayons respectifs varient du simple au double (voir figure 3.7). En vue de face, ces deux cylindres sont disposés horizontalement, et sont centrés verticalement dans l'image. Ils occupent deux positions latérales symétriques, l'un à gauche et l'autre à droite de l'image, et se recouvrent partiellement, la zone de superposition dessinant un carré au centre de l'écran. Trois dispositions sont implémentées : le gros cylindre (LC) devant le petit (condition « Gros devant »), le petit cylindre (SC) devant le gros (condition « Petit devant »), ou les deux cylindres à la même profondeur (condition « Emboîtés »). Dans ce dernier cas, l'extrémité du petit cylindre se trouve à l'intérieur de LC.

Les deux objets sont pleins, leur densité étant représentée par un scalaire s qui les fait apparaître clair ($s = 1$) ou foncé ($s = 0.2$). Dans la scène, tous les voxels extérieurs à SC et LC sont considérés comme vides ($s = 0$). Afin que la partie centrale de l'image corresponde toujours à la même quantité de matière accumulée par transparence, quelle que soit l'organisation en profondeur des deux objets, un trou cylindrique ($s = 0$) de même rayon que SC est creusé à l'extrémité du gros cylindre.

Les cylindres sont rendus en projection orthographique, ce qui supprime les variations de leur taille apparente d'une condition de profondeur à l'autre, cette information étant de nature à faciliter la résolution de la tâche expérimentale. Cela garantit également que les informations de contour sont limitées à quelques jonctions en X, celles-ci constituant un élément clé de la perception de transparence (voir sections 2.2 et 2.3). Cependant, des essais préliminaires ont fait ressortir le besoin de renforcer l'apparence volumique des cylindres sombres ($s = 0.2$). Pour ce faire, un pseudo-éclairage est simulé dans les volumes en modulant la densité de chaque voxel x :

$$s(x) = s_0 + a\vec{R(x)}.\vec{L} \tag{3.8}$$

où $\vec{R(x)}$ désigne la direction normalisée de x par rapport à l'axe du cylindre et $\vec{L}$ la direction d'éclairage

($s_0 = 0.2$ ou $s_0 = 1$, $a = 0.1$, et la lumière vient du haut). Les valeurs de $s(x)$ sont restreintes à l'intervalle $[0;1]$.

Pour réaliser le rendu, les données volumiques ont été précalculées sur une grille régulière et sont chargées dans une texture 3D au moment de leur affichage. Le rendu volumique est réalisé au moyen de cent plans régulièrement espacés, les images produites étant accumulées de l'arrière vers l'avant. La fonction de transfert (voir section 1.2 p.66) choisie est linéaire en luminance et constante en opacité :

$$phi : s \rightarrow (s, 0.25) \tag{3.9}$$

Afin de réduire l'influence de la contribution de la couleur du fond sur la perception de la profondeur, une image contenant un bruit blanc par pixel est placée à l'arrière plan. Ce bruit est défini dans l'intervalle $[0.2 : 1]$, ce qui correspond approximativement à la distribution des luminances induites par les rendus des deux cylindres. De plus, ce bruit permet de forcer la transparence apparente des deux objets. En effet, dans un stimulus comme celui de la figure 3.6, Kersten et al. [KBSK92] ont noté que les participants choisissent toujours d'interpréter l'image comme ne contenant qu'une seule surface transparente, la deuxième apparaissant opaque et occultant le fond de la scène.

Au cours de l'expérience, trois facteurs sont manipulés : l'ordre en profondeur des deux cylindres (3 conditions), leur arrangement latéral (2 conditions, le gros à droite et le petit à gauche et inversement) et leurs luminances respectives (2 conditions, SC clair et LC foncé et inversement). Il y a donc au total 12 conditions différentes, la moitié d'entre elles pouvant être observées sur la figure 3.7(b). Dans tous les cas, les inversions de contrastes au niveau des jonctions en X suivent un trajet en « Z » : d'après les règles décrites en section 2.3, les stimuli induisent donc bien une perception de transparence pour chacun des deux cylindres, ce qui génère a priori une ambiguïté concernant leur disposition en profondeur.

3.1.2 Déroulement de l'expérience

Le protocole principal correspond à un choix forcé à trois alternatives. Pour chaque image, il est demandé au participant de déterminer l'ordre en profondeur des deux cylindres, parmi les trois réponses proposées : « GAUCHE DEVANT », « IMBRIQUES », « DROIT DEVANT ». Aucun retour n'est fourni aux participants sur les réponses données.

Au cours de l'expérience, les fixations oculaires des sujets sont enregistrées au moyen d'un capteur de regard EyeLink®II, qui permet également de s'assurer qu'ils regardent effectivement la scène. Les sujets sont assis à 80 cm d'un écran LCD de 19 pouces (résolution 1280x1024), la tête maintenue immobile au moyen d'un repose-menton. L'image du stimulus, un carré de 800 pixels de côté, occupe approximativement 17 ° d'angle visuel, et l'intersection des deux cylindres 2 ° . Pour chaque participant, un calibrage du capteur de regard est réalisé, et l'expérience peut commencer.

Pendant une session expérimentale, les douze configurations possibles sont présentées chacune cinq fois, l'ordre des 60 essais correspondants étant généré aléatoirement. A chaque essai, une croix rouge sur fond blanc apparaît d'abord au centre de l'écran, pour une durée aléatoire d'une à trois secondes. Il est demandé au participant de fixer le centre de la croix. Ensuite, le stimulus est présenté pendant deux secondes. Il est suivi par un écran de réponse, qui permet au sujet de faire son choix en cliquant sur l'une des trois réponses affichées à l'écran. Une fois la réponse sélectionnée, l'essai suivant est enclenché. Les positions du regard sont enregistrées pendant chaque présentation du stimulus, et les réactions et impressions du sujet sont recueillies à la fin de la session.

De plus, un petit test est mené après l'expérience pour déterminer si les différentes configurations génèrent des images perceptivement distinctes en termes de luminances. Pour chacune des deux conditions de densité (LC avec $s_0 = 0.2$ et $s_0 = 1.$), les carrés centraux correspondant aux trois configurations de profondeur sont isolés et placés côte à côté. Il est alors demandé au sujet de les classer,

du plus clair au plus foncé. Tous les participants ont réalisé cette tâche correctement et sans temps de réflexion. Par conséquent, les difficultés potentielles à réaliser la tâche principale, à savoir l'estimation de l'arrangement spatial des cylindres, ne peuvent trouver leur origine dans des problèmes de discrimination de contrastes.

La présentation de chaque stimulus dure 2 secondes exactement. Un temps plus long de 5 secondes a été expérimenté dans une étude pilote avec deux sujets. Ceux-ci ont reporté avoir utilisé tout ce temps pour rechercher des indices dans l'image leur permettant de consolider leur choix, ce qui au final n'a fait qu'augmenter leur confusion au fur et à mesure que l'expérience avançait. Nous avons donc réduit le temps de présentation à 2 secondes, afin de renforcer la part de perception dans la décision finale au détriment du raisonnement conceptuel, le but de l'étude étant de déterminer la capacité du SVH à percevoir une organisation en profondeur correcte d'objets semi-transparents. Les entretiens réalisés avec les participants à la fin de l'expérience ont montré qu'ils n'avaient effectivement pas eu assez de temps pour élaborer des stratégies complexes reposant sur une exploration exhaustive de l'image ou sur la comparaison d'images consécutives. Par ailleurs, des essais informels ont montré que des temps de présentation trop réduit, de l'ordre de quelques dixièmes de seconde, généraient une confusion trop importante. Les images présentées se révélaient confuses et il était plus que difficile de prendre une décision. Par conséquent, une durée de deux secondes est apparue comme un compromis acceptable entre ces deux extrêmes.

3.1.3 Participants

Dix sujets, âgés de 22 à 39 ans et naïfs quant aux objectifs de l'étude, ont participé à l'expérience. Parmi eux, deux avaient un peu d'expérience quant à l'utilisation des techniques de rendu volumique. Tous avaient une vision normale, éventuellement après correction. Par ailleurs, quatre sujets choisis au hasard parmi les dix ont passé l'expérience une deuxième fois après quelques minutes de repos, l'ordre de présentation des stimuli ayant été modifié.

Avant de passer la séquence de 60 essais, les sujets lisaient le protocole expérimental sur lequel une vue de trois-quart de la scène était représentée (figure 3.7 (a)). Ils s'entraînaient ensuite sur quelques essais dont l'ordre différait de ceux de la séquence principale, sans qu'aucun retour sur la qualité de leurs réponses ne leur soit donné. Enfin, la vue de trois-quart était à nouveau présentée avant de commencer. Tout ceci permettait de s'assurer autant que possible que les sujets avaient correctement compris la tâche expérimentale, et qu'ils garderaient en tête les propriétés volumiques des objets affichés.

3.2 Résultats

3.2.1 Performances des sujets

Les taux de bonnes réponses sont calculés pour chaque participant, par condition et pour l'ensemble de l'expérience. Dans le cas des quatre sujets ayant passé deux sessions, on considère la moyenne des réponses sur les deux sessions, ce qui fait donc au final dix feuilles de résultats.

Les résultats moyens sont illustrés sur la figure 3.8. Dans cette tâche où les sujets devaient classer deux objets en fonction de leurs profondeurs respectives, les performances obtenues sont faibles, avec seulement 48,1% de bonnes réponses en moyenne, même si elles apparaissent significativement supérieures au niveau du hasard qui se situe ici à 33% (test t de Student, $p < 0.01$). Une analyse plus en détail des résultats en fonction des trois configurations de profondeur montre des taux de bonnes réponses respectivement de 71%, 44% et 29.3% pour les cas « Gros devant », « Emboîtés » et « Petit devant ». Ces résultats sont au-dessus du niveau du hasard lorsque le gros cylindre est présenté devant

(significatif, $p < 0.01$) ou quand les deux s'intersectent (peu significatif, $p = 0.14$), mais quasiment au même niveau quand le petit cylindre est en avant de la scène ($p = 0.55$).

Une première explication de ces résultats peut être trouvée en analysant les moyennes des réponses données par les participants (figure 3.9). En effet, il apparaît qu'au cours de l'expérience les sujets ont une nette préférence pour voir le gros cylindre en avant (48.4% de l'ensemble des réponses) ou les deux cylindres emboîtés (40.4%). En revanche, ils sont peu disposés à imaginer le petit cylindre plus proche de l'œil (seulement 11.2% des réponses). Nous avons par ailleurs examiné l'importance de la couleur respective de chacun des deux objets. La figure 3.10 montre que les performances des participants lorsque le gros cylindre est clair ou foncé sont très proches les unes des autres.

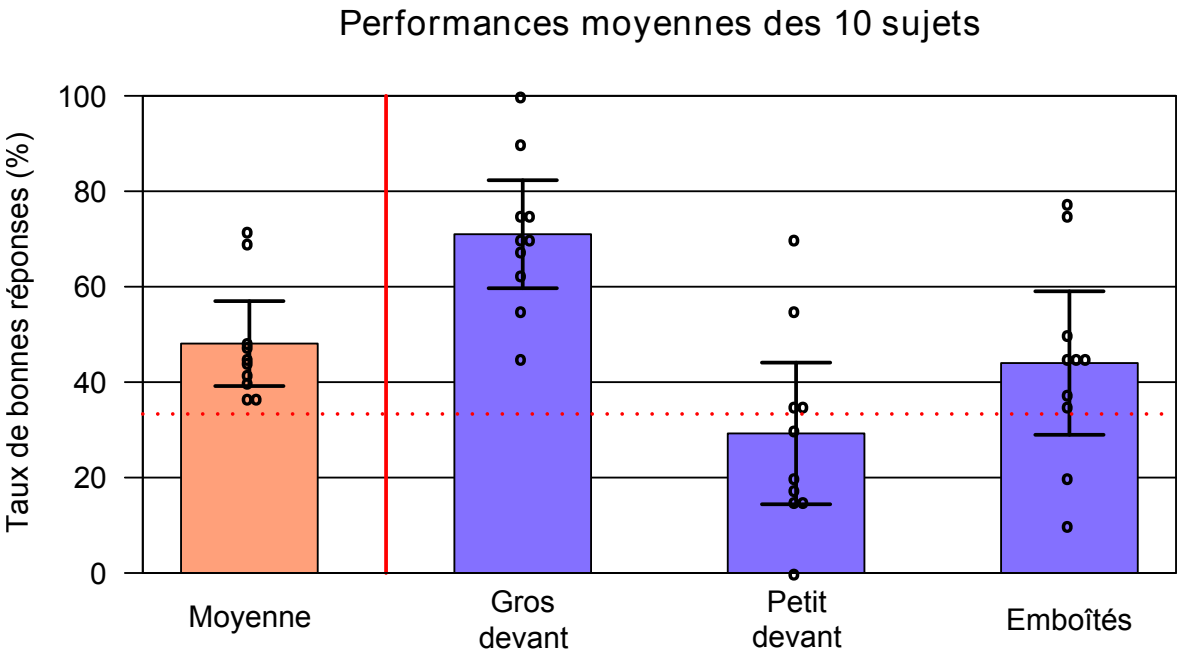


FIG. 3.8: Résultats moyens des 10 sujets concernant la perception de la profondeur dans un rendu transparent statique. La ligne rouge en pointillés indique le niveau correspondant à des réponses qui seraient données au hasard. Les barres d'erreurs indiquent les intervalles de confiance à 5%, les cercles les moyennes de chacun des 10 participants. Les performances des sujets sont assez faibles, les difficultés étant plus importantes quand le petit cylindre se trouve en avant plan ou au niveau du plus gros.

3.2.2 Analyse des fixations

A première vue, on pourrait supposer que le faible nombre de réponses indiquant la perception du petit cylindre en avant plan s'explique par le fait que le gros cylindre, occupant deux fois plus de place dans l'image, est plus à même d'attirer l'attention des sujets, ce qui pourrait forcer le choix de la réponse « Gros devant ». Cependant, une analyse détaillée des trajectoires oculaires des participants pendant la session expérimentale révèle qu'en moyenne plus de saccades sont dirigées vers le petit cylindre que vers le gros. Principalement quatre types de trajectoires sont observées : des fixations limitées à la zone d'intersection (F_{inter}), une exploration du centre vers le petit cylindre (F_{petit}) ou du centre vers le gros cylindre (F_{gros}) et enfin des saccades alternées entre les deux objets (F_{double}). Des exemples de chacune de ces catégories sont présentés sur la figure 3.11. Pour neuf participants sur les dix, F_{petit} est plus fréquent que F_{gros}, avec un rapport moyen de 3 à 1 (sur 60 essais, $\overline{F}_{petit} = 15.1$ et

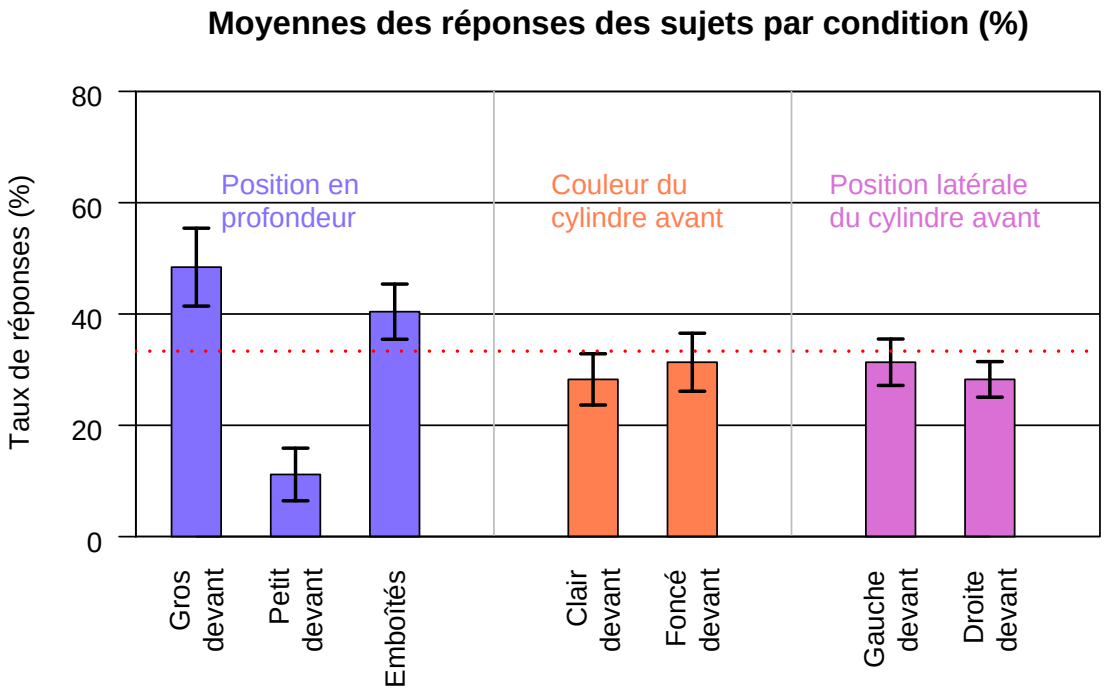


FIG. 3.9: Réponses moyennes des sujets en fonction des différents facteurs. La ligne rouge en pointillés indique une répartition égale entre les différents niveaux de chaque facteur ce qui peut correspondre au cas idéal de succès total dans la tâche demandée, mais également à une répartition purement aléatoire des réponses. La principale information est la nette préférence des sujets à percevoir le gros cylindre en avant plan, et dans une moindre mesure les deux cylindres emboîtés, plutôt que le petit devant. En revanche, la couleur ou la position latérale des cylindres n'ont pas d'incidence sur les réponses observées.

$\overline{F}_{gros} = 5.2$). Les trajectoires F_{double} sont en moyenne aussi fréquentes que F_{gros}, avec des variations importantes selon les sujets. Enfin, les fixations limitées à la partie centrale de l'image représentent approximativement la moitié des essais. Dans ce cas, le regard est dirigé soit vers l'un des contours latéraux des cylindres, soit vers l'une des jonctions en X.

3.2.3 Rôle éventuel de l'apprentissage

Les stimuli étant répétés cinq fois pour chacune des 12 conditions, on ne peut exclure a priori que les participants s'habituent aux images proposées et apprennent à distinguer les différentes configurations. Ils pourraient alors répondre à la tâche principale en s'appuyant sur un raisonnement utilisant par exemple des comparaisons entre paires d'images. Cela fausserait les performances enregistrées, qui relèveraient alors au moins autant de critères cognitifs que perceptifs.

La figure 3.12 illustre l'évolution moyenne des performances des 10 sujets au fur et à mesure des 60 essais. En la présence d'un effet d'apprentissage important, on s'attendrait à observer une courbe globalement convexe. Or ici la progression est linéaire, ce qui révèle une certaine constance dans le taux de bonnes réponses, quelle que soit l'avancée dans la session. La figure 3.13 montre quelques différences suivant les sujets, certains s'améliorant en deuxième partie de session (par exemple *AB* ou *AO*) alors que d'autres au contraire voient leurs performances diminuer à mesure que le temps passe

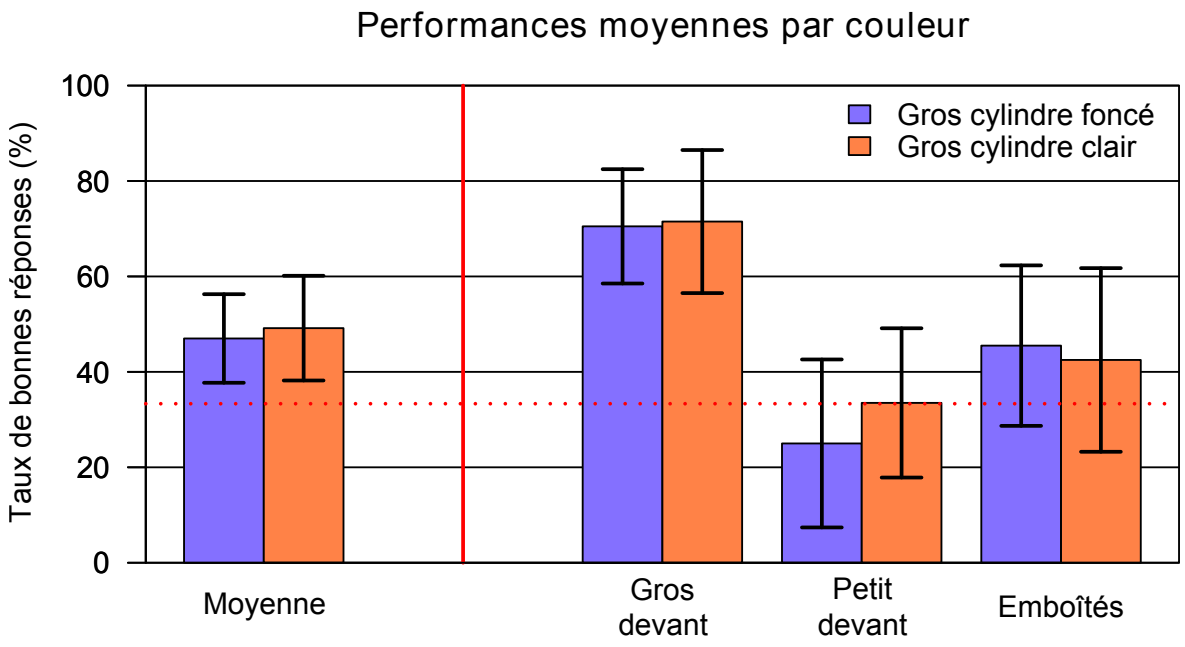


FIG. 3.10: Moyennes des bonnes réponses des dix sujets en fonction de la couleur des cylindres. Ce facteur n'influence pas la perception de la profondeur relative dans les stimuli présentés.

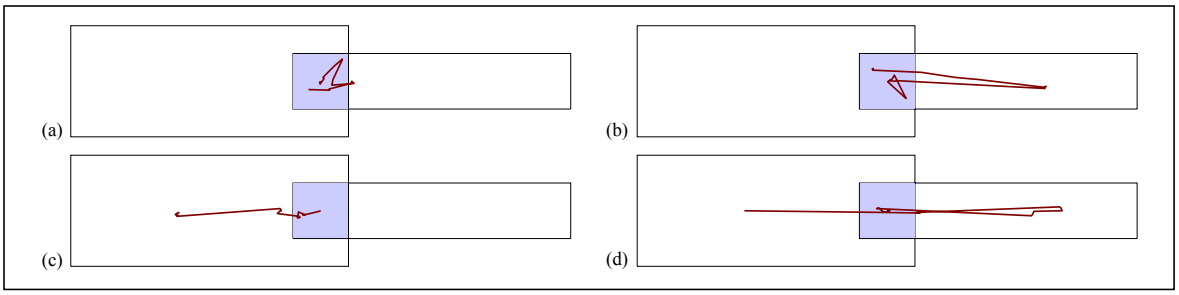


FIG. 3.11: Exemples d'enregistrements illustrant les principaux profils de fixations oculaires observés au cours de l'expérience. Les sujets explorent principalement l'un des deux cylindres (le gros en (a), F_{gros}, ou le petit en (b),F_{petit}), se limitent à leur intersection dans l'espace image (c), F_{inter}, ou sautent de l'un à l'autre (d), F_{double}.

(FF,SR). Ceci peut s'expliquer éventuellement par une certaine fatigue suite à la présentation répétée des différentes images, mais également par la confusion engendrée par la tâche expérimentale, jugée globalement difficile par l'ensemble des participants. Cette confusion, qu'elle se produise au début ou à la fin de la session, pourrait ici jouer un rôle inverse de celui de l'apprentissage, en réduisant les performances dans la tâche de perception de la profondeur.

Par ailleurs, quatre sujets ont passé une deuxième session expérimentale après quelques minutes de repos, avec un ordre différent des 60 essais. La comparaison des moyennes de leurs bonnes réponses montre une légère progression entre les deux sessions (+6.2% en moyenne). Cependant, une analyse détaillée montre que les progressions individuelles sont principalement enregistrées dans les configurations spatiales pour lesquelles le sujet présentait des performances inférieures à la moyenne. Il est intéressant de noter que l'un des quatre participants a montré des résultats meilleurs que la moyenne pendant la première session, mais qu'il ne s'est pas amélioré d'une session à l'autre (-1.7%).

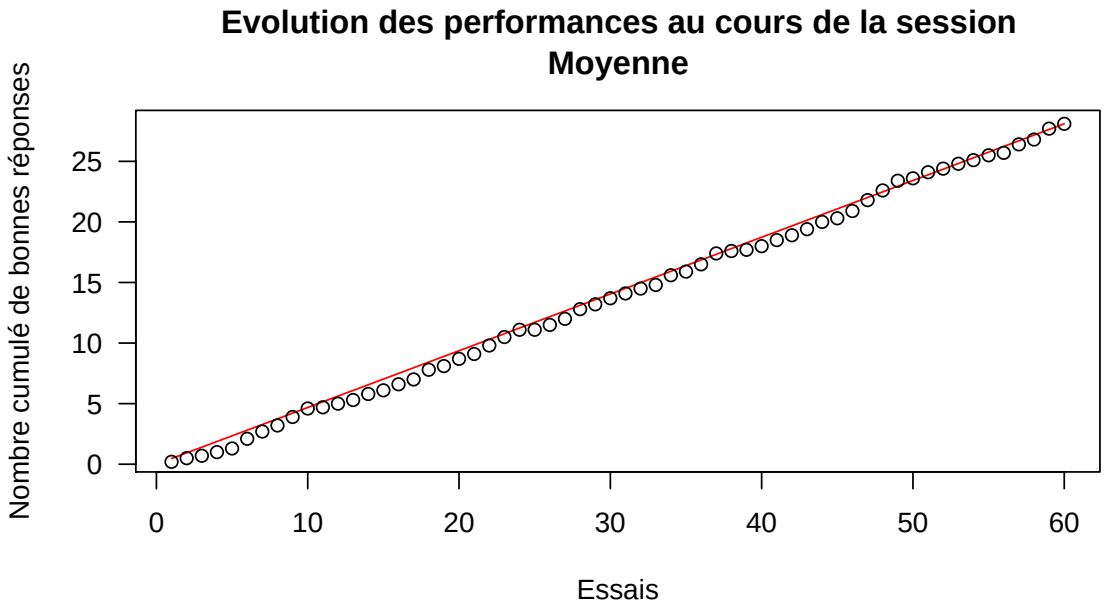


FIG. 3.12: Evolution du nombre de bonnes réponses au cours des 60 essais constituant une session expérimentale, moyennée sur l'ensemble des 10 sujets. La ligne rouge indique un taux de bonnes réponses constant, reflétant une absence d'évolution dans la capacité à réaliser la tâche demandée. Les profils observés expérimentalement sont très proches de cette tendance.

3.3 Discussion

3.3.1 Résultat principal

Dans cette expérience, nous avons essayé d'évaluer la capacité de notre système visuel à correctement percevoir l'ordre en profondeur d'objets semi-transparents rendus statiquement avec un modèle de type DVR. En visualisation scientifique, l'immobilité de la scène est souvent une conséquence inévitable pour les gros ensembles de données en raison du coût computationnel associé aux méthodes de rendu volumique ordonné.

Nous avons utilisé un stimulus volontairement simple, et sélectionné des conditions de rendu générant un contraste suffisant pour distinguer visuellement les différentes images produites. Cependant, les performances d'ensemble sur cette tâche simple sont relativement faibles, même si les participants ont répondu significativement mieux qu'au hasard. Ces résultats montrent les limites du SVH lorsqu'il s'agit d'appréhender la nature de la superposition d'objets semi-transparents rendus avec un modèle optique approché. De plus cette perception semble fortement guidée par des facteurs autres que la seule combinaison ordonnée des luminances, comme le montre la préférence pour la proximité du gros cylindre.

Les informations visuelles utiles ont ici été limitées aux variations de contrastes aux niveaux des contours et intersections des deux objets, essentiellement dans la zone centrale de l'image. Les jonctions en X, éléments importants pour la perception de transparence, apparaissent clairement à l'intersection du bord vertical du gros cylindre et des bords horizontaux du plus petit. La projection orthographique élimine tous les indices géométriques potentiels. La simplicité du stimulus a permis de réduire l'influence d'éventuelles stratégies cognitives dans la prise de décision. Dans des scènes de visualisation complexes, le nombre de jonctions visibles a de grandes chances d'être plus important,

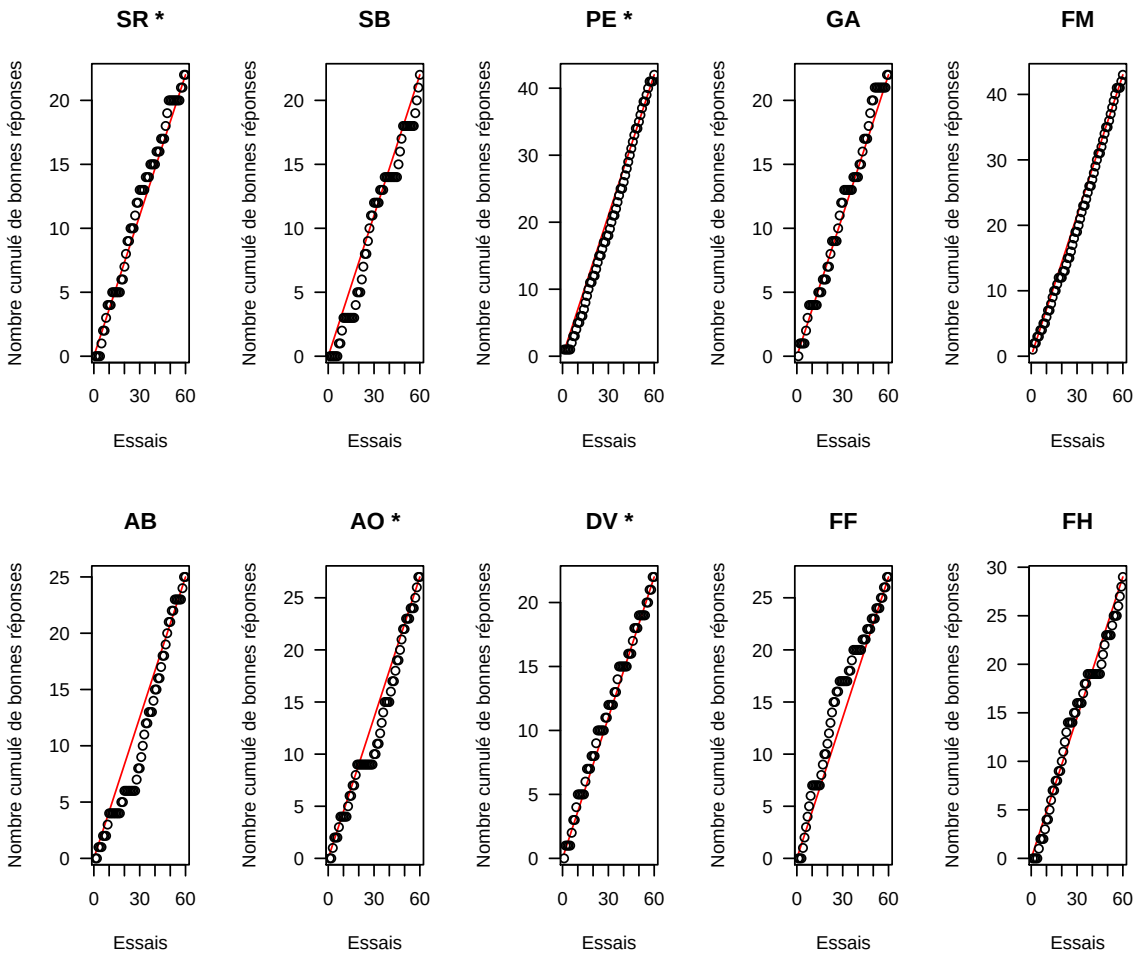


FIG. 3.13: Evolution du nombre de bonnes réponses au cours des 60 essais constituant une session expérimentale pour chacun des 10 sujets. Les performances individuelles ne présentent que quelques légères variations par rapport à la tendance moyenne.

ce qui peut en théorie faciliter la perception de profondeur par transparence. En revanche, la relative continuité spatiale des données dans de telles scènes tend à brouiller ces contours, ce qui affaiblit la perception des objets et de la profondeur.

Par conséquent, un rendu statique en DVR ne semble pas être très efficace pour fournir une compréhension non ambiguë de la structure spatiale de données volumiques, même si l'équation utilisée pour le rendu tient compte de l'ordre en profondeur des différents éléments.

3.3.2 Du plus gros au plus proche

Les réponses des participants dénotent une préférence importante pour voir le plus gros cylindre comme étant plus proche de l'œil (figure 3.9). Celle-ci se manifestait notamment au moment où les sujets devaient cliquer sur la réponse de leur choix, et où dans un certain nombre de cas ils positionnaient la souris sur « Petit devant » avant de finalement sélectionner une autre réponse. Plusieurs causes peuvent expliquer ce biais.

En premier lieu, les deux objets étant de forme identique, cela pourrait révéler un rôle de la

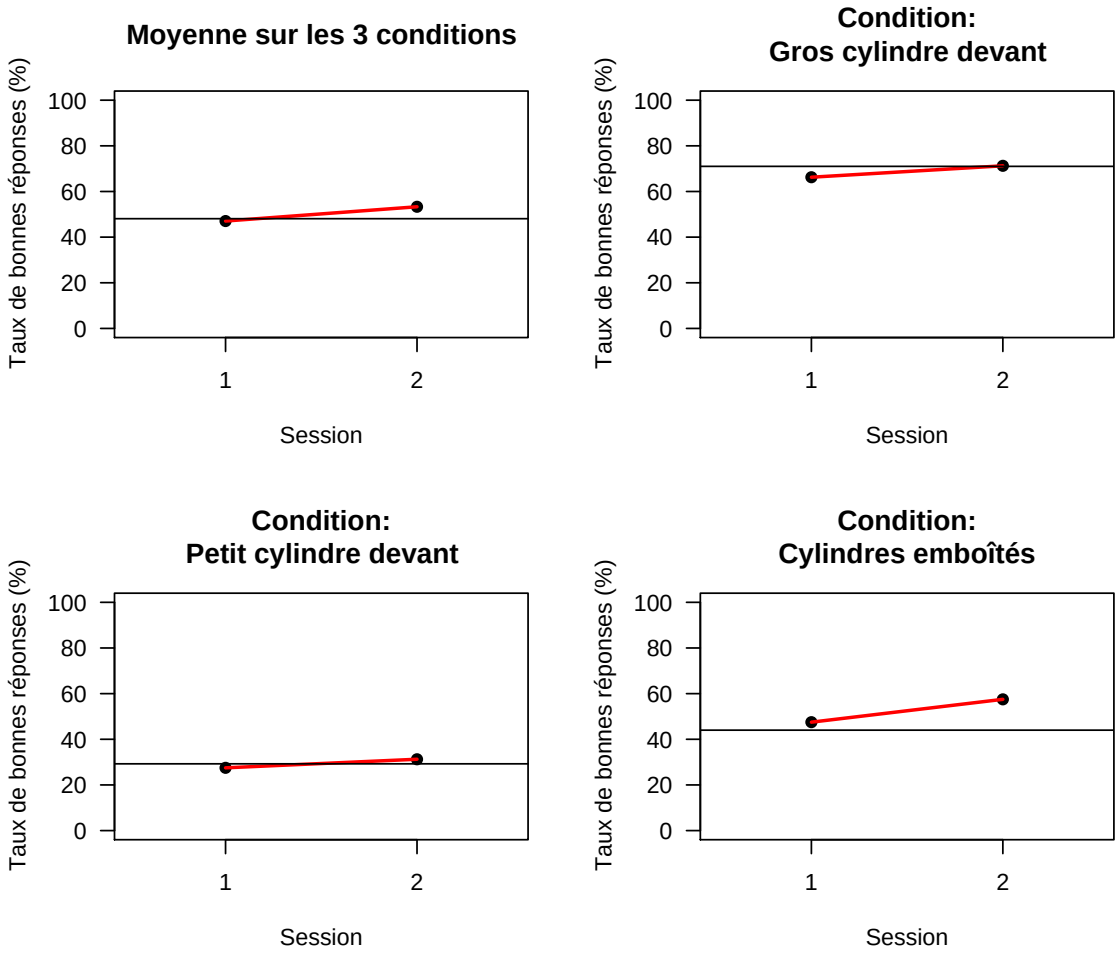


FIG. 3.14: Pour les quatre sujets ayant passé deux sessions expérimentales, performances moyennes comparées entre la session 1 et la session 2. La ligne noire horizontale indique la performance moyenne de l'ensemble des sujets. Globalement, la deuxième session présente des résultats légèrement meilleurs, mais essentiellement dans la configuration où les deux cylindres sont emboîtés. Les sujets ayant passé deux sessions expérimentales sont indiqués par une astérisque dans la figure 3.13.

constance de taille. Dans un environnement naturel, deux objets parfaitement semblables en taille et en forme vont générer des projections rétiniennes de tailles différentes s'ils ne se trouvent pas à la même distance de l'œil. Il a été montré que le SVH utilise cette information picturale pour déterminer la profondeur des éléments dans une image. Par conséquent, on ne peut exclure que les sujets aient parfois utilisé inconsciemment le rapport des tailles des deux cylindres pour prendre leur décision. Malgré tout, nous avons essayé de réduire l'influence de ce biais en affichant à plusieurs reprises des vues de 3/4 de la scène lors de la présentation de l'expérience.

Une autre explication possible réside dans la quantité de transparence perçue, et la simplification potentiellement réalisée par le SVH. En effet, comme dans le stimulus ambigu des deux rectangles superposés (figure 3.6), il est possible que les participants n'aient parfois perçu qu'un seul des deux objets comme transparent, même si la présence de bruit en arrière plan devait limiter cet effet. Dans ce cas, la plus grande surface dans l'image du gros cylindre peut avoir provoqué une asymétrie en sa

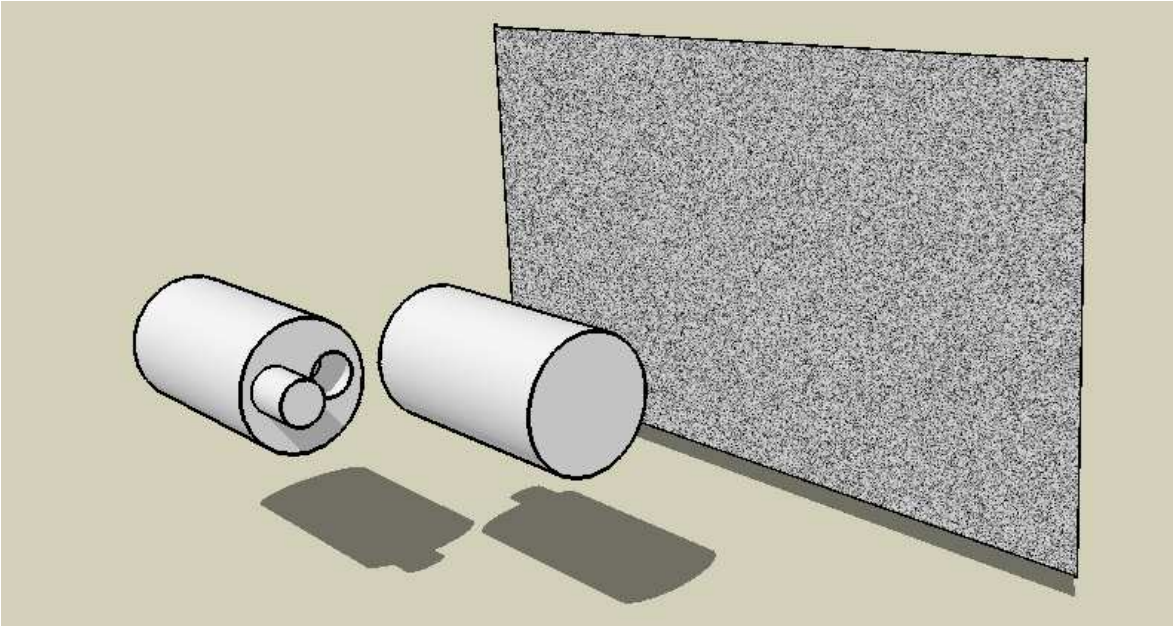


FIG. 3.15: Proposition de modification de l'expérience d'évaluation du DVR en statique. Dans cette expérience, les réponses des sujets révèlent un biais de positionnement du plus gros cylindre en avant plan. Pour éviter cet artefact, il serait possible de modifier la scène, en y intégrant deux cylindres modifiés comme sur l'image ci-dessus, ce qui permettrait de reproduire les trois configurations en profondeur de l'expérience avec des objets parfaitement identiques.

faveur dans cette perception de transparence, le projetant en conséquence en avant de la scène.

Pour éviter ce biais de taille, quel que soit sa cause, il serait envisageable de modifier la scène utilisée pour cette expérience, par exemple en utilisant deux cylindres identiques tels que ceux représentés sur la figure 3.15. Ce choix permet de reproduire exactement les trois configurations étudiées ici, en éliminant le facteur taille. En revanche, l'image apparaîtrait encore plus complexe, rendant la tâche de détermination des profondeurs relatives potentiellement plus délicate à réaliser.

3.3.3 Temps de présentation et de réponse

En limitant la présentation de chaque image à seulement deux secondes, nous avons essayé de restreindre autant que possible l'influence d'approches plus cognitives dans le processus de décision. Cependant, ceux-ci ne peuvent pas être totalement rejetés, comme l'ont montré les discussions réalisées avec les participants à la fin de chaque session. Ainsi, de nombreux sujets ont noté que la présentation d'un stimulus pouvait influencer le choix réalisé sur l'essai suivant, ce qui était possible quand deux essais consécutifs présentaient des configurations proches (uniquement modification de la profondeur des cylindres). Malgré tout, à part deux d'entre eux, tous ont déclaré être peu sûrs de leurs réponses, certaines séquences d'essais semant même le doute dans leur esprit quant aux réponses données précédemment. Le temps de présentation relativement court limitait dans les faits la possibilité de construire des stratégies solides basées sur les classifications des configurations de luminances. Nous ne pouvons bien sûr pas rejeter complètement l'hypothèse que la mémoire à court terme ait pu jouer un rôle dans cette expérience, mais nous pensons que son effet n'a pas dû modifier les performances de façon consistante.

Le temps de réponse, c'est-à-dire le temps écoulé entre l'apparition de l'écran de réponse et le

choix du participant, ont été également enregistrés pour chaque essai, même si nous n'en avons pas tenu compte dans les analyses développées précédemment. Les participants avaient l'instruction de répondre aussi vite que possible, ce qui était généralement le cas (temps de réponse moyen ; 1.9 secondes) malgré l'absence de mécanisme pour les forcer en ce sens. Le temps de réponse moyen ne dénote pas de corrélation linéaire avec les performances globales par sujet ($R^2 = 10^{-4}$).

4 Expérience 2 : DVR et effet cinétique de profondeur

Nous nous intéressons maintenant à l'apport d'informations visuelles dynamiques pour la perception de la profondeur dans des rendus volumiques. Notre hypothèse de travail est que l'effet cinétique de profondeur (KDE) peut permettre de générer une perception non ambiguë des structures spatiales dans les rendus transparents ordonnés. Nous souhaitons également évaluer l'importance du choix de la fonction de transfert sur la profondeur perçue.

Dans une expérience de type KDE, la mise en rotation de la scène suivant un axe non parallèle à la direction de vue génère une perception de profondeur, celle-ci étant directement couplée au sens de rotation ressenti. Si la projection utilisée est orthographique ou correspond à une perspective faible et que l'image contient peu d'indices spatiaux, comme c'est le cas avec un nuage de points peu dense, un observateur regardant la rotation continue de la scène éprouvera souvent des inversions de la direction du mouvement se produisant aléatoirement, et s'accompagnant d'inversions de la profondeur perçue des éléments projetés.

La très grande majorité des travaux relatifs au KDE reposent sur des stimuli constitués de lignes opaques ou de points. L'influence de la perception de transparence sur l'extraction de profondeur par le mouvement a été peu étudiée. Daniel Kersten et al. [KBSK92] ont montré que la perception de transparence pouvait altérer le mouvement tridimensionnel perçu lorsque deux rectangles superposés sont mis en rotation autour d'un axe vertical. Dans le cas de rendus de type X-RAY, Marta Kersten et al. [KSTE06] ont analysé les apports de la vision stéréoscopique et de la perspective atmosphérique pour résoudre l'ambiguïté de perception de profondeur inhérente à ce rendu transparent non ordonné. L'expérience proposée dans cette section s'inspire du travail de [KSTE06], à la différence que nous nous intéressons à la perception de la profondeur évoquée par le mouvement pour des rendus de type DVR, MIP et X-RAY en l'absence d'indices visuels autres que l'accumulation de transparences.

4.1 Protocole expérimental

4.1.1 Description du stimulus

La scène est constituée d'un cylindre vertical tournant autour de son axe à une vitesse de 34 °/s et rendu avec une projection orthographique. La rotation est réalisée vers la gauche ou vers la droite, cette direction étant définie par le mouvement des points les plus proches dans la scène : « vers la gauche » correspond ainsi à une rotation dans le sens horaire lorsque l'objet en vue de dessus. La vitesse de rotation choisie appartient aux plages de vitesses dont il a été montré qu'elles induisent un effet cinétique de profondeur [Gre61, RHS96].

Le volume du cylindre est rempli avec un champ scalaire correspondant à un bruit de Perlin, l'extérieur étant fixé à 0 (voir figure 3.16(a)). Le bruit de Perlin génère des données non homogènes variant continûment mais ne présentant pas de structures reconnaissables, ce qui présente l'avantage de limiter les biais cognitifs influant la perception des observateurs. Le niveau de bruit $P(x)$ d'un voxel x est défini par

$$P(x) = \sum_{i=0}^{n-1} \frac{N(b^i x)}{a^i} \tag{3.10}$$

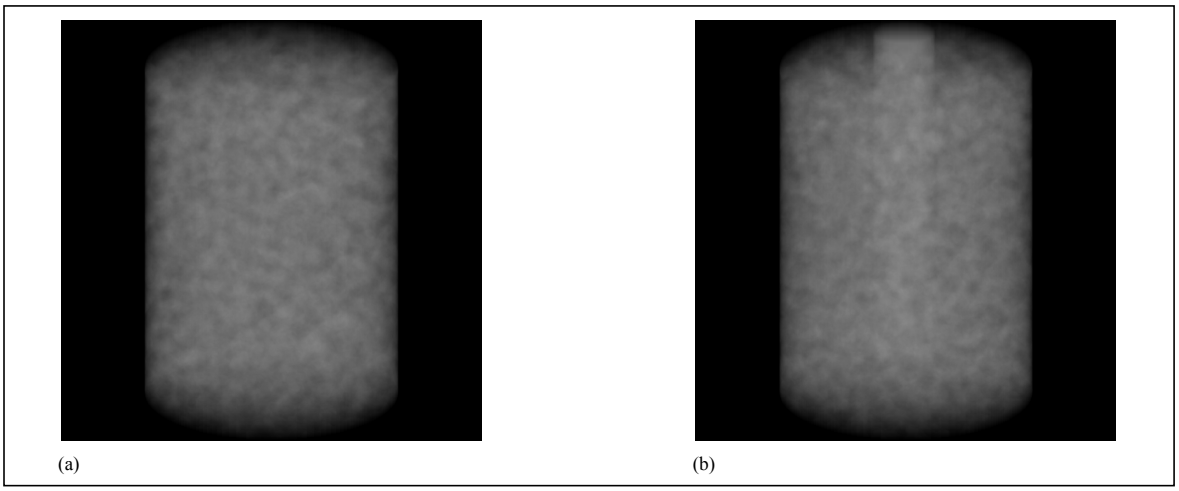


FIG. 3.16: Vue de face et légèrement du dessus du stimulus mis en place pour étudier la perception de la profondeur dans une scène transparente en mouvement (inspiré de [KSTE06]). Le stimulus présenté aux sujets est constitué d'un cylindre allongé rempli avec un bruit de Perlin. Pour choisir des fonctions de transfert adéquates, un plus petit cylindre (b), de niveau scalaire maximal ($s = 1$), est positionné à l'arrière du cylindre principal, et les paramètres de rendu sont réglés de telle sorte que cette sonde soit toujours visible. Ce petit cylindre n'est pas utilisé au cours de l'expérience, mais seulement au préalable.

où $N(x)$ est la fonction de Perlin de base, et b et $1/a$ définissent respectivement les fréquences relatives et la persistance des harmoniques sommées. Nous choisissons ici $a = b = 2$ et $n = 4$. Les données volumiques sont enregistrées dans une texture 3D, et le rendu est réalisé par accumulation de l'arrière vers l'avant de 100 plans régulièrement espacés. Les images de l'animation sont précalculées, ce qui garantit que le stimulus est effectivement affiché avec un taux de rafraîchissement de 60 images par seconde.

En présentation statique, l'image paraît être celle d'un rectangle plat, mais la mise en rotation de l'ensemble fait percevoir le cylindre. Comme signalé précédemment, la direction de mouvement ressentie, horaire ou antihoraire, est directement corrélée aux relations de profondeur perçues dans le volume.

Six conditions de rendu distinctes ont été implémentées : MIP, X-RAY, et quatre DVR avec des fonctions de transfert différentes. Le réglage de la FT demeure un problème important en rendu volumique, et la quantité de FT potentielles ne permet pas d'explorer exhaustivement l'espace des paramètres correspondant. Dans cette expérience, nous avons décidé de nous concentrer sur deux formes particulières de FT, la première (DVR_l) linéaire en luminance et constante en opacité et la deuxième (DVR_α), inversement, linéaire en opacité et constante en luminance, ce qui correspond à

$$(DVR_l) \quad l(s) = s \quad , \alpha(s) = \alpha_l \tag{3.11}$$

$$(DVR_\alpha) \quad l(s) = 1 \quad , \alpha(s) = \alpha_a s. \tag{3.12}$$

Ces deux fonctions sont illustrées sur la figure 3.17. Dans chaque condition, deux valeurs ont été retenues pour α, sur la base d'un simple test de transparence. Un cylindre quatre fois plus fin que le cylindre principal, mais de même hauteur, est positionné à l'intérieur du cylindre principal, tangent à son bord (voir figure 3.16(b)). Le champ scalaire dans cette sous-région est fixé à 1, et les valeurs de α sont choisies de telle sorte que cet élément intérieur soit visible, quel que soit l'angle de vue. Cela correspond approximativement à une situation où l'on souhaite que les maxima du champ analysé

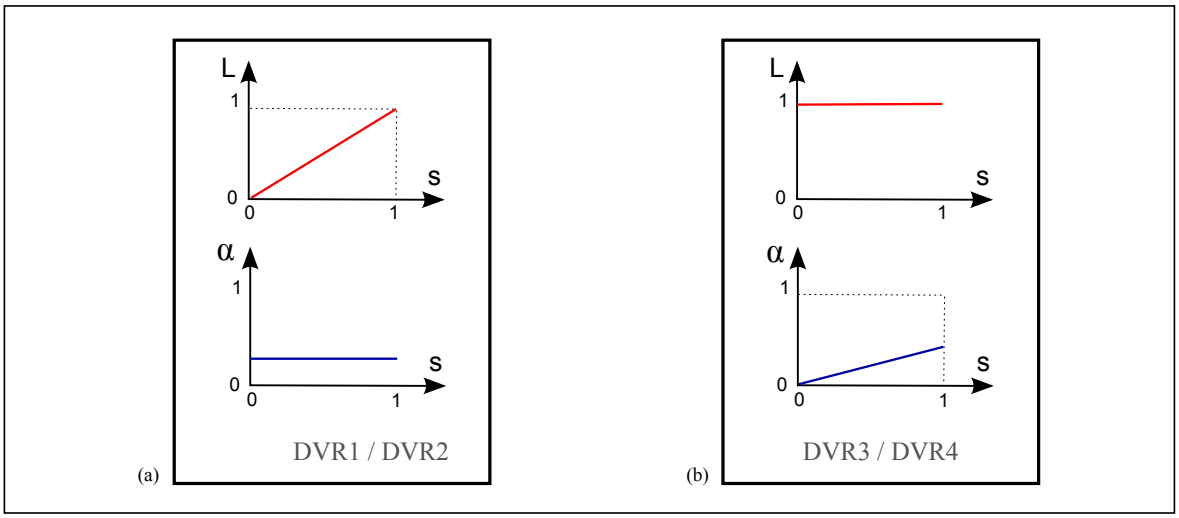


FIG. 3.17: Choix des fonctions de transfert.

soient toujours détectables. Les valeurs retenues sont respectivement $\alpha_l = 0.01$ et $\alpha_l = 0.025$ pour DVR_l, et $\alpha_a = 0.015$ et $\alpha_a = 0.025$ pour DVR_α. Par la suite, nous désignerons respectivement ces quatre conditions de rendu par $DVR1$ et $DVR2$ (FT DVR_l) et $DVR3$ et $DVR4$ (FT DVR_α).

Des exemples d'images produites avec les six conditions de rendu sont reproduites sur la figure 3.18. Pour chacun des six modes de rendu, une rotation vers la gauche et vers la droite sont implémentées, ce qui génère un total de 12 conditions expérimentales.

4.1.2 Déroulement de l'expérience

Il est demandé aux sujets de déterminer le sens de rotation du cylindre, en cliquant avec la souris sur l'icône de leur choix. Les deux réponses proposées sont « Vers la gauche » et « Vers la droite », le sens de ces expressions étant explicité préalablement au moyen d'un rendu opaque ne présentant aucune ambiguïté. Pour rendre claire la configuration spatiale de la scène, un affichage du cylindre solide légèrement vu de dessus est proposé avant le début de la session. Aucun retour sur les réponses fournies n'est donné au cours de l'expérience.

Le dispositif expérimental est identique à celui de la première expérience (section 3.1.2 p.74), la position du regard des participants étant enregistrée pendant la présentation du stimulus. Le cylindre occupe dans le champ visuel 14 ° verticalement et 10 ° horizontalement. La rotation (34 ° /s) induit une vitesse maximale de 3 ° /s dans le champ visuel, pour un observateur situé à 80 cm de l'écran.

Au cours d'une session expérimentale, chacune des 12 configurations possibles (6 rendus, 2 sens de rotation) est présentée 5 fois, la séquence des essais étant déterminée aléatoirement. Chacun des 60 essais commence par une présentation de la croix de fixation au centre de l'écran, suivie de l'affichage d'une des séquences précalculées pendant une demi-seconde. Puis l'écran de réponse est affiché, le sujet sélectionne la direction perçue, ce qui déclenche l'essai suivant. A la fin de la session, les réactions et impressions du sujet sont recueillies pendant un entretien avec l'expérimentateur.

Le temps de présentation très court du stimulus, 0.5 s, permet de prévenir l'occurrence d'inversions dans le sens de direction perçu. Cela élimine l'influence de biais perceptifs liés à de tels événements, ainsi que le trouble éventuellement engendré par l'apparition d'états bistables. Ainsi, si deux directions étaient éprouvées successivement pour un même essai, le sujet pourrait choisir préférentiellement une réponse dans ce genre de cas, comme « Vers la droite ». Notons malgré tout que ce biais à été observé pendant l'expérience, 65% des réponses correspondant à la perception d'un

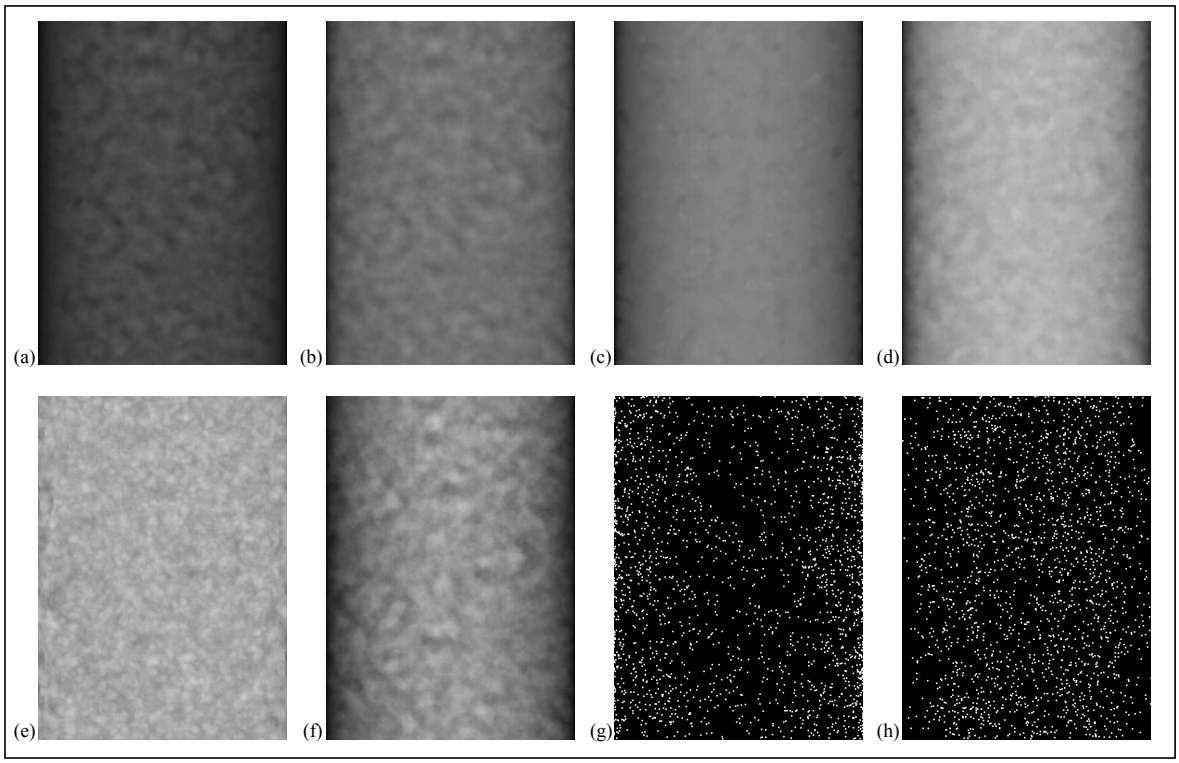


FIG. 3.18: Les six rendus sélectionnés pour le cylindre avec bruit de Perlin, en vue frontale. Ceux-ci correspondent à un DVR avec fonction de transfert linéaire en luminance ((a) et (b), conditions $DVR1$ et $DVR2$), linéaire en opacité ((c) et (d), conditions $DVR3$ et $DVR4$), et aux rendus non ordonnés en profondeur, de type MIP (e) et X-RAY (f). (g) et (h) montrent respectivement les nuages de points surfacique et volumique ajoutés pour l'expérience sur l'influence du mode de projection (section 5).

cylindre tournant vers la droite.

4.2 Résultats

Dix sujets ont participé à cette expérience, les mêmes que pour la première expérience. Leurs résultats moyens, correspondant au nombre de fois où la réponse donnée correspondait au sens réel de rotation, sont illustrés sur la figure 3.19.

Les performances obtenues lorsque le rendu était du MIP ou du X-RAY ne sont pas distinctes du niveau du hasard (respectivement 48% et 54% de bonnes réponses, test t de Student : $p = 0.34$ et $p = 0.17$). Ces résultats ne sont pas surprenants, dans la mesure où les images générées par ces deux techniques ne dépendent pas de l'ordre en profondeur des éléments du volume. Ce niveau proche du hasard valide le fait que le stimulus élaboré ne contient pas d'indices visuels non contrôlés qui pourraient aider à résoudre la tâche ou en tout cas biaiser significativement les résultats.

Le DVR donne des résultats très différents, suivant la nature de la fonction de transfert implémentée. Dans le cas de la condition DVR_l (cas $DVR1$ et $DVR2$, α constant), le mouvement permet clairement de lever les ambiguïtés en profondeur, les sujets ayant répondu correctement pour 99.5% des essais. En revanche, les performances enregistrées lorsque la FT est linéaire en opacité sont beaucoup plus contrastés, avec respectivement 11% et 79% de bonnes réponses pour les conditions $DVR3$ ($\alpha(s) = 0.015$) et $DVR4$ ($\alpha(s) = 0.025$). Le $DVR3$ montre plus qu'une ambiguïté importante dans l'affichage, avec des performances moyennes nettement inférieures au niveau du hasard : il s'agit en

fait d'une inversion importante de la perception de profondeur, les éléments les plus proches étant en fait le plus souvent perçus comme se trouvant à l'arrière du volume. Pour le $DVR4$, la perception de profondeur est plus correcte, mais les ambiguïtés demeurent, les performances moyennes étant significativement distinctes de la perfection (test t de Student, $p < 0.01$).

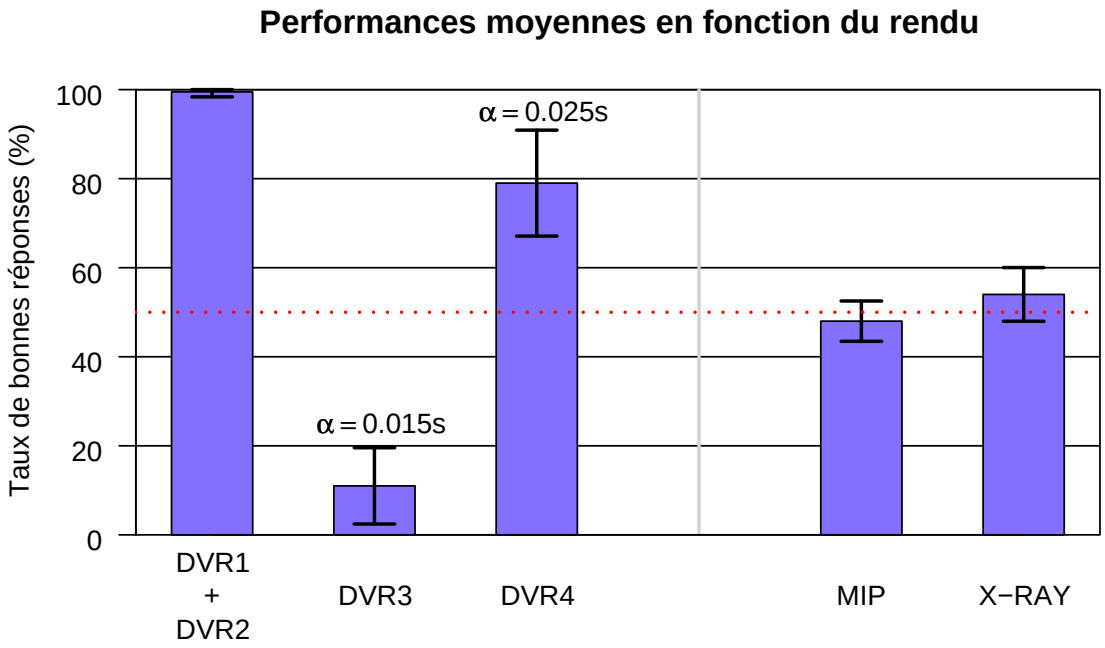


FIG. 3.19: Performances moyennes des sujets pour l'estimation du sens de rotation d'un cylindre avec rendu volumique, en fonction du modèle optique retenu. Les conditions $DVR1$ et $DVR2$ sont considérées ensemble. Dans cette expérience, seuls les rendus ordonnés en profondeur et avec une FT linéaire en luminance génèrent une perception correcte et non ambiguë.

4.3 Discussion

4.3.1 Importante de la fonction de transfert

Les résultats obtenus montrent que l'effet cinétique de profondeur peut grandement contribuer à faciliter la perception des relations spatiales dans des rendus volumiques ordonnés, mais seulement si la fonction de transfert est réglée avec soin. Etant donnée la taille de l'espace des FT potentielles [5], nous avons restreint notre analyse à deux catégories bien particulières, les FT linéaires en luminance et constantes en opacité, et l'inverse, les paramètres étant choisis dans chaque cas pour permettre une vision complète du champ volumique simulé. Nous avons vu que les rendus utilisant les premières bénéficient sans équivoque du KDE, celui-ci induisant une perception de profondeur claire, alors que les rendus reposant sur les dernières sont perceptivement plus ambigus, et provoquent même parfois des inversions de profondeur importantes.

[5]Pour un champ scalaire contenant des valeurs entières dans l'intervalle [0 :255] auxquelles on associe une luminance et une opacité codées sur un octet, il existe déjà $(256 * 256)^{256}$ FT différentes !

Rendu	MIP	X-RAY	$DVR1$	$DVR2$	$DVR3$	$DVR4$
Intensité moyenne des pixels ($\bar{I} \in [0:255]$)	170	145	73	114	128	176
Contraste RMS ($\frac{\sigma_I}{\bar{I}}$)	0.051	0.079	0.069	0.047	0.035	0.033

TAB. 3.1: Intensité moyenne et contraste RMS (*root-mean-square contrast*) pour les différents rendus volumiques implémentés.

Dans la pratique, le réglage de la FT demeure un problème délicat, souvent résolu empiriquement même si certaines méthodes semi-automatiques ont pu être proposées [PLS+00]. De plus, les FT deviennent de plus en plus complexes, elles comportent souvent des non-linéarités, et leur calcul repose sur de multiples dimensions [KPI+03]. On ne peut prétendre généraliser les résultats obtenus ici dans le cas de fonctions linéaires et unidimensionnelles à toutes ces fonctions de transfert potentielles. Cependant, nous avons montré que, même dans le cas de rendus ordonnés de transparence, une FT garantissant la discrimination des structures dans le volume peut engendrer des ambiguïtés de perception de profondeur lorsque l'observateur est confronté passivement au mouvement de la scène.

4.3.2 Influence du contraste

Les différentes techniques implémentées ici, et les paramètres associés, produisent des rendus de contrastes variables, aucune égalisation n'ayant été introduite entre les différentes conditions. Cela a pu modifier la perception du mouvement du cylindre, et en conséquence influencer la perception de profondeur. En effet, une région de l'image faiblement contrastée est susceptible d'apparaître quasiment immobile pour le SVH. L'existence de telles régions peut ralentir le mouvement global perçu.

Dans cette expérience, les participants ont affirmé avoir toujours perçu un mouvement du cylindre, et seulement certains d'entre eux ont eu l'impression que la vitesse de rotation variait d'une condition à l'autre, avec notamment une rotation plus lente dans la condition $DVR3$. Les intensités et contrastes RMS (*root-mean-square*) moyens dans la partie centrale de l'image sont reportés dans le tableau 3.1. On ne peut totalement exclure que les différences de contrastes aient pu jouer un rôle dans les performances enregistrées, par exemple entre $DVR1$ et $DVR4$. Cependant, les niveaux de contraste dans les différentes images se situent amplement au niveau des seuils de détection. Pour des luminances et fréquences spatiales proches de ceux rencontrés ici, [vNB67] font ainsi état d'un seuil de contraste d'approximativement 0.005.

En outre, les mesures de contraste dans l'image ne peuvent rendre compte des performances observées pour $DVR3$. L'inversion significative de profondeur perçue pour cette condition est plus délicate à interpréter. Un examen attentif du stimulus révèle en fait une inversion de contraste entre les éléments d'arrière-plan et d'avant-plan, avec notamment la présence de taches sombres parmi les premiers cités. Quand des taches sombres et claires se croisent dans une direction de vue particulière, le contraste local de l'image est réduit. Cela peut être interprété par le SVH comme un indice léger d'occultation généré par les structures sombres, donnant en conséquence l'impression que ces éléments se trouvent plus proches de l'œil. Les rendus volumiques présentent une certaine évanescence : lorsque le point de vue change, l'ensemble des voxels contribuant à la couleur d'un pixel est modifié. Ces changements peuvent induire un mouvement de deuxième ordre lié à la pseudo-texture résultant du rendu volumique, mouvement éventuellement opposé à la direction d'ensemble.

Le phénomène d'atténuation locale de contraste est une propriété intrinsèque des rendus par transparence, mais son influence sur la perception de la profondeur peut varier grandement en fonction de

la FT sélectionnée et de la structure spatiale des données. Dans des cas concrets de visualisation scientifique, le choix des paramètres de rendu, conjugué à la nature des données à analyser, peut influencer de nombreux éléments tels que la luminance, le contraste et les fréquences spatiales apparentes, ce qui en retour peut affecter la capacité à correctement percevoir les relations spatiales dans le volume.

5 Expérience 3 : apport de la perspective linéaire

Les stimuli utilisés dans l'expérience précédente ont été produits au moyen d'une projection orthographique, ce qui a permis de limiter l'information de profondeur disponible aux seules accumulations ordonnées de transparences. De nombreuses études relatives au KDE utilisent une telle projection parallèle, celle-ci constituant une bonne approximation des projections perspectives pour des angles de vue limités. Cependant, si le mouvement de l'objet projeté induit effectivement une perception tridimensionnelle dans ces conditions, la profondeur perçue dans l'image apparaît souvent ambiguë en l'absence d'autres indices. De tels stimuli génèrent souvent une perception bistable, une observation prolongée du mouvement s'accompagnant d'inversions simultanées du sens de rotation et de la profondeur perçus.

L'ambiguïté inhérente aux projections parallèles résulte du fait que, lors de la rotation du cylindre autour de son axe, les vitesses projetées de points situés à l'arrière et à l'avant du volume sont identiques, au signe près. Remplacer la projection orthographique par une perspective linéaire induit une réduction après projection des tailles des objets les plus lointains, et par conséquent une modification des valeurs absolues des vitesses dans l'image, les points les plus proches se déplaçant plus rapidement que les éléments plus lointains. Ces informations sont a priori de nature à faciliter une extraction correcte de la profondeur par le mouvement. Rogers et Rogers [RR92] ont effectivement observé que la perspective linéaire aidait à clarifier la perception de la profondeur relative de bandes de points en translation. Eagle et Hogervorst [EH99] ont quant à eux démontré que l'utilisation d'affichage à angle de vue élargi avec projection perspective permettait aux participants de mieux discriminer les angles diédraux. Plus récemment, Petersik et Dannemiller [PD04] ont présenté à des observateurs naïfs quant au sujet de l'expérience une sphère en rotation remplie de points, la direction de rotation étant parfois inversée pendant le mouvement. En variant le niveau de perspective, ils ont montré que le nombre moyen de changements de sens correctement détectés par les participants était clairement dépendant de la projection choisie.

Dans l'expérience présentée ci-après, nous cherchons à estimer si l'utilisation d'une perspective naturelle ou exagérée conjuguée à un mouvement de rotation de la scène est de nature à faciliter significativement la perception de la profondeur dans des rendus par transparence, qu'ils soient ordonnés en profondeur ou non. En plus des méthodes de rendu présentées dans la section 4, nous introduisons ici des stimuli composés de nuages de points proches de ceux que l'on retrouve traditionnellement dans la littérature relative au KDE/SfM.

5.1 Protocole expérimental

5.1.1 Modification du stimulus de l'expérience 2

Comme dans l'expérience précédente, le stimulus est constitué d'un cylindre empli d'un bruit de Perlin, et le rendu est réalisé au moyen de MIP, X-RAY ou DVR, avec quatre fonctions de transfert différentes dans ce dernier cas. A cela viennent s'ajouter deux conditions dites de nuages de points, inspirées par les expériences classiques en KDE depuis Green [Gre61]. Pour ces deux conditions, 2000 points blancs opaques sont placés aléatoirement soit sur la surface du cylindre (condition CLOUD-SURF), soit dans l'ensemble du volume intérieur (condition CLOUD-VOL). Les points sont

de taille constante égale à un pixel, ils ne fournissent donc aucun indice de profondeur autrement que par leur mouvement dans l'image. Ces deux nouveaux rendus sont illustrés sur les figures 3.18 (g) et (h).

En plus de la projection orthographique utilisée précédemment (condition O), deux vues perspectives du cylindre sont implémentées. Dans l'espace de l'observateur elles correspondent respectivement à une distance de vue de 80 cm (condition P1, approximation de la perspective naturelle pour l'observateur) et de 30 cm (condition P2, perspective amplifiée ou exagérée), l'écran passant par le milieu du cylindre. Les déformations induites par ces différentes projections sont illustrées par la figure 3.20, et la position virtuelle de l'observateur est représentée dans la figure 3.21 (a). La projection perspective induit des différences de vitesses notables entre les points situés à l'arrière et à l'avant du cylindre, fournissant une information directement exploitable par le système visuel pour reconstituer correctement le volume. Le rapport entre les vitesses maximales de points situés à l'avant et à l'arrière du cylindre sont respectivement de 0.86 et 0.62 pour les conditions P1 et P2.

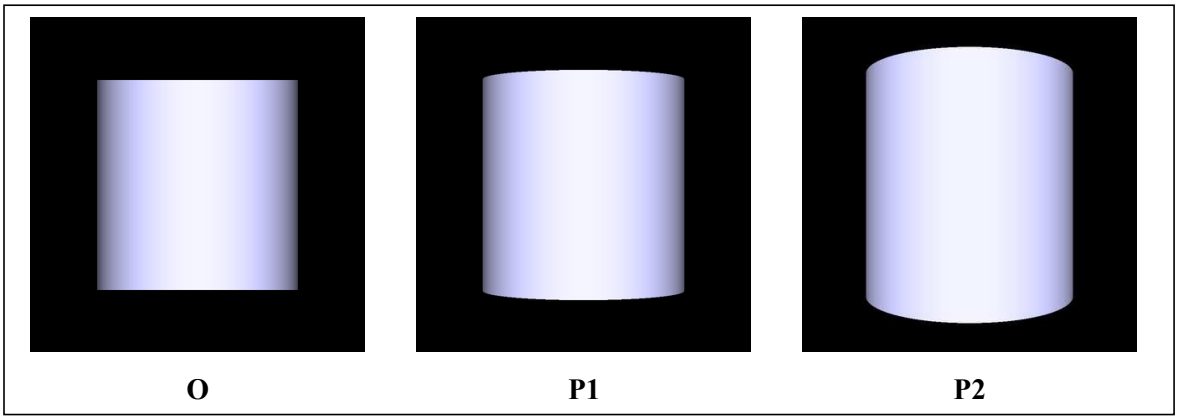
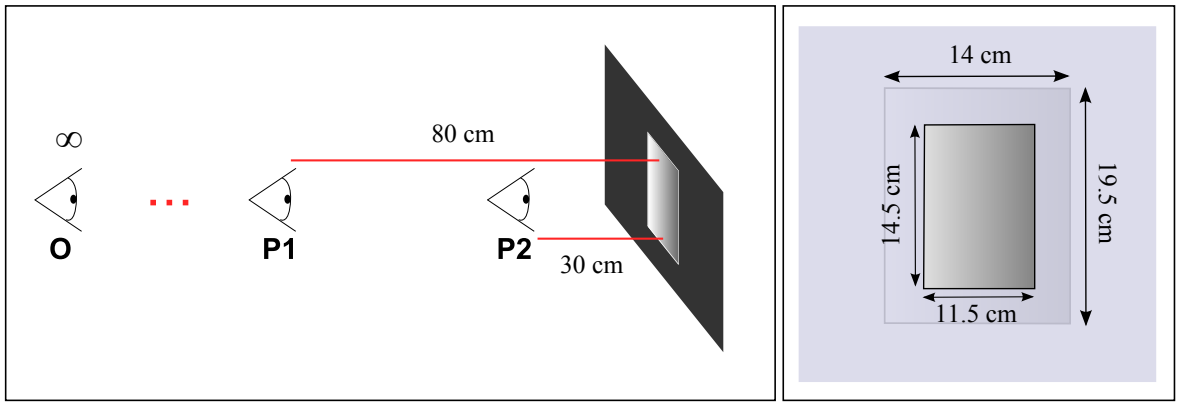


FIG. 3.20: Déformations induites par la projection perspective. Le cylindre plein est montré en projection orthographique (O), en perspective normale (P1) et en perspective exagérée (P2).

Avec une projection orthographique, les bords supérieur et inférieur du cylindre ne sont pas visibles, car ils sont parallèles à la direction de vue. En revanche, l'utilisation de perspectives linéaires les fait apparaître à l'écran, que ce soit pour les rendus transparents ou par nuages de points. Cette information supplémentaire est susceptible de fournir des indices non contrôlés facilitant la détection de la direction de rotation. A contrario, l'interprétation de ces discontinuités visibles par transparence peut également provoquer une certaine confusion chez le sujet regardant, et limiter sa capacité à réaliser la tâche demandée, ce que plusieurs observations informelles ont mis en évidence. Afin de prévenir ces artefacts liés à la structure de l'objet, un masque est introduit dans l'image afin de cacher les extrémités du cylindre (voir figure 3.21 (b)). Ce masque dissimule aussi les bords latéraux du cylindre, de manière à préserver approximativement le rapport largeur/hauteur de la partie visible de l'objet projeté. Cette coupure latérale présente l'avantage de permettre de réaliser l'expérience de contrôle décrite ci-après sans modifier la taille apparente du stimulus. En revanche, il peut résulter de ce masquage une perte partielle de l'information perceptivement pertinente pour résoudre la tâche principale, l'importance de cette perte étant difficile à évaluer.

5.1.2 Contrôle de la présence d'artefacts 2D dans les réponses

Dans les conditions P1 et P2, les éléments du volume les plus proches de l'œil vont bouger plus rapidement à l'écran que les plus éloignés. Il s'agit de l'information essentielle pouvant aider à résoudre



(a) (b)

FIG. 3.21: Configuration expérimentale utilisée pour mettre en évidence l'apport de l'information de perspective. Trois projections sont implémentées (a), qui correspondent à une caméra virtuelle placée à l'infini (O), ou à respectivement 0.8 m (P1, distance réelle des sujets à l'écran) et à 0.3 m (P2) du cylindre. Pour limiter l'extraction d'information au niveau des bords du cylindre en projection perspective, seulement la partie centrale de l'objet est visible, le reste de l'écran étant masqué par un cadre noir, schématisé ici en bleu (b).

l'ambiguïté de profondeur rencontrée par exemple dans l'expérience 2 pour les rendus MIP et X-RAY. Cependant, l'expérience menée ne mesure qu'indirectement la profondeur perçue, les participants rapportant pour chaque essai le sens de rotation apparent du cylindre. On ne peut dès lors écarter qu'une amélioration des performances dans la tâche expérimentale puisse résulter simplement d'une corrélation plus forte entre le sens de rotation du cylindre et la vitesse moyenne des points de l'image, et non pas d'une perception de tridimensionnelle plus précise. Comme le soulignent avec justesse Sperling et Landy [SL89], les participants pourraient très bien répondre, consciemment ou non, sur la base d'indications purement 2D, et ignorer voire ne pas percevoir la structure tridimensionnelle sous-jacente au stimulus représenté.

Pour vérifier si de tels artefacts pourraient expliquer une éventuelle amélioration des performances liée à l'augmentation du niveau de perspective, nous avons implémenté une session de test. Au cours de celle-ci, une translation latérale est appliquée au cylindre en plus de la rotation autour de son axe, de manière à compenser les différences de vitesse induites par la perspective. Nous faisons l'hypothèse que les sujets utilisent les mêmes mécanismes pour inférer le mouvement 3D de l'objet, quel que soit le rendu. Le test est ainsi réalisé uniquement avec la condition CLOUD-VOL en projection perspective amplifiée (P2). En effet, une étude préliminaire a montré que cette configuration provoque en général une réponse subjective tranchée et correcte. De plus, le rendu de points permet une estimation fiable et précise de la vitesse moyenne à l'écran. Trois conditions de translation sont définies : pas de translation du tout (T0) ; une translation annulant la vitesse moyenne des points à l'écran (T1) ; et une translation égalisant en valeur absolue les vitesses maximales projetées des points situés à l'avant et à l'arrière du volume (T2). Notons que, le cylindre étant partiellement occulté par le masque rectangulaire décrit précédemment, plus de points sont visibles à l'arrière du volume qu'à l'avant en perspective, si bien que la condition T2 inverse de fait la direction de la vitesse moyenne à l'écran. Si les réponses des sujets sont principalement guidées par cette information 2D, alors le sens de rotation ressenti ne correspondra pas au mouvement 3D réel du cylindre. Les détails des calculs déterminant l'amplitude des translations à appliquer dans les conditions T1 et T2 sont présentés dans l'annexe A.

Ces conditions de test sont présentées aux participants en deux blocs identiques, l'une avant

et l'autre après la session principale. Chaque bloc est composé de 5 répétitions par condition de translation et pour chacune des deux directions de rotations, ce qui fait en tout 30 essais mélangés aléatoirement. Des erreurs de réponse survenant au cours de la première session reflèteront probablement une difficulté à réaliser la tâche demandée, alors qu'une augmentation des erreurs sur la deuxième session peut résulter aussi d'une certaine fatigue visuelle. Ce contrôle ne vise pas à mesurer si une translation supplémentaire détériore la perception de forme tridimensionnelle, pas plus qu'à quantifier la précision éventuelle de cette perception de l'espace. Il est avant tout destiné à rejeter l'explication selon laquelle une amélioration significative des performances serait dû uniquement à une meilleure corrélation entre la vitesse moyenne dans l'image et le sens de rotation du cylindre.

5.1.3 Déroulement de l'expérience

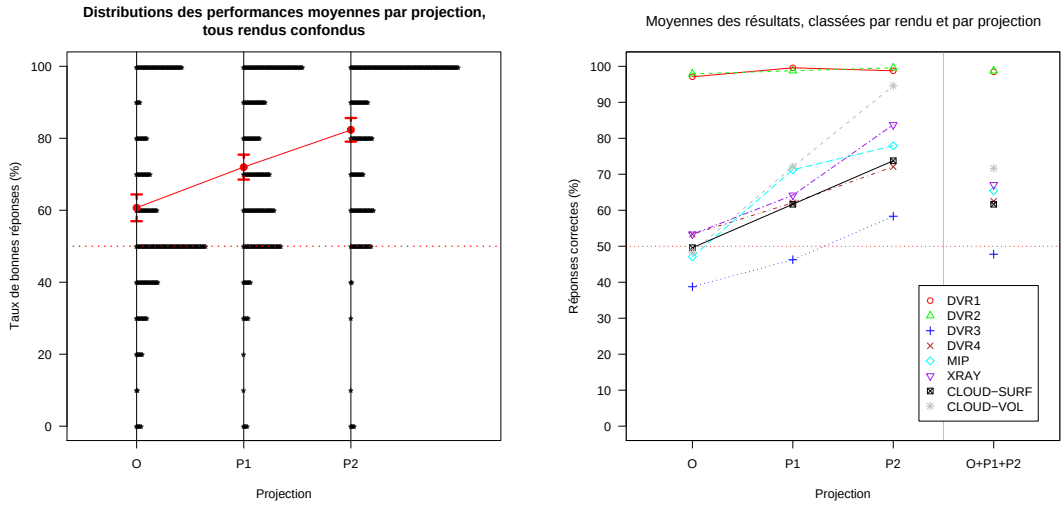
Aucun capteur de regard n'est utilisé ici. Les participants sont assis à 80 cm d'un écran LCD de 19 pouces (résolution 1280x1024), leur tête étant libre. Il leur est demandé de déterminer la direction de rotation de l'objet perçu, en sélectionnant la réponse correspondante avec la souris, parmi les deux choix proposés : « Vers la gauche » et « Vers la droite ». La partie visible du cylindre à travers le masque occupe, dans le champ visuel, 10.4 $^\circ$ verticalement et 8.4 $^\circ$ horizontalement.

Avant que l'expérience ne commence, la configuration spatiale du stimulus est décrite précisément aux sujets, au moyen d'un cylindre en rotation présenté avec le masque et sans le masque, toujours en projection orthographique. Afin de limiter d'éventuels effets d'hystérésis, il est précisé que les nombres d'essais correspondant à une rotation horaire et antihoraires ne sont pas nécessairement égaux. En effet, l'expérience 2 avait montré que certains sujets pouvaient parfois, en cas de doute, modifier leur réponse dans le but d'équilibrer les occurrences des deux directions de rotation, ce qui est susceptible de se produire notamment lorsque des rendus identiques se succèdent. Il est également demandé aux participants de répondre le plus rapidement possible, sur la base de leur première impression. Ces deux points visent à réduire l'influence du raisonnement dans les performances finales. Enfin, l'utilisation d'une perspective linéaire normale ou amplifiée n'est pas mentionnée, les seules images présentées relevant d'une projection orthographique.

Une session expérimentale est constituée de trois blocs : la session de contrôle initiale, la session principale et la session de contrôle finale. Les sessions de contrôle sont celles décrites dans la section précédente. Des pauses sont prévues entre les sessions ainsi qu'au milieu de la session principale pour réduire la fatigue visuelle causée par la présentation rapide de nombreuses images aux contrastes variables. La session principale comporte 240 essais (3 conditions de projection, 2 de rotation, 8 rendus, 5 présentations par configuration) dont l'ordre est généré aléatoirement. Pour chaque essai, une croix rouge de fixation est affichée au centre de l'écran pendant 1 à 3 secondes, puis elle laisse sa place au stimulus présenté pendant 0.5 s, puis l'écran de réponse est affiché. Une fois que le sujet a sélectionné sa réponse, l'essai suivant démarre. A la fin de l'expérience, les réactions et impressions du sujet sont recueillies. Il lui est en particulier demandé s'il a perçu plus de rotations dans une direction spécifique, si la vitesse de rotation lui a paru égale d'un essai à l'autre, et s'il a ressenti un mouvement autre qu'une simple rotation pendant les sessions de contrôle.

5.1.4 Participants

Vingt-quatre personnes âgées de 22 à 58 ans (moyenne : 37.8 ans) ont participé à cette expérience. Aucune d'entre elles n'avaient passé l'une des deux expériences précédentes. Toutes étaient naïves quand aux objectifs de l'étude, et leur expertise relative à la 3D assistée par ordinateur variait de « totalement naïf » à « utilisateur quotidien ». Tous avaient une vision normale, éventuellement après correction.



(a) (b)

FIG. 3.22: (a) Moyennes, intervalles de confiance à 5% et distributions de l'ensemble des résultats moyens pour chacune des trois projections. (b) Classement des différents rendus en fonction des performances moyennes des sujets.

5.2 Résultats

5.2.1 Session principale

Globalement, l'augmentation dans l'affichage des différences de vitesse et de taille dus à l'introduction de perspective normale et exagérée facilitent la perception de la profondeur par le mouvement, comme l'illustrent les figures 3.22 et 3.23. Hormis pour les rendus DVR1 et DVR2, pour lesquels le taux de bonnes réponses est proche de 100% quelle que soit la projection, le nombre moyen de réponses correctes augmente pour chaque rendu entre les conditions O à $P1$, et entre $P1$ à $P2$. Cette tendance est confirmée par une série de tests orientés de Wilcoxon comparant pour chaque rendu les paires de conditions O vs $P1$ et $P1$ vs $P2$, avec comme hypothèses alternatives respectives : $P1 > O$ et $P2 > P1$. Sauf pour DVR1 et DVR2, tous les tests sont significatifs ($p < 0.05$).

Pour les rendus ne présentant normalement aucun indice de profondeur en projection orthographique, à savoir MIP, X-RAY, CLOUD-SURF et CLOUD-VOL, les performances moyennes enregistrées sont proches de 50%, ce qui est cohérent. En ce qui concerne les conditions DVR3 et DVR4, les taux moyens de bonnes réponses diffèrent clairement de ceux enregistrés dans l'expérience 2 (section 4.2 et figure 3.19 p.86), avec ici respectivement 38.7% et 53.3% de réponses correctes contre 11% et 79% précédemment. Ce rapprochement vers le niveau du hasard reflète une plus grande difficulté à réaliser la tâche dans la nouvelle configuration, vraisemblablement en raison de la présence du masque qui cache une partie importante de l'information sur les côtés. Cette hypothèse est renforcée par les entretiens réalisés en fin d'expérience, au cours desquels les participants ont signalé un trouble plus important que précédemment face aux stimuli présentés. Notamment, le mouvement apparaissait dans certaines configurations plutôt plan que réellement tridimensionnel, ce qui rendait le choix plus incertain.

L'importance du niveau de projection dépend fortement de la condition de rendu, comme le montre la figure 3.22(b). Par exemple, la perspective exagérée (P2) génère une perception correcte à 94.6% dans le cas du nuage de points volumique, mais seulement à 73.5% pour le nuage surfacique. Ce dernier cas est en effet plus problématique, les participants ayant reporté a posteriori que le

mouvement de rotation du cylindre laissait souvent place à une impression de plans glissant l'un sur l'autre. Les performances en DVR4 sont proches de celles du nuage surfacique, alors que celles obtenues avec le DVR3 demeurent clairement inférieures aux autres, quelle que soit la projection. Pour le MIP et le X-RAY, l'analyse est un peu plus délicate. Le MIP bénéficie clairement de la perspective normale (P1), mais l'amplification en condition P2 n'apporte que peu d'améliorations. Au contraire, le rendu X-RAY est significativement affecté par l'exagération de perspective, les performances étant meilleures en condition P2 que pour les rendus MIP, DVR4 et CLOUD-SURF.

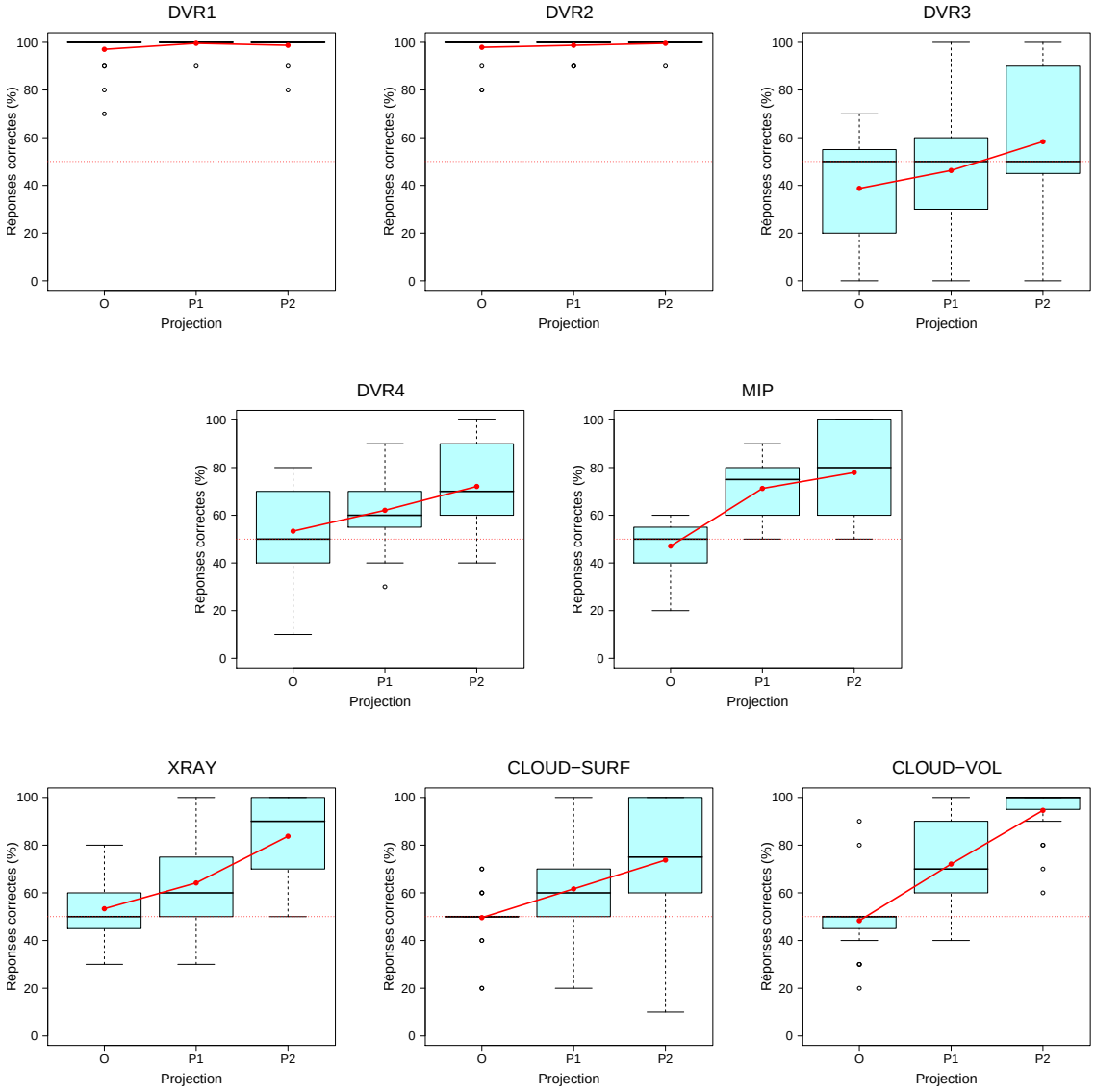


FIG. 3.23: Moyennes des réponses correctes des sujets pour la détermination du sens de rotation du cylindre, pour chaque rendu transparent et pour les nuages de points. Les boîtes à moustaches représentées illustrent la variabilité des réponses (médiane, écarts interquartiles, outliers). Les lignes rouges indiquent quant à elles les variations des moyennes de bonnes réponses entre les différentes conditions. Ces résultats révèlent une influence très nette du mode de projection, une exagération de la perspective facilitant la perception de la profondeur par le mouvement. La force de cet effet varie en fonction du rendu.

5.2.2 Sessions de contrôle

Les performances des sujets sur les essais de contrôle incluant une translation compensatoire sont très hautes. La session de début a donné lieu en moyenne à respectivement 95.2%, 93.8% et 93.3% de réponses correctes pour les conditions de translation T0, T1 et T2, et la session de fin à 97.1%, 95.7% et 94.3% de réponses correctes. Ces résultats sont très proches de la moyenne de 94.6% observée pour la condition CLOUD-VOL en projection P2 pendant la session principale. Cela montre non seulement que les participants ont correctement assimilé la tâche expérimentale, mais également que leurs réponses relevaient au moins pour ce rendu d'une réelle perception tridimensionnelle générée par KDE. Par ailleurs, un seul sujet s'est rendu compte qu'une translation avait été ajoutée à la rotation dans ces essais. Parmi les autres, certains ont perçu parfois un mouvement de torsion, notamment pour le cas T2 présentant la translation la plus importante, le reste ayant eu l'impression que la rotation du cylindre demeurait naturelle.

5.3 Discussion

5.3.1 Résultat principal

L'emploi d'une perspective linéaire induit une réduction des tailles projetées des objets les plus distants, et par conséquent une diminution des vitesses à l'écran fonction de la profondeur lorsqu'un volume est soumis à un mouvement rigide. L'expérience mise en place ici, reposant sur un effet cinétique de profondeur, a montré que cette information peut être utilisée par le système visuel humain pour lever les ambiguïtés de profondeur dans des rendus transparents, qu'ils soient ordonnés (DVR) ou non (MIP, X-RAY). Pour le rendu X-RAY, ces résultats complètent ceux obtenus par Kersten et al. [KSTE06] qui ont mis en évidence que les disparités binoculaires et l'atténuation de contraste avec la distance (perspective atmosphérique) permettaient de lever l'ambiguïté de profondeur inhérente à ce rendu. Il est important de noter que la projection joue un rôle non négligeable même si le rendu ne présente pas de sous-structures visuellement stables au fur et à mesure du mouvement, ce qui est le cas avec tous les rendus volumiques et plus particulièrement avec le X-RAY et le MIP. En effet, le MIP, ne retenant que les maxima le long de chaque rayon, certains pseudo-contours visibles dans l'image sont susceptibles de disparaître lorsqu'ils passent devant ou derrière une région portant un scalaire plus important. Cependant, ces résultats sont cohérents avec le travail de Todd [Tod85] qui a démontré que l'extraction de structure par le mouvement est particulièrement robuste à la présence d'un bruit visuel même important dans le mouvement d'ensemble.

L'amplitude du gain perceptif procuré par l'ajout de perspective varie grandement d'un rendu à l'autre, ainsi qu'entre les sujets ayant participé à l'expérience. Cela peut être expliqué en partie par le degré de rigidité présent dans les successions d'image, les points permettant de suivre facilement le volume alors que des rendus plus évanescents offrent une perception plus confuse au cours de la rotation de la scène. D'autre part, le remplissage du volume, ou du moins l'impression d'occupation de l'espace induite par un rendu particulier, pourrait influencer la perception du mouvement, comme le montrent les différences importantes relevées entre les nuages de points surfaciques et volumiques, ces derniers induisant une perception de profondeur bien moins ambiguë à mesure que la perspective est amplifiée. Si l'on regarde les résultats moyens obtenus pour les rendus volumiques MIP, X-RAY et DVR4, on constate qu'ils se trouvent compris pour les conditions P1 et P2 entre ceux correspondant aux nuages de points surfaciques et volumiques. Les conditions DVR4 et CLOUD-SURF induisent des réponses relativement proches en moyenne, et il est possible que le cylindre rendu par transparence soit perçu dans le cas du DVR4 non pas comme un volume plein, mais plutôt comme une simple surface transparente en mouvement. Pour le X-RAY avec une perspective exagérée, les résultats sont proches de ceux obtenus avec un nuage volumique, et on peut supposer que cette configuration induit

une véritable perception volumique. Ces remarques sont en accord avec les retours des participants recueillis après l'expérience, même si la méthodologie employée ne permet pas ici de caractériser autrement que subjectivement les notions de perceptions surfacique et volumique. Les données recueillies ici, avec un stimulus bien particulier et des contrastes variant fortement d'un rendu à l'autre, ne peuvent permettre de se prononcer définitivement sur ce point, et une étude plus approfondie sur la nature de la perception induite par des rendus transparents en mouvement serait certainement intéressante.

5.3.2 Une perception 2D ou 3D ?

Pour dissimuler les bords supérieur et inférieur du cylindre en projection perspective, nous avons appliqué un masque à l'image ne révélant que la partie centrale du cylindre. Lorsque, une fois l'expérience terminée, la rotation du cylindre rendu par transparence était montré aux sujets, ceux-ci étaient très troublés de pouvoir détecter les deux faces, et ils reportaient une perception instable et non rigide, ce qui en soi justifie l'emploi de ce masquage dans l'expérience principale. Cependant, ce dernier réduit la quantité d'informations disponibles, et rend la tâche plus complexe à réaliser. Notamment, de nombreux sujets ont déclaré avoir parfois perçu un stimulus 2D composé de surfaces flottant dans le plan image, plutôt qu'un vrai cylindre 3D. Les sessions de contrôle réalisées au début et à la fin ont montré que les participants réalisaient la tâche expérimentale sur la base d'une perception 3D dans le cas du nuage de points volumique, mais nous ne pouvons affirmer sur la base des mesures réalisées que ce soit toujours le cas pour les autres conditions de rendu. En observant le suivi oculaire d'un objet particulier au sein d'un mouvement 3D d'ensemble, Ringach et al. [RHS96] ont montré que le KDE induit une modification de la vergence. Une mesure précise de l'angle entre les axes des deux yeux pourrait ainsi permettre de déterminer la nature de la perception induite par le mouvement dans les rendus par transparence.

5.3.3 Biais latéral

La variabilité des résultats s'explique en partie par le fait que les sujets montrent souvent une préférence marquée pour une certaine direction de rotation dans les cas où le stimulus affiché est relativement ambigu. Par exemple, lorsque le nuage de points volumique est affiché en projection orthographique, la perception est normalement bistable, la direction de rotation ressentie pouvant indifféremment correspondre à un mouvement horaire (« vers la gauche ») ou antihoraire (« vers la droite »). On constate cependant que certains sujets sont biaisés envers une direction particulière, ils répondent beaucoup plus fréquemment « vers la gauche » ou « vers la droite ». Bien qu'une telle asymétrie dans les réponses pour un stimulus de KDE ait déjà été rapportée, par exemple par Sereno et Sereno [SS99], nous n'avons pas connaissance d'une quelconque explication de ce phénomène.

Pour déterminer si la force de ce biais latéral a une influence sur les performances des sujets, nous choisissons les quatre rendus totalement ambigus en projection orthographique, c'est-à-dire les nuages de points surfacique et volumique, le MIP et le X-RAY. Pour chacun, on calcule par sujet le pourcentage moyen de réponses correspondant à « vers la gauche », la moyenne des quatre étant appelée N_G. Le biais latéral d'un sujet est déterminé par le maximum de N_G et $100 - N_G$: c'est l'amplitude absolue du biais, quelle que soit sa direction. Un calcul du coefficient de corrélation de Pearson montre que le taux global de réponses correctes des sujets dans l'expérience est significativement lié à l'importance de ce biais directionnel ($R = -0.55$). La corrélation entre ces deux variables est explicitée sur la figure 3.24.

En d'autres termes, dans cette expérience, les sujets sont d'autant plus à même d'utiliser les indices de perception de la profondeur présents dans la séquence animée qu'ils ne présentent pas de préférence forte pour une direction de rotation particulière.

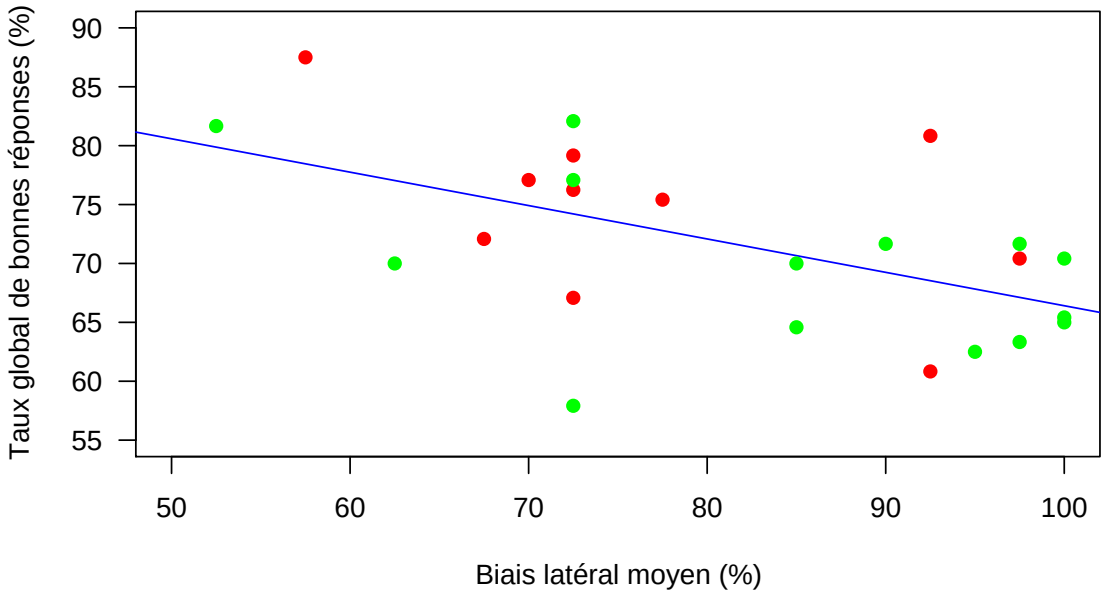


FIG. 3.24: Lorsque le rendu est totalement ambigu, les deux directions de rotation générant la même stimulation, certaines personnes montrent malgré tout une préférence pour percevoir un mouvement « vers la gauche » (points verts) ou « vers la droite » (points rouges). Les performances des sujets dans l'ensemble de l'expérience sont corrélées avec la force de ce biais directionnel (coefficient de Pearson, $R = 0.55$).

6 Conclusion et perspectives

6.1 Discussion des résultats

6.1.1 Rendu volumique et perception de l'espace

L'exploration de données volumiques denses pose un problème fondamental, dans la mesure où une méthode de rendu idéale devrait permettre d'une part d'afficher le plus de données possibles, en dépit des occultations survenant nécessairement au moment de la projection, tout en offrant une compréhension claire des phénomènes observés et de leurs relations spatiales. De nombreuses méthodes ont été proposées, parmi lesquelles les modèles optiques de DVR occupent une place importante. De tels rendus inspirés par des modèles physiques supposent que le système visuel humain est capable d'appréhender et de comprendre la structure des volumes partiellement transparents, alors que de tels objets ne se rencontrent que rarement dans notre environnement.

Il est intéressant de noter que les premières études de psychophysique sur la perception de la transparence partagent un même ancrage physique, avec des équations issues du modèle descriptif de l'épiscotistère. Mais comme le soulignent Singh et Anderson [SA02], le lien entre la théorie de la perception et de tels modèles est plutôt vague, leurs travaux insistant sur la déviation existant entre la transparence perçue et le modèle de l'épiscotistère de Metelli. De manière analogue, les expériences présentées dans ce chapitre montrent que des modèles inspirés de la propagation de la lumière dans un matériau dense ne génèrent pas nécessairement une perception correcte de la profondeur même pour

des objets relativement simples présentés statiquement. Le SVH ne semble que faiblement en mesure de s'appuyer sur les réductions de contraste induites par l'accumulation ordonnée de transparences pour interpréter correctement de telles images, et peut être fortement influencé par d'autres indices.

Les résultats présentés ici sont liés à une tâche spécifique de jugement de profondeurs relatives entre objets sans signification particulière pour l'observateur. Il est montré que le mouvement de la scène ainsi que l'utilisation d'amplifications perspectives facilitent la perception spatiale. Cependant, en visualisation scientifique, l'utilisation du mouvement n'est pas toujours souhaitable ou possible en pratique, et des vues quasi statiques sont plus à même de favoriser la reconnaissance d'objets ou de sous-structures particuliers dans les données. De même, la perspective amplifiée introduit des distorsions importantes de forme et de taille, et peut poser problème lorsqu'il s'agit de comparer l'extension de phénomènes ou d'éléments éparpillés dans le volume. En d'autres termes, il faut veiller à ce que l'amélioration de la perception de profondeur ne réduise pas la capacité à accomplir d'autres tâches importantes en visualisation scientifique. Par ailleurs, rappelons que notre questionnement principal concerne l'utilisation de rendus transparents dans la phase exploratoire de la visualisation, lorsque les résultats de simulation à analyser ne sont pas connus. Une fois cette étape passée, l'ingénieur peut être plus à même de trouver les vues et paramètres de rendu permettant de générer par DVR des images claires illustrant les principaux phénomènes physiques simulés.

6.1.2 Contrastes et couleurs

Dans les expériences 2 et 3, le choix des paramètres de rendu génère des contrastes très différents d'une image à l'autre, ce qui peut notamment influer sur la capacité à détecter le mouvement à l'écran. De plus, l'écran LCD n'a pas été calibré avec une sonde externe, ce qui reflète la réalité des pratiques quotidiennes en visualisation. Par conséquent, la relation entre l'intensité du signal d'affichage et la luminance de l'écran n'était pas ici linéaire. Par rapport au cas idéal, cela peut avoir modifié une partie de la réponse perceptive des participants, mais pas les conclusions principales des expériences menées. Par ailleurs, en dehors des problèmes de calibrage de l'écran, les conditions lumineuses ambiantes peuvent également jouer un rôle sur la perception dans les rendus volumiques. En effet, Purves et al. [PWNL04] ont montré que l'intensité lumineuse perçue par le SVH ne dépend pas linéairement de la luminance physique, et peut en outre être modulée par la luminosité environnante. Dans notre cas, les participants étaient placés dans une pièce comportant des sources de lumières atténuées mais pas totalement supprimées, ce qui correspond à peu près aux pratiques usuelles en visualisation.

Par ailleurs, nous avons restreint notre étude au cas de rendus achromatiques. En visualisation scientifique, l'utilisation d'échelles de couleurs est une pratique courante, et les fonctions de transfert peuvent associer à chaque voxel une couleur et une opacité. Si ces couleurs peuvent permettre de mieux isoler certains objets dans le volume, nous ne pensons pas que l'intégration ordonnée de ces couleurs dans un rendu volumique puisse constituer un indice de profondeur véritablement accessible au SVH.

6.2 Améliorer l'exploration de données volumiques

6.2.1 Faciliter la perception de profondeur par le mouvement

Les expériences 2 et 3 reposent sur l'ambiguïté de perception de profondeur relative à l'extraction de structure par le mouvement. Typiquement, lorsque le stimulus est composé d'un nuage de points en projection orthographique, le mouvement de la scène induit par KDE un état bistable dans lequel le sens de rotation apparent est directement corrélé avec les relations de profondeur perçues. En visualisation, une telle perception ambiguë peut être assez courante lorsque l'objet visualisé comprend

relativement peu d'indices d'occultation, par exemple pour des rendus transparents ou pour des scènes éparses. Une manifestation directe de cette ambiguïté se produit par exemple lorsque l'utilisateur tourne la scène au moyen de la souris. Pour un rendu non opaque, le monde virtuel peut sembler se déplacer dans la direction opposée au mouvement commandé, et les structures spatiales deviennent confuses. Disposer de méthodes prévenant l'occurrence de ce phénomène présenterait clairement un avantage pour un outil de visualisation.

Plusieurs études ont montré que la présentation d'une stimulation visuelle dynamique corrélée avec la rotation principale de l'objet, que ce soit dans le centre du champ visuel ou en vision périphérique, permet de résoudre l'incertitude de profondeur d'un nuage de points en rotation à l'écran. Rogers et Rogers [RR92] font tourner une texture horizontale posée devant l'écran dans la même direction que la surface avant du stimulus présenté à l'écran et observent que cette configuration permet de produire une perception de la rotation beaucoup plus claire. Sereno et Sereno [SS99] rapportent que le mouvement 2D de points autour d'une sphère échantillonnée par un nuage de points facilite une perception correcte si la direction de ce mouvement est opposée à celle des points avants de la sphère en rotation. L'utilisation d'une surface occultant une partie restreinte des points arrières peut aussi stabiliser la perception, à condition que ce masque soit visible explicitement dans l'image [FH04]. Par ailleurs, Freeman et Driver [FD06] montrent que la corotation d'un stimulus de contexte permet de lever l'ambiguïté de profondeur sur le nuage de points principal, ce qui révèle l'importance des mécanismes à longue portée dans la perception du mouvement. Cependant, cela n'est véritablement efficace que si ce contexte est partiellement ambigu et non opaque, avec par exemple des points contrastés en avant-plan et plus transparents en arrière-plan.

Il serait intéressant de vérifier comment de tels protocoles pourraient être utilisés en pratique pour clarifier la profondeur perçue par KDE dans des rendus autres que des nuages de points. L'affichage de points en translation dans l'image, proposé par [SS99] pourrait par exemple être testé avec des rendus volumiques denses ou simplement des surfaces transparentes superposées, comme illustré dans la figure 3.25. L'expérience pourrait être réalisée pour des surfaces plus ou moins familières afin de vérifier l'importance de l'a priori dans une telle stabilisation de la perception, et pour plusieurs valeurs de paramètres du nuage de points arrière, tels que sa densité, son contraste par rapport au fond et la vitesse de déplacement de l'ensemble. L'objectif serait ici de déterminer comment une telle solution peut être appliquée concrètement dans une application de visualisation, l'affichage supplémentaire de points devant être le moins intrusif possible.

6.2.2 Trouver de nouveaux rendus perceptivement adaptés

Le rendu de données volumiques denses demeure un défi important en visualisation scientifique. Les méthodes de DVR simples génèrent des rendus difficiles à régler et souvent compliqués à appréhender, que ce soit au niveau de l'identification des formes présentes dans la scène ou des profondeurs relatives des différentes sous-structures.

Les techniques classiques de rendu volumique peuvent être améliorées afin de mieux faire ressortir certaines structures caractéristiques, que ce soit en seuillant sur les transparences accumulées [RSK06] ou en implémentant des techniques non photoréalistes comme la détection de silhouette [RE01] ou l'ajout de halos [BG07]. De manière générale, la quantité de transparence présente dans l'image au moment de l'exploration des données ne peut être que limitée, car notre système visuel semble peu adapté à utiliser efficacement cette information. Il faut sans doute se restreindre à l'affichage d'un nombre restreint de surfaces superposées, ou à un volume présentant un faible extension de profondeur, la transparence devant alors aider à mieux comprendre ce qui se passe au voisinage d'une surface moyenne dans les données. Par ailleurs, le DVR peut être remplacé par des techniques non photoréalistes inspirées de rendus artistiques, en échantillonnant le volume par des primitives

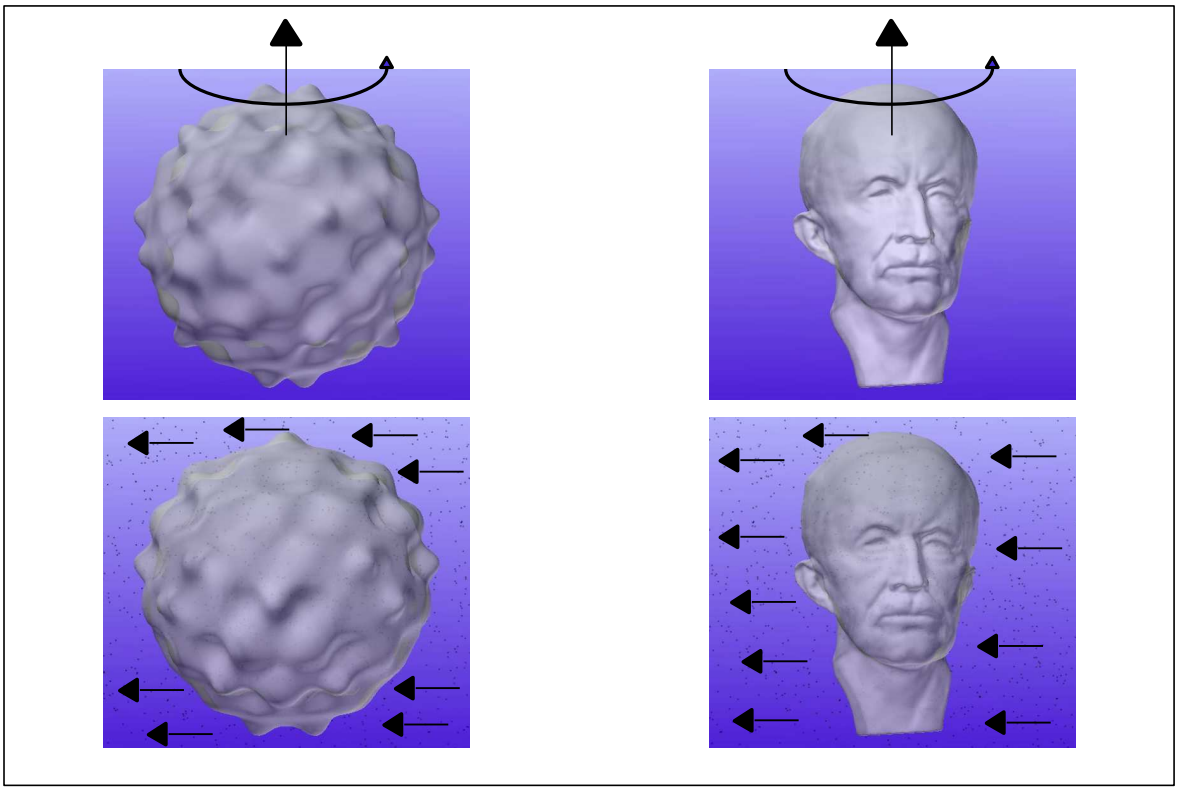


FIG. 3.25: En l'absence d'indices visuels forts, l'effet cinétique de profondeur génère une perception bistable dans laquelle les surfaces avants et arrières peuvent être inversées pendant la rotation de la scène. Dans de tels cas, des informations visuelles supplémentaires peuvent permettre de clarifier la perception de profondeur. Par exemple, un nuage de points bidimensionnel dont la direction de translation est opposée au mouvement des points avants de la surface du stimulus permet de lever les ambiguïtés relatives au sens de rotation d'une sphère échantillonnée par des points [SS99]. Il serait intéressant d'étudier si cet effet fonctionne avec d'autres types de rendus ambigus, comme des surfaces transparentes superposées. L'étude peut être menée avec des surfaces quelconques (à gauche) ou des objets pour lesquels un biais cognitif amplifie l'ambiguïté, voire biaise la perception, le cas le plus connu étant l'observation d'un visage en rotation (illusion du masque, à droite).

élémentaires telles que des points [LME+02] ou des traits [BKR+05]. Si les images produites peuvent paraître convaincantes à première vue, il n'est pas certain qu'elles permettent une meilleure exploration des données, des rendus aussi épars manquant fortement d'indices de profondeur.

Le mouvement de la scène ou de l'observateur peut être utilisé pour faciliter la perception des données visualisées, par effet de parallaxe ou de profondeur. Pour améliorer l'exploration, on pourrait également penser à utiliser des informations visuelles dynamiques pour un point de vue stable, en faisant varier un certain nombre de paramètres du rendu de façon continue. Par exemple, pour le rendu volumique, il s'agirait de modifier continûment la fonction de transfert de manière à faire ressortir successivement différentes structures dans le volume. Il serait alors intéressant de déterminer s'il existe des modes de variations tels que nous soyons capables d'intégrer localement l'information perçue dans le volume, et si oui quelles en sont les limites. De plus, de tels rendus reposant sur notre perception du mouvement nécessiteraient une interactivité importante, les algorithmes utilisés doivent donc absolument être les plus efficaces possibles en termes de temps de calcul.

CHAPITRE 4

EyeDome Lighting : un éclairage interactif et perceptivement adapté

> La clarté, c'est une juste répartition d'ombres et de lumière.
>
> Johann Wolfgang von Goethe

En modulant les intensités lumineuses du monde qui nous entoure, l'éclairage constitue une importante source d'information sur la forme et l'orientation des objets environnants, ainsi que sur leurs relations spatiales. S'il résulte simplement de la propagation de la lumière et de son interaction avec les objets rencontrés, les effets produits sont divers par nature : intensité fonction de l'orientation relative des surfaces par rapport aux sources de lumières, ombres portées, interréflexions, importance de l'étendu des sources de lumière. A partir de là, des causes différentes peuvent produire des effets très divers, ce qui rend difficile l'extraction de forme à partir de l'ombrage. Il s'agit en fait d'un problème mal défini, car largement sous-contraint. L'ambiguïté constituée par la perception de bas-reliefs en est un exemple typique, et il a été montré [BKY97] qu'une même image pouvait donner lieu à une infinité d'interprétations différentes quant à la surface correspondante, celle-ci se déformant en fonction de la direction supposée de la source de lumière.

Cependant, quand nous nous retrouvons dans des situations où l'éclairage est la principale source d'information, notre perception de l'espace nous paraît parfaitement déterminée, même si elle peut finalement s'avérer fausse. En dépit des incertitudes, notre système visuel est capable de résoudre le problème de l'extraction de forme par l'ombrage, et les formes perçues témoignent de contraintes importantes, comme l'hypothèse d'un éclairage par le haut ou la préférence pour des formes convexes. Il semble également que notre perception soit sensiblement différente en fonction de la nature même de l'éclairage, avec des effets différents suivant que la journée est plutôt ensoleillée ou nuageuse. La connaissance du fonctionnement du SVH est ainsi très importante si l'on souhaite mettre au point des algorithmes d'éclairage efficaces en termes de perception des objets.

La restitution de l'éclairage en informatique graphique est un domaine de recherche toujours actif, et des solutions très nombreuses ont été proposées, avec pour principale caractéristique le compromis réalisé entre niveau de réalisme et degré d'interactivité. Les méthodes de lancer de rayon [Whi80] visent à reproduire un éclairage global, c'est-à-dire à calculer la couleur de chaque pixel de l'image en tenant compte le plus précisément possible des caractéristiques du champ de lumière, des propriétés physiques des surfaces et du trajet des rayons lumineux entre les différentes sources primaires et secondaires et le point considéré, ce qui peut s'avérer très coûteux en temps de calcul suivant la complexité de la scène considérée. A l'autre extrême, les éclairages les plus couramment utilisés en rendu interactif [Gou71, Pho75] ne tiennent compte que de l'orientation des surfaces par rapport aux principales sources de lumières supposées ponctuelles (ombrage diffus) et par rapport à l'observateur (taches de spécularité). En dehors de ces approches, de nombreux travaux sont focalisés sur la restitution d'un effet particulier de l'éclairage, comme la restitution des ombres portées ou le calcul de la quantité de lumière visible en chaque élément de surface. Par ailleurs, de nombreux modèles, qualifiés de non photoréalistes, dévient explicitement des modèles physiques afin de mettre en évidence des caractéristiques spécifiques des surfaces, tels que la courbure ou la rugosité apparentes.

Dans ce chapitre, nous proposons un nouveau modèle d'éclairage non photoréaliste, l'EyeDome Lighting (EDL), dont la particularité est de ne nécessiter aucune information géométrique préalable sur la scène rendue. Cette contrainte s'appuie sur le constat que la quantité de données à visualiser est souvent trop importante pour avoir recours à des prétraitements importants, d'autant que de nouveaux objets peuvent être créés à la volée pour les besoins de la visualisation. Nous proposons donc une solution reposant uniquement sur la carte de profondeur obtenue au moment du rendu de la scène. Le modèle proposé vise à faciliter la perception des orientations et courbures, ainsi que la séparation en profondeur des différents objets d'une scène, tout en limitant la quantité de calculs à réaliser. Etant implémenté directement sur le GPU, l'EDL permet un calcul interactif d'éclairage pour une scène tridimensionnelle quelconque. Après une courte revue des travaux relatifs à la perception de la forme par l'éclairage (section 1.1) et aux algorithmes d'éclairage en informatique graphique (section 1.2), nous détaillons le fonctionnement de l'EDL, dont le principe repose sur une source de lumière fictive située autour de l'œil de l'observateur (section 2). Nous montrons ensuite comment ce modèle peut être étendu pour simuler une direction d'éclairage quelconque (section 3.1), et pour intégrer à moindre coût une zone de focus stylisé. Nous concluons en apportant notamment quelques lumières supplémentaires sur l'évaluation possible des modèles d'éclairage (section 4).

1 Perception de l'éclairage, modèles infographiques réalistes et approximatifs.

1.1 Eclairage et perception de la forme

L'idée de cette section n'est pas de préciser exactement les règles qui définissent l'extraction de forme par l'ombrage, mais plutôt de mettre en lumière les a priori employés par le système visuel pour résoudre cette tâche complexe, ainsi que les biais perceptifs qui peuvent en découler. De nombreuses expériences ont été menées en psychophysique pour essayer de comprendre cet aspect particulier de la perception de la forme. Celles-ci posent parfois problème, comme le souligne [KvD03], quand elles reposent sur des scènes synthétiques générées par ordinateur sur la base de modèles d'éclairage simplifiés, par exemple ne tenant pas compte des interréflexions qui contribuent à rééclairer certaines surfaces. Cette remarque est juste si l'objectif est d'appréhender précisément notre capacité à exploiter l'information d'éclairage dans des situations écologiques. Cependant, de telles études, en simplifiant les phénomènes simulés, permettent non seulement de mettre en valeur des propriétés singulières de

notre perception visuelle, mais surtout peuvent également se révéler particulièrement utiles lorsque l'on s'intéresse aux rendus d'éclairage par ordinateur, ce qui est notre cas. Il faut néanmoins toujours conserver en mémoire que les résultats obtenus ne se généralisent pas toujours, et peuvent être influencés par les informations contextuelles présentes dans la scène rendue.

Notre expérience quotidienne témoigne de la capacité de notre système visuel à extraire des informations spatiales à partir des variations de luminance induites par l'éclairage ambiant sur le monde qui nous entoure. Elle met également en évidence l'importance que peut avoir le type d'éclairage sur notre perception de la forme : les contrastes et le relief apparent des surfaces peuvent nous sembler modifiés lorsque le soleil passe brusquement derrière les nuages. Plus généralement, un simple gradient de luminance sur une surface plane peut être interprété comme une information d'ombrage, et générer une perception de forme tridimensionnelle. C'est ce que l'on appelle l'extraction de *forme par l'ombrage* (*shape from shading*) [MT86, Ram88]. Cette capacité est notamment utilisée depuis longtemps par les artistes pour produire une impression de relief simplement à partir d'un dégradé d'intensités lumineuses sur leur toile. Ainsi, si l'on présente à un sujet trois images d'un même objet, la première ne contenant que la silhouette de l'objet, la deuxième étant un dessin au trait et la dernière une photographie de l'objet éclairé, on constate que le relief perçu se rapproche très nettement de la surface réelle dans ce dernier cas [Koenderink96a], ce qui illustre bien cette capacité à inférer la forme à partir de la distribution des luminances dans l'image. Cependant, la direction d'éclairage peut modifier le relief apparent, que la scène représentée soit issue d'une photographie [KvDCL96] ou une image de synthèse résultant d'un modèle physique d'éclairage simplifié [CF07].

Il n'en reste pas moins que l'extraction de forme à partir de l'ombrage est un problème sous-contraint [Oli91]. Ainsi, différentes surfaces peuvent donner la même image, pour peu que la direction d'éclairage ne soit pas clairement accessible, ce que Belhumeur et al. [BKY97] ont appelée l'ambiguïté du bas-relief. Le système visuel est donc obligé d'utiliser un certain nombre d'a priori pour extraire une information de forme, et ceux-ci peuvent en retour engendrer des perceptions incorrectes. L'illusion du cratère en est un exemple typique, dans lequel une simple inversion verticale d'une image transforme un creux en bosse ou réciproquement (voir Fig.4.15(a)). Afin de mettre au point des méthodes d'éclairage facilitant la perception des formes de l'objet et de ses relations spatiales dans une scène tridimensionnelle complexe, il est donc intéressant de comprendre comment le cerveau utilise les différents indices liés à l'éclairage, quels sont les a priori auxquels il recourt pour lever les ambiguïtés éventuellement présentes dans l'image rétinienne, et quels sont les biais perceptifs qui en découlent éventuellement. En effet, un éclairage photoréaliste, c'est-à-dire produisant une image du monde ne pouvant être distinguée d'une photographie, ne constitue pas nécessairement la panacée : Koenderink et van Doorn [KvD03] illustrent bien cette idée en montrant que les interréflexions se produisant dans une cavité peuvent faire apparaître celle-ci non seulement comme une bosse, mais également en apparence plus « haute » que son homologue convexe de même courbure réelle.

La quantité de lumière émise par un élément de surface varie en fonction de l'éclairage de ce point, de la nature de la surface, de son orientation par rapport aux différentes sources de lumière primaires et secondaires et de la direction sous laquelle on la regarde. Les photons atteignant une surface sont soit absorbés, soit réfléchis dans différentes directions où ils pourront à nouveau être réfléchis, et ainsi de suite. Ces interactions complexes donnent lieu à différents phénomènes qui reflètent les orientations locales des surfaces et leur arrangement dans l'espace. Par exemple, les ombres portées apportent de l'information quant à la présence d'objets entre la surface et la ou les source(s) d'éclairage. Les reflets sur une surface non totalement mate dépendent de l'orientation de celle-ci par rapport à la source de lumière, mais également de l'angle sous lequel on la regarde, alors que l'ombrage diffus sur une surface mate est uniquement fonction de l'angle entre la normale à la surface et la direction de la source. La figure 4.1 illustre ces différentes situations.

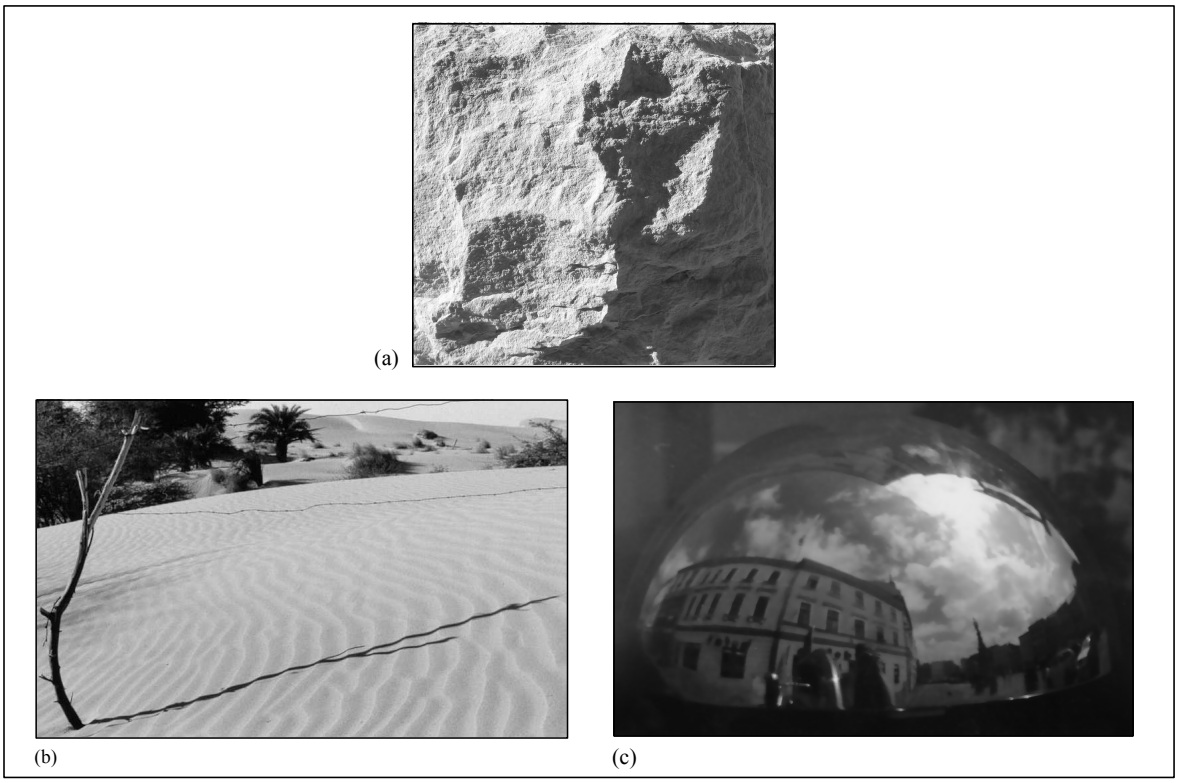


FIG. 4.1: Suivant les conditions d'éclairage et la nature des objets observés, plusieurs effets peuvent être observés. (a) L'ombrage diffus donne une information importante, quoique parfois ambiguë, sur la forme de la surface observée. (b) Les ombres portées permettent essentiellement de déduire la position des structures dans l'espace, sous réserve d'interpréter correctement la direction d'éclairage, et dans une moindre mesure leur forme. (c) Les reflets contribuent également à la perception de la forme, même si cette information est généralement plus délicate à exploiter.

Dans ce qui suit, nous nous concentrerons sur les connaissances relatives à quelques points précis de la perception de la forme par l'ombrage. Nous les considérons ici séparément, tout en gardant à l'esprit que toute forme perçue résulte d'une interaction de ces différentes composantes, et que le contexte dans lequel une surface est observée joue un rôle non négligeable dans la perception qui en découle.

1.1.1 Biais de convexité

La détermination de la forme nécessite d'intégrer l'ensemble des géométries locales en un percept global cohérent. Il est donc intéressant de savoir comment le système visuel humain traite les informations locales et globales de courbure. A ce niveau, l'illusion du masque (*hollow-face illusion*) nous montre une préférence importante de notre système visuel pour percevoir des formes convexes plutôt que concaves : la rotation d'un masque autour d'un axe vertical semble s'inverser lorsque la partie creuse de l'objet est visible. Cette illusion n'est pas spécifique à un objet aussi particulier qu'un visage, mais se retrouve également pour des patatoïdes quelconques [HB94].

En utilisant des quadriques et un modèle d'ombrage purement diffus, Erens et al. [EKK93] ont observé une confusion importante entre les différentes formes proposées, elliptiques et hyperboliques, à laquelle s'ajoute un biais important envers la perception de surfaces convexes en l'absence d'indications indirectes sur la direction d'éclairage (ombres portées). De même, l'inclinaison locale sur une

surface plus complexe est une information extraite avec plus de précision lorsque le point observé se trouve dans une région localement elliptique plutôt qu'hyperbolique [MK96]. Langer et Bülthoff [LB01] ont par ailleurs observé, dans une expérience où les sujets devaient discriminer creux et bosses, que la perception locale de la forme est fortement influencée par la forme globale de la surface, la tâche étant mieux réalisée sur des surfaces globalement convexes que concaves. En contrôlant finement les paramètres d'éclairage, Liu et Todd [LT04] ont confirmé que les jugements relatifs au signe de la courbure d'une surface sont réalisés avec un taux de réussite proche du hasard, avec un biais de réponse important en faveur de surfaces convexes. Par ailleurs, augmenter la profondeur des surfaces simulées a un effet sur la profondeur perçue relativement faible lorsque cette surface est concave : tout se passe comme si le biais de convexité entrait en conflit avec la profondeur perçue même lorsqu'un creux est correctement détecté. Si ces expériences sont menées sur des scènes synthétiques, avec un modèle d'éclairage simplifié, il est à noter que ce biais peut également être éprouvé en conditions naturelles. Koenderink et al. [KvD03] montrent ainsi que les interréflexions se produisant dans une concavité ont pour conséquence de rééclairer son centre, ce qui peut l'amener à être perçue comme plus convexe qu'une bosse de même courbure.

Dans l'image d'une surface continue, les *contours occultants* correspondent aux limites de l'objet dans la direction de vue. Ils jouent un rôle important dans l'extraction de forme, car ils indiquent des courbures localement convexes dans la direction de vue [Koe84]. S'ils se trouvent masqués, l'extraction locale de forme à partir d'un gradient de luminance est moins bien réalisée [TR89, RT90].

Le biais de convexité, local ou global, concerne la courbure de la surface observée. Certains auteurs [RT90, ML98, ML01, LB01] ont par ailleurs relevé un biais relatif à l'attitude des surfaces, à savoir une préférence du système visuel humain pour observer des surfaces par le haut, ce qui en terme d'image se traduit par : « plus un point est haut dans l'image, plus il est éloigné en profondeur ». Cette hypothèse influence les formes perçues à partir de dessins au trait [ML98] mais également pour des surfaces ombrées [RT90, ML01, LB01].

1.1.2 Conditions d'éclairage

Comme nous l'avons mentionné précédemment, les conditions d'éclairage peuvent modifier drastiquement le profil d'ombrage d'une surface. Une lumière très directionnelle, comme celle résultant d'un ensoleillement direct, produira des ombres plus nettes qu'un éclairage diffus, par exemple lors d'une journée très nuageuse. De même, une lampe pointée sur le haut ou sur le côté d'un objet ne produira pas les mêmes nuances, et cela peut conduire à des perceptions de forme assez différentes [KvDCL96, CF07].

Dans le cas d'une source de lumière dirigée dont les caractéristiques ne peuvent être extraites précisément, il a été observé depuis longtemps que notre cerveau traite l'information d'ombrage comme si l'éclairage venait du haut de l'image. D'un point de vue écologique, cela peut simplement refléter notre adaptation à un milieu dans lequel la lumière naturelle se trouve presque toujours au-dessus de notre tête. Cet a priori de *lumière venant d'en haut* influence notre perception des surfaces tant pour des scènes naturelles, comme dans l'illusion du cratère, que dans le cas de simples gradients 2D de luminance [Ram88]. Lorsqu'il s'agit de définir un modèle d'éclairage pour des scènes synthétiques, il est nécessaire de tenir compte de cette information. Nous y reviendrons donc un peu plus en détails en section 3.1.1. Il est important de remarquer que, au moins dans le cas de stimuli simples, les biais de convexité et d'éclairage venant d'en haut interagissent fortement.

L'autre caractéristique importante de l'éclairage est son étendue, c'est-à-dire la quantité de lumière venant de sources primaires et secondaires plus ou moins concentrées dans l'espace. A l'opposé de la lumière directionnelle, une journée nuageuse est caractérisée par une répartition à peu près constante de la quantité de lumière dans toutes les directions de la demi-sphère supérieure de l'espace. Dans

le cas extrême d'un éclairage directionnel, la luminance d'un élément de surface « voyant la source de lumière » est principalement déterminée par l'angle entre la normale à la surface et la direction d'éclairage. Au contraire, un éclairage diffus génère des profils de luminance principalement guidés par la proportion de la source de lumière visible depuis chaque élément de surface, les orientations jouant un rôle mineur. Par conséquent, les cavités vont apparaître plus sombres que les bosses, ce qui n'est pas nécessairement le cas avec une lumière collimatée.

En se basant sur des stimuli produits avec un éclairage diffus, Langer et Bülthoff [LB00] ont montré que le SVH utilise en partie un a priori correspondant à l'observation physique précédente, qu'ils ont appelé *dark means deep* (« sombre égal profond »). Koenderink et al. [KvDCL96] ont également observé que les zones sombres génèrent plus souvent une impression de concavité. Plus récemment, Nefs et al. [NKK05] ont relevé que la variabilité des surfaces perçues lorsqu'on change la direction d'éclairage comporte une composante non linéaire déviant de l'ambiguïté du bas-relief [BKY97], et qui caractérise justement cette impression de profondeur plus importante en lien avec une région sombre.

1.1.3 Perception et importance de la direction d'éclairage

Si la direction de l'éclairage joue un rôle important dans notre perception des formes, notre système visuel est-il capable de la déterminer précisément ? En utilisant des stimuli assez simples et orientés à l'identique, Enns [ER90] et Sun [SP96b] notamment ont conclu que le système visuel était capable de détecter rapidement un objet dont la direction d'éclairage diffère de celle des autres présents dans une scène, et ce de façon préattentive. Cette propriété avait déjà été relevée dans un autre contexte par [Ram88] sur la base de gradients de contrastes induisant la perception de formes convexes et concaves. Ramachandran avait alors observé que l'image donnait lieu à un groupement perceptif basé sur la direction d'éclairage inférée, les formes des surfaces perçues dans l'image ne pouvant s'inverser que toutes en même temps et non isolément, ce qui est cohérent avec la perception d'un éclairage global cohérent. Cependant, Ostrovsky et al. [OCS05] ont mis en place une expérience semblable à celle de [SP96b], mais en utilisant des surfaces plus complexes et en faisant varier aléatoirement l'orientation de chaque objet. Ils ont alors observé que la capacité à isoler un objet éclairé est inversement liée au nombre d'objets (distracteurs) présents, ce qui tend à prouver que le profil d'éclairage d'une scène complète n'est pas détecté de façon préattentive, mais que cette information serait plus vraisemblablement extraite localement. Ainsi, le déplacement artificiel de certaines ombres dans une photographie de paysage ou dans un tableau au moyen d'un simple traitement d'image génère des incohérences qui ne sont que très rarement perçues [OCS05]. Cependant, cet effet peut dépendre de la familiarité de la scène représentée ou de la tâche à résoudre : la même opération opérée sur un visage induit une déformation notable. Une observation de tableaux de grands maîtres comme ceux de Fra Carnevale montre d'ailleurs que les artistes ont exploité cette particularité pour réaliser des œuvres physiquement fausses, mais dont l'impossibilité ou l'incohérence de certaines ombres échappe totalement à l'attention de celui qui les contemple naïvement [Cav05, Cav06]. Ces observations ne mettent pas en cause la capacité du système visuel humain à extraire localement une information de haut niveau comme le « champ de lumière » (voir à ce sujet [KPvD+07]), information pouvant être utilisée pour déterminer la forme d'une surface, mais discutent plutôt de la possibilité qu'une somme d'informations locales (directions d'éclairage) soient intégrées dans un percept global cohérent.

1.1.4 Importance des reflets

Une des autres caractéristiques d'un éclairage concerne la production de reflets spéculaires. Leur présence est déterminée par l'aspect microscopique du matériau, sa rugosité : une surface parfaitement

lisse, comme un miroir, produira d'importants reflets, alors qu'une surface mate n'en comportera aucun. Leur position dépend non seulement de l'angle entre la source lumineuse et la normale à l'objet, mais également de la direction de vue. En théorie, les reflets fournissent donc une information supplémentaire sur l'orientation des surfaces observées.

En utilisant un objet dont la surface se comporte comme un miroir pur, Fleming et al. [FTA04] ont même montré que la forme de l'objet peut être perçue avec une précision importante, et ce en dépit de l'absence de connaissance sur la structure du monde environnant se reflétant dans l'objet. Cependant, Savarese et al. [SFFP04] ont obtenu un résultat contraire. En masquant les contours occultants d'un miroir elliptique ou hyperbolique réfléchissant une grille régulière, ils ont en effet relevé de très mauvaises performances concernant la capacité à discriminer la courbure du miroir.

Dans des configurations plus courantes, c'est-à-dire pour des matériaux non totalement lisses, plusieurs études ont montré que les reflets de spécularité apportent une information effectivement utilisée par le SVH pour améliorer l'extraction d'information de forme [TM83, MT86, NTO04] par l'ombrage.

1.1.5 Ombres portées

Jusqu'à présent, nous nous sommes principalement intéressés au profil d'ombrage d'une surface engendrée par sa forme et la nature de l'éclairage. Nous avons vu que le système visuel est capable, avec certaines limitations, de résoudre le problème d'extraction de la forme d'une surface continue.

Nous allons maintenant considérer l'ombrage d'une scène comportant plusieurs objets distincts : un éclairage non complètement diffus génère des *ombres portées* plus ou moins douces qui sont une source d'information non négligeable quant à la position des différents objets dans l'espace.

La perception des ombres portées constitue en fait un problème relativement complexe, car elle nécessite d'une part de reconnaître une ombre en tant que telle (identification), et d'autre part de l'associer à sa source, l'objet l'ayant projetée (correspondance). A ce niveau, l'étude de nombreux tableaux montrent que nous sommes relativement tolérants tant en ce qui touche à la consistance de l'ombre elle-même [Cav05] qu'au réalisme physique de sa forme exacte [Mam04, Cas08]. En imagerie de synthèse, ce dernier point peut être exploité pour produire des ombres perceptivement satisfaisantes sur la base de représentations géométriques très simplifiées [SSMK05].

Une fois ces problèmes résolus, il a été démontré qu'une ombre portée apporte principalement de l'information quant à l'arrangement spatial d'une scène, plutôt que sur la forme des objets ou de la surface sur laquelle elle se trouve [MKK98]. Ainsi, le simple déplacement d'une ombre suffit à donner l'impression du mouvement en profondeur d'un objet pourtant statique à l'écran [KKMB96], et la modification du trajet d'une ombre altère en conséquence le déplacement perçu de l'objet projetant cette ombre. Comme dans le cas de l'interprétation de l'ombrage d'une surface, le système visuel doit résoudre un problème sous-contraint. En effet, le déplacement d'une ombre peut être dû au mouvement d'un objet ou de la source de lumière elle-même. Il semble ici que l'hypothèse d'une source de lumière stable constitue le principal a priori à l'œuvre [MKK98]. Dans une scène virtuelle statique, plusieurs études ont mis en évidence l'apport important des ombres portées pour la perception de la profondeur de différents éléments de la scène [WFG92, HWSB99]. Cependant, le nombre d'objets visualisé est ici relativement faible, et ces résultats seraient sans doute plus nuancés dans le cas d'affichages plus denses pour lequel la résolution du problème de correspondance serait moins aisée. Remarquons pour finir que la présence d'ombres portées permet également de discriminer plus facilement des surfaces convexes et concaves qui seraient confondues en leur absence [EKK93, LT04].

1.2 Eclairage en informatique graphique

La modélisation de l'éclairage constitue un domaine de recherche très actif depuis les débuts de l'informatique graphique. De nombreuses techniques ont été développées, répondant à des contraintes différentes en terme de propriété des images générées, de temps de calcul ou d'espace de travail (objet ou image).

Ainsi, il peut être nécessaire de générer une image la plus proche possible de celle que l'on obtiendrait via un appareil photographique, par exemple pour comprendre quel sera l'éclairage dans un bâtiment n'existant encore que virtuellement ; a contrario, le but recherché peut être un rendu plus expressif, s'éloignant délibérément du réalisme physique afin de renforcer des caractéristiques particulières de la scène 3D par des méthodes souvent inspirées d'œuvres d'artistes et d'illustrateurs, et indirectement de connaissances sur la perception visuelle chez l'homme. Dans la littérature, ces deux catégories sont généralement désignées respectivement par les termes de rendu photoréaliste et rendu non photoréaliste (NPR, *Non-Photorealistic Rendering*).

Par ailleurs, le temps de rendu peut être un facteur critique. Si une image de film d'animation peut être calculée en plusieurs heures, l'interaction avec un objet virtuel nécessite de limiter drastiquement la quantité de calculs à réaliser, ce qui se traduit le plus souvent par l'application de modèles simplifiés, limités à un aspect spécifique de l'éclairage. L'exemple le plus connu et le plus couramment implémenté dans les logiciels traitant de 3D est le modèle d'éclairage de Phong, qui décrit simplement et localement la façon dont un rayon de lumière est réfléchi par une surface.

L'espace dans lequel les calculs sont réalisés permet également de distinguer les différents algorithmes d'éclairage. Si la plupart travaillent dans l'espace objet 3D, l'éclairage de la scène étant calculé avant le rendu ou au moment de la projection de la géométrie, certains réalisent les calculs comme une étape de posttraitement de l'image produite, éventuellement enrichie d'informations géométriques telles que la profondeur des pixels ou la normale associée [DWS+88, ST90]. Le travail en espace image permet de limiter le nombre de calculs à réaliser aux objets effectivement visibles dans l'image ; en contrepartie, il ne permet pas de traiter certains aspects comme le rendu de surfaces transparentes.

Si une description et une taxonomie complètes sont ici hors de propos, nous présenterons dans cette section une vue générale des différents types d'algorithmes existants, et nous détaillerons plus particulièrement certains types de modèles.

1.2.1 Méthodes d'éclairage global

Le rendu photoréaliste d'une scène nécessite de modéliser précisément le trajet de la lumière depuis les sources primaires jusqu'au plan image, en considérant les phénomènes de réflexions et réfractions opérant au niveau des différentes surfaces rencontrées. En plus d'une définition précise des émetteurs, il est nécessaire de caractériser la manière dont la lumière est réfléchie par les matériaux rencontrés, par exemple au moyen de fonctions 4D dépendant de la direction d'incidence et de réflexion des rayons lumineux, les *BRDF* (Bidirectional Reflectance Distribution Function). Ces paramètres étant connus, différentes méthodes existent pour calculer un *éclairage global* (*global illumination*), c'est à dire tenant compte de toutes les réflexions et interréflexions des rayons lumineux. Les techniques de *lancer de rayon* (ray tracing [Whi80]) parcourent à l'inverse le chemin de lumière, du plan image vers les sources. Dans un autre registre, les méthodes de *radiosité* [GTGB84, NN85] abordent le problème sous un autre angle : elles calculent l'état d'équilibre radiatif de la scène 3D considérée comme un système fermé, le résultat pouvant être utilisé pour produire des images depuis un point de vue quelconque. Contrairement au lancer de rayons, les méthodes de radiosité permettent notamment de tenir compte de la dispersion de la lumière.

1.2.2 Méthodes d'éclairage local

Même si certaines accélérations existent (par exemple [DSDD07]), les techniques d'éclairage global sont très gourmandes en calcul et ne peuvent être appliquées lorsque l'interactivité est une contrainte forte. Il faut alors recourir à des méthodes locales, c'est à dire ignorant le parcours complet des rayons lumineux, et souvent basées sur une simplification des modèles physiques de réflexion. On trouve ici les éclairages les plus couramment utilisés pour le rendu interactif, à savoir les techniques de Gouraud [Gou71], Phong-Blinn [Pho75, Bli77] et Cook-Torrance [CT82] qui définissent des modèles empiriques de réflexion et proposent différents schémas d'interpolation des calculs sur les surfaces à éclairer. Par exemple, l'ombrage I d'un élément de surface de normale $\vec{N}$ éclairé par une source lumineuse de direction $\vec{L}$ et vu d'une direction $\vec{V}$ (voir figure 4.2 (a)) est calculé par Phong comme la somme de trois composantes, ambiante, diffuse et spéculaire :

$$I = A + k(D\vec{N}.\vec{L} + S(\vec{R}.\vec{V}))^n \text{ avec } \vec{R} = 2(\vec{N}.\vec{L})\vec{N} - \vec{L}$$

Le champ de lumière est ici représenté très localement, simplement par sa direction, en ne tenant pas compte du blocage éventuel des rayons lumineux par des surfaces interposées. Le rééclairage induit par l'ensemble des interréflexions est considéré comme constante en chaque point de l'espace (A). Par conséquent, deux faces d'un objet orientées identiquement produiront le même éclairage, à la spécularité près qui dépend de l'angle de vue, même si l'une se situe dans un creux et l'autre sur une bosse de la surface, rendant plus difficile l'interprétation du relief entourant ces faces.

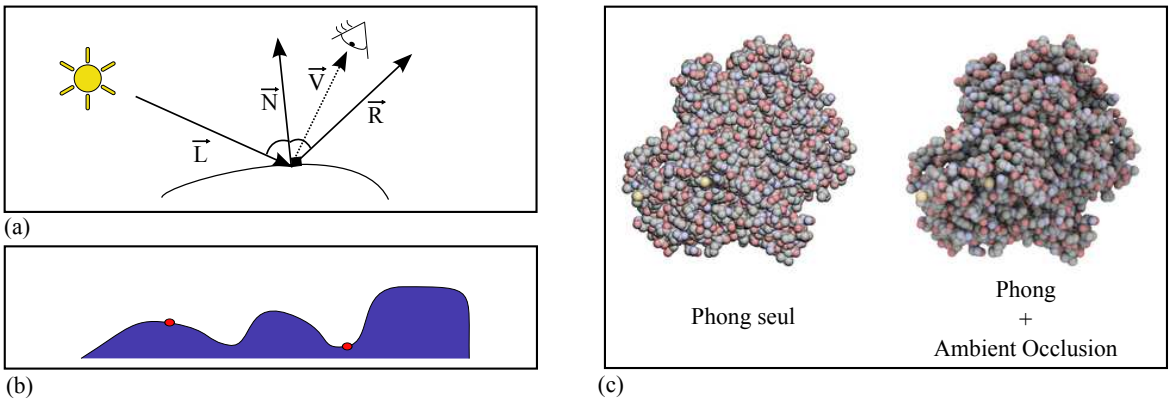


FIG. 4.2: (a) Dans le modèle d'éclairage local de Phong, l'ombrage d'un élément de surface dépend seulement de l'orientation de la surface $\vec{N}$, de la direction de lumière $\vec{L}$ (composante diffuse et spéculaire) et du point de vue $\vec{V}$ (composante spéculaire). (b) Le calcul d'ambient occlusion consiste à déterminer pour chaque élément de surface la portion de ciel visible, la quantité d'ombrage étant définie à partir de cette fraction. (c) Ce terme supplémentaire, ajouté au modèle local de réflexion de Phong, permet de mieux distinguer le relief en obscurcissant les creux (images otenues à partir du logiciel de visualisation de molécules Qutemol [TCM06]).

1.2.3 Eclairage ambiant

Pour améliorer le rendu généré par un éclairage purement local de type Phong, de nombreux auteurs ont proposé de remplacer la constante affectée à l'éclairage ambiant, A, par un terme dépendant de la géométrie environnant l'élément de surface considéré. L'idée générale consiste ici à considérer que la scène représentée est entourée d'une sphère apportant un éclairage de fond isotrope. Le terme ambiant est alors défini par la portion de ce ciel de lumière visible en un point P (voir figure 4.2(b)),

ce qui s'écrit

$$A(P,\vec{N}) = \frac{1}{\pi}\int_{\Omega}(1-V(\vec{\omega}))(\vec{N}.\vec{\omega})d\omega$$

où $V(\vec{\omega})$, défini sur la demi-sphère Ω orientée par $\vec{N}$, renvoie 1 si le rayon de direction $\vec{\omega}$ partant de P intersecte une surface occultante, et 0 s'il atteint la sphère de lumière. La contribution ambiante ainsi calculée procure une indication supplémentaire facilitant la perception de la profondeur, de la courbure locale et de la proximité spatiale entre objets (cf figure 4.2(c)). En effet, il répond à l'observation qu'une partie de la perception de la forme par l'ombrage est déterminée par l'attribution d'une profondeur plus importante à des zones relativement plus sombres de l'image [LB00, KvDCL96, NKK05].

L'idée de l'éclairage ambiant, communément désigné par *ambient occlusion* ou rendu en portion de ciel visible (*skydome rendering*), a été introduite par Miller [Mil94] sous le nom d'*accessibility shading*. Zhukov et al. [ZIK98, IKSZ03] s'en inspirent pour précalculer des cartes d'obscurance, la fonction de visibilité introduite précédemment étant modulée par la distance des surfaces occultantes. Ces cartes sont utilisées interactivement pour moduler localement l'intensité de l'éclairage de type Phong, par exemple dans les jeux vidéos.

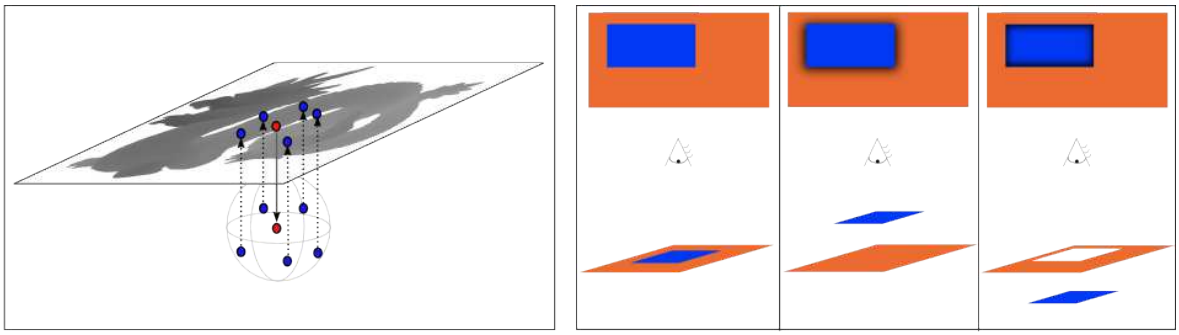
Si l'ambient occlusion apporte une information visuelle importante, le temps nécessaire à son calcul demeure problématique. En effet, il nécessite en théorie, pour chaque élément de surface, de déterminer la portion de ciel cachée par l'ensemble des autres éléments de la scène. De nombreuses méthodes et optimisations ont été proposées suivant la nature des objets représentés, par exemple pour le rendu de molécules [TCM06], d'arbres [HPAD06] ou de gros nuages de points [DGM04]. Lorsque l'objet visualisé est assimilable à un champ de hauteur, le calcul peut ainsi être restreint à la détermination des hauteurs maximales entourant un point dans chaque direction du plan (*horizon mapping* [Max88, Ste98, DDGM+04]). Une autre optimisation, proposée par Borgo et al. [BCS01], consiste à rendre la scène depuis un nombre restreint de points de vue échantillonnés sur la sphère, et à approcher l'éclairage ambiant en un point de surface par le rapport du nombre d'images où ce point est visible ; cela permet de tirer parti de l'accélération de la carte graphique, mais le calcul demeure trop long pour être réalisé interactivement.

L'essentiel des méthodes mentionnées jusqu'ici nécessitent des calculs trop lourds pour être exécutés au moment du rendu, ce qui limite leur usage à des objets statiques. Pour des scènes dynamiques, Kontkanen et Laine [KL05] ont proposé une méthode de précalcul de champ volumique de distances par objet permettant par la suite de reproduire en temps réel les ombres liées aux occlusions locales. Plus récemment, l'évolution importante de la puissance de calcul des cartes graphiques a permis l'éclosion de nouvelles méthodes génériques de calcul du terme d'ambient occlusion. Contrairement aux précédentes implémentations sur GPU qui travaillaient en espace objet et s'appuyaient sur un stockage optimisé de versions simplifiées de la géométrie (par exemple des disques [Bun05]), ces nouveaux algorithmes réalisent un ombrage en posttraitement (*deferred shading* [DWS+88]) s'appuyant sur les cartes de profondeur et de normales enregistrées au moment de la projection. Shanmugan et Arikan [SA07] calculent ainsi pour chaque pixel l'occultation induite par quelques voisins échantillonnés dans l'espace image et reprojetés dans l'espace objet sous forme de sphères. Pour cela, ils font remarquer que l'ombrage principal d'une surface est engendré par les voisins proches d'un objet, la projection induisant nécessairement leur proximité dans l'image ; la réciproque n'est pas vraie, mais une atténuation des contributions en fonction de la distance 3D permet de limiter les artefacts liés à cette approximation. Une approche légèrement différente est à l'œuvre dans le moteur de jeu *CryEngine2* de la société Crytek [Mit07] : pour chaque pixel, K points sont échantillonnés dans une boule représentant le voisinage 3D du point avant projection, et l'ambient occlusion correspond au nombre de ces points visibles après projection, c'est à dire dont la profondeur est inférieure à celle enregistrée dans l'image (voir figure 4.4 (a)). Cette technique, ajoutée à d'autres astuces et à certains prétraitements, appliquée dans le jeu *Crysis*, permet de produire un éclairage convaincant en

temps réel sur des scènes dynamiques complexes. Une autre méthode interactive adaptée aux nouvelles cartes graphiques a été proposée récemment par Dimitrov et al. [DBS08] : elle consiste en une adaptation de l'horizon mapping [Max88, Ste98] au GPU, en considérant l'image de profondeur comme un champ de hauteur. Le point commun de ces techniques interactives sur GPU réside dans la définition d'un échantillonnage du voisinage de chaque pixel, le nombre de voisins utilisés définissant un compromis entre le temps de calcul et la précision ou le niveau de bruit de l'ombrage. En général, elles sont réalisées en variant la définition du voisinage d'un pixel à l'autre, pour éviter les effets de bande (*banding effect*), et en lissant légèrement l'image obtenue (voir figure 4.3).

(a) (b) (c) (d) (e)

FIG. 4.3: Exemple de calcul d'ambient occlusion en espace image (algorithme adapté de [Mit07]), utilisant uniquement la carte de profondeur (a). Pour éviter les effets de bande (b), les points définissant le voisinage varient d'un pixel à l'autre (c). Le bruit généré par cette variation aléatoire peut être limité en augmentant le nombre de voisins (8 en (b) et (c), 16 en (d)) et en lissant l'image finale ((e), 16 voisins).



(a) (b)

FIG. 4.4: (a) Principe de l'algorithme de calcul d'ambient occlusion en espace image [Mit07] : le pixel rouge est rétroprojeté dans l'espace 3D, où son voisinage est échantillonné (points bleus). Les points correspondants sont reprojetés en espace image, où leur profondeur est comparée à celle de la carte de profondeur enregistrée. (b) L'unsharp masking basé sur le tampon de profondeur proposé par [LCD06] permet de détourer les objets en fonction des différences de profondeur dans l'espace image, ce qui facilite la perception des relations spatiales entre objets occultants.

1.2.4 Autres algorithmes

Les modèles d'ombrage diffus local de type Phong et d'ambient occlusion constituent des modèles très approchés d'ombrage de surface visant à reproduire certaines des caractéristiques d'un éclairage global. Le premier définit l'ombrage en fonction de l'orientation de la surface relativement à une source lumineuse supposée directionnelle (de type journée ensoleillée), alors que le second suppose un dôme d'éclairage uniforme (éclairage diffus, jour nuageux, ou beaucoup d'interréflexions dans l'environnement) et calcule la fraction de cette source effectivement visible depuis la surface.

Pour terminer ce rapide tour d'horizon, nous mentionnerons simplement l'existence d'algorithmes spécifiquement dédiés au rendu d'ombres portées. Ceux-ci visent à reproduire les zones d'ombre et de pénombre causées sur une surface en raison de l'occultation totale ou partielle de la source de lumière par un objet interposé. Les sources lumineuses sont supposées ponctuelles ou peu dispersées ; dans le cas contraire, la configuration serait en effet proche d'un rendu d'ambient occlusion. Il existe essentiellement deux types d'algorithmes de calcul d'ombres portées. Dans le cadre du *shadow mapping*, la scène est rendue une première fois du point de vue de la source de lumière, puis du point de vue de la caméra et les obtenus dans cette deuxième passe sont transformés dans le repère de la source de lumière. Leur profondeur est alors comparée à celle enregistrée lors de la première passe, ce qui détermine s'ils voient ou non la lumière. La méthode des *shadow volumes* définit dans un premier temps les volumes d'ombres associés à chaque objet de la géométrie, et utilise cette information 3D pour calculer l'ombrage au moment du rendu. Généralement, les shadow maps sont plus rapides à calculer que les shadow volumes, mais produisent plus d'artefacts en raison notamment de la pixellisation qu'elles induisent. Ici encore, il est possible d'utiliser des méthodes de lancer de rayon dans l'espace image pour produire un résultat approché convenable en un temps raisonnable (voir par exemple [ARHM00]) ; cependant, une telle technique ne permettra pas de générer l'ombre portée d'un objet masqué du point de vue de la caméra mais visible depuis la source de lumière.

1.2.5 Modèles d'ombrage non photoréalistes

Si la plupart des rendus mentionnés dans les paragraphes précédents ne peuvent pas être objectivement qualifiés de photoréalistes au regard de la qualité des images produites, ils cherchent néanmoins à rendre compte d'aspects particuliers de la réalité physique de l'ombrage. Leurs limitations sont dues principalement aux ressources disponibles, qu'il s'agisse de temps de calcul, d'espace mémoire ou de connaissances sur les propriétés physiques des objets représentés. A contrario, les rendus non photoréalistes d'ombrage dévient explicitement des lois physiques, dans le but de faciliter la perception des caractéristiques principales des objets visualisés, qu'il s'agisse de leur forme ou de leur organisation dans l'espace. Pour cela, ils s'appuient soit directement sur des critères perceptifs, soit sur des techniques à l'œuvre dans des tableaux, dessins et illustrations techniques qui dénotent souvent de la part des artistes une connaissance empirique relative à notre perception visuelle [Cav05]. Comme le dit Gombrich : « L'artiste ne s'intéresse pas aux causes, mais aux mécanismes de certains effets. Le problème qui se pose à lui est de nature psychologique : il faut que l'image évoquée soit convaincante, bien qu'aucun des tons dont le peintre peut disposer ne corresponde à ce que nous appelons le 'réel'. » ([Gom02],p.44). Au delà du problème de la tonalité évoqué ici, les artistes ont depuis longtemps dû trouver des techniques permettant de résoudre un de leurs problèmes majeurs : communiquer le sens de l'espace au moyen d'une représentation sur un support plat.

Une première approche en ombrage NPR consiste à produire des images de type dessin animé (*cartoon-like rendering*), par exemple en restreignant le dégradé de luminances résultant d'un calcul de type Phong à un nombre restreint de teintes, avec des transitions brusques [Dec96], et en renforçant les silhouettes des objets [ST90, Dec96]. Pour cela, une fonction de transfert 1D définit le lien entre la quantité d'ombrage et la teinte produite. Cette idée a été étendue en utilisant des fonctions de transfert plus continues ou plus complexes. Barla et al. [BTM06] ajoutent par exemple la profondeur des points relativement à la caméra en entrée de cette fonction, ce qui produit un ombrage NPR liant éclairage relatif aux attitudes des surfaces et perspective atmosphérique. Dans le même ordre d'idées, Sloan et al. [SMGG01] proposent d'utiliser une sphère texturée reliant directement les informations de normales à la teinte désirée. Dans ce qui précède, l'ombrage est rendu par des changements de teinte, et pas simplement de luminance. Ce principe, directement inspiré de tableaux et d'illustrations techniques, a été proposé par Gooch et al. [GGSC98] afin notamment de limiter les zones trop sombres

et d'augmenter les nuances résultant de l'éclairage. Si tous ces algorithmes améliorent les contrastes et le rendu d'un simple éclairage Phong, ils conservent la localité de l'algorithme, reposant principalement sur l'orientation de l'élément de surface considéré. Plus récemment, Vergne et al. [VBGS08] ont implémenté un descripteur de forme inspiré de celui de Koenderink et van Doorn [KvD92]. Ce descripteur repose sur deux paramètres dérivés de mesures de courbure en espace objet et image, et une texture 2D est ensuite définie pour associer à chaque pixel une valeur d'ombrage en fonction de cette mesure.

S'appuyant sur l'incapacité du système visuel à percevoir les incohérences d'éclairage pour une scène un peu complexe, Lee et al. [LHV04] découpent la surface à éclairer en sous-régions auxquelles ils appliquent des directions d'éclairage différentes. Celles-ci sont générées en maximisant une fonction dépendant des intensités d'ombrage diffus et spéculaire en chaque vertex. Cette idée est inspirée des travaux de Shacked et Lischinsky [SL01] qui contrôlent la position de l'éclairage en optimisant une métrique perceptive relative aux images produites. Rusinkiewicz et al. [RBD06] reprennent cette idée de direction d'éclairage variable, en remarquant par ailleurs que les détails d'un relief sont plus facilement perceptibles si la lumière est rasante. Ils proposent d'adapter localement la direction d'éclairage orthogonalement à la normale, et en lien avec la direction de courbure principale. Les champs de normales et courbures utilisés sont filtrés à plusieurs résolutions, et l'ombrage global est une somme pondérée des différents niveaux, en utilisant un toon-shader adouci [Dec96], ce qui permet de faire ressortir les différentes fréquences de variations du relief. L'idée du lissage du champ de normales avait été mis en œuvre auparavant par Cignoni et al. [CST05], pour générer des transitions douces d'ombrage au niveaux des angles saillants d'une géométrie de type CAO, un éclairage Phong basique ayant tendance à aplatir l'image perçue dans ce cas. Si les méthodes précédentes donnent des résultats visuellement convaincants pour les objets étudiés, le plus souvent une unique surface continue et relativement lisse, elles nécessitent cependant des prétraitements assez lourds (calculs de courbures, lissage de champs de normales ou optimisation). Par ailleurs, d'autres auteurs se sont intéressés à la même problématique, mais sur la base de photographies d'un même objet mais prises avec différentes directions d'éclairage. L'objectif est ici de proposer des algorithmes automatiques [FAR07] ou guidés par l'utilisateur [ALK+03] permettant d'améliorer la perception du relief à différentes échelles en mélangeant localement les différentes sources enregistrées. Toutes ces méthodes, en espace objet ou image, utilisent soient un éclairage réel ou un modèle d'ombrage local proche de celui de Phong. A contrario, pour faciliter l'analyse visuelle de surfaces lisses, Weiskopf et Hauser [WH06] proposent une déviation du modèle de Phong qu'ils appellent *cycle shading*. L'idée consiste à introduire dans le modèle plusieurs directions principales de spécularité au lieu d'une seule, ce qui génère un ombrage variant par cycles en fonction de l'inclinaison de la surface. Une étude utilisateurs montre l'apport de cette technique pour détecter une discontinuité de la surface visualisée, mais il parait peu envisageable d'utiliser ce modèle dans d'autres cas, la texture générée en surface de l'objet perturbant la perception de la forme d'ensemble.

Les techniques exposées jusqu'ici s'attachent essentiellement à améliorer la perception du relief dans les images générées, les illustrations proposées étant le plus souvent réduites à une simple surface. Dans des scènes plus complexes, de nombreux objets sont situés à des profondeurs différentes, et même un modèle d'éclairage global n'apporte pas toujours une séparation nette entre deux objets décalés dans l'espace mais se recouvrant partiellement dans l'image (occultation). Luft et al. [LCD06] observent que de nombreux peintres altèrent localement le contraste autour d'une silhouette afin de la détacher plus nettement de son arrière plan. Ils proposent alors de transposer cette idée pour améliorer le rendu d'images avec information de profondeur, en utilisant un opérateur de contraste classique en traitement d'images, l'*unsharp masking*. Leur opérateur calcule la différence entre la profondeur d'un pixel et celle obtenue par filtrage gaussien. Cette valeur peut alors être utilisée pour créer un effet de halo autour des objets les plus proches, ce qui facilite effectivement la perception en profondeur dans

l'image (voir illustration de l'effet sur la figure 4.4). L'utilisation de halos pour faciliter la séparation en profondeur dans des scènes complexes de visualisation scientifique a par ailleurs été proposée pour le rendu volumique de champs vectoriels [IFP97] ou scalaires [BG07], ainsi que pour la visualisation de molécules [TCM06]. Elle est également pleinement intégrée aux outils de fenêtrage de la plupart des systèmes d'exploitation récents. Ces effets halos ne constituent pas en soi un ombrage, ils représentent plutôt une information supplémentaire que l'on peut apporter à une scène éclairée autrement. Cependant, il est possible de les obtenir photographiquement, en utilisant un éclairage diffus positionné autour de l'objectif (ils correspondent alors à des zones de pénombre).

1.3 Eclairage en visualisation scientifique

La visualisation scientifique occupe une place à part dans le domaine de l'infographie, en raison de la diversité des modèles de représentation à l'œuvre. Il ne s'agit pas seulement de représenter des surfaces continues, telles que les limites du domaine étudié ou quelques isosurfaces, mais également des objets de dimension inférieure, comme des lignes de courant ou des points de Gauss. Pour cela, il n'est pas rare que les rendus fassent appel à des glyphes ou imposteurs, *point sprites* par exemple. La mise en place d'un modèle d'éclairage peut alors se révéler assez problématique, et les algorithmes proposés dans les publications sont très généralement adaptés à une représentation bien particulière. En ce qui concerne les principaux logiciels de visualisation (VTK, Ensight...), ils ignorent bien souvent le problème de l'éclairage, se restreignant au modèle local de Phong.

Pour la visualisation de lignes de courant, Zöckler et al. [ZSH96] proposent une adaptation du modèle de Phong permettant d'améliorer la perception de nombreuses lignes à l'écran. Dans le cadre d'un rendu de type LIC (Line Integral Convolution) volumique, Interrante et Grosch [IFP97] facilitent la séparation visuelle des éléments dans le volume en ajoutant un effet de halo. Plus récemment, d'autres rendus mélangeant Phong, halos et effets d'ombres ont été développés [MTHG03, SGS05].

Pour certaines données, il est possible de précalculer l'éclairage global ou certaines de ses caractéristiques, et d'utiliser le résultat de manière interactive. Ce principe a été appliqué notamment pour des isosurfaces extraites de grilles régulières. Stewart [Ste03] réalise ainsi un calcul d'ombrage de type ambient occlusion localisé (*vicinity shading*) en tout point de la grille, les valeurs enregistrées pouvant ensuite être utilisées interactivement pour n'importe quelle isosurface extraite de la grille. Banks et Beason [BB07] utilisent le même principe, mais en précalculant une contribution d'éclairage global plus complexe. Pour la visualisation de particules représentées par des sprites, Gribble et Parker [GP06] réalisent le même type de prétraitement, les valeurs d'éclairage étant stockées et compressées dans des textures 2D.

Notons enfin que plusieurs études perceptives ont été menées dans le cadre de la visualisation pour évaluer l'impact du modèle d'éclairage sur la perception de la forme et des relations spatiales entre objets. Toutes concluent qu'un éclairage global apporte une information non négligeable par rapport à un modèle purement local, qu'il s'agisse de discriminer deux formes presque identiques [FWSP04] ou de comprendre la structure globale et les relations spatiales opérant dans un champ dense de particules [GP06] ou de lignes de courant [WB08].

1.4 Conclusion : caractéristiques importantes pour un éclairage

Après cette revue assez large des modèles d'éclairage existants, il est important de définir les objectifs que l'on souhaite voir remplis par un éclairage destiné à la visualisation scientifique.

En premier lieu, il doit contribuer à faciliter la perception de la forme ainsi que des relations spatiales entre objets. De ce que nous avons vu, un éclairage local basé uniquement sur les informations de direction de lumière et de surface n'est pas satisfaisant, et il paraît nécessaire de disposer d'un

terme permettant de distinguer les bosses et les creux. En ce qui concerne les relations spatiales entre objets, les effets de halo inspirés de principes artistiques ont prouvé une certaine efficacité visuelle.

La nature même des objets visualisés impose une double contrainte. D'une part, il s'agit souvent de représentations tirées de concepts mathématiques (isosufaces, lignes de courant...). L'idée de photoréalisme est ainsi à écarter : les images produites doivent convaincre par l'information qu'elles véhiculent, et non par leur ressemblance possible avec une véritable photographie. De plus, nous avons déjà mentionné la diversité des objets représentés, une même scène pouvant comporter des surfaces, des lignes ou divers glyphes (flèches, sprites) ; beaucoup de techniques ont été développées pour une représentation particulière, mais il serait utile de disposer d'un modèle plus général.

Par ailleurs, l'interactivité constitue un des points clés de la visualisation, au moins dans la phase d'exploration des données. Comme l'ombrage ne représente qu'une sous-partie du rendu total d'une scène 3D, il parait nécessaire ici de disposer d'une méthode consommant le moins de temps de calcul possible. En raison de la quantité de plus en plus grande de données générées par les simulations numériques actuelles, nous écarterons l'idée d'un prétraitement qui s'avérerait vraisemblablement trop coûteux, d'autant plus que de nouvelles structures peuvent être extraites au moment même de la visualisation.

Enfin, suivant les objets représentés, une lumière fixe frontale ou par le haut ne représente pas nécessairement le meilleur choix. On constate intuitivement qu'un déplacement de la source lumineuse peut aider à la compréhension de structures complexes, permettant de mieux mettre en évidence certaines régions autrement peu contrastées. Ce déplacement nécessite en outre une certaine interactivité, ce qui rejoint l'impératif précédent d'un modèle d'éclairage peu coûteux en calculs.

2 Algorithme de l'EyeDome Lighting

2.1 Aperçu

Nous proposons ici un nouvel algorithme d'éclairage interactif inspiré des considérations détaillées précédemment. Pour répondre aux impératifs d'interactivité et de généralité au regard des données représentées, les calculs reposent uniquement sur une exploitation de la carte de profondeur obtenue après projection de la scène. Il s'agit donc d'une technique d'ombrage différé (*deferred shading* [DWS+88]). Le nombre d'opérations à réaliser est volontairement restreint, avec un temps de calcul dépendant principalement de la taille de l'image. De plus, une approche multirésolution permet de considérer différentes échelles spatiales à moindre coût. Dans ce qui suit, nous ne nous préoccuperons donc pas de la taille des données géométriques affichées, dans la mesure où celle-ci ne détermine en rien l'évaluation de l'ombrage.

Dans son implémentation de base, cet éclairage se comporte un peu comme si chaque pixel accédait à un dôme de lumière très local orienté par la direction de vue, l'œil de l'utilisateur, d'où le nom « EyeDome Lighting » (EDL). L'analogie s'arrête ici, l'EDL n'ayant pas d'équivalent physique : il s'agit bien d'un ombrage non photoréaliste, dont les caractéristiques sont au croisement d'un ombrage diffus, d'un modèle d'ambient occlusion et d'une génération de halos. Afin de permettre une exploration de la scène par la lumière, par exemple pour rééclairer interactivement des zones plus sombres, une extension directionnelle sera également présentée.

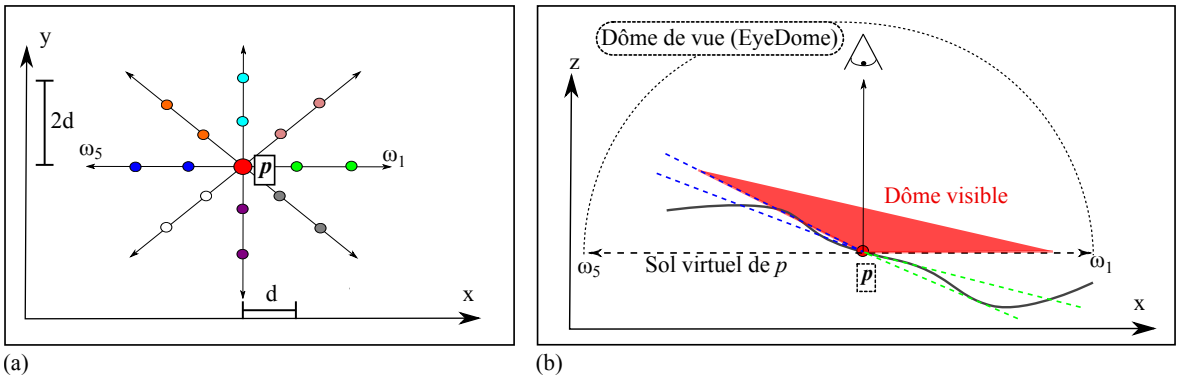


FIG. 4.5: (a) Exemple de voisinage d'un pixel p, pour $k = 8$ directions et $n = 2$ points par direction. (b) L'ombrage en un pixel p est fonction de la portion de surface du pseudo-dôme de lumière visible depuis p.

2.2 Algorithme de base

2.2.1 Notations

Tous les calculs sont réalisés sur l'image de profondeur de la scène, obtenue après projection de la scène à visualiser lors d'une première passe de rendu hors écran et enregistrée dans une texture. Classiquement, l'image de profondeur est assimilée à un champ de hauteur discrétisé et normalisé :

$$Z = \{\, z = f(x,y) \,/\, (x,y) \in [0:1]^2 \,,\, z \in [0:1] \,\} \,,$$

(x,y) désignant les coordonnées d'un pixel dans l'espace image. Par convention, la verticale est orientée vers le point de vue, les objets les plus distants ayant alors pour hauteur $z = 0$, les plus proches $z = 1$. Cette définition en espace image est indépendante de la nature de la scène à ombrer, et peut s'appliquer tout autant à des objets triangulés, à des représentations implicites de surfaces ou à une image de profondeur obtenue par un autre moyen (photographie stéréo par exemple). Elle ne tient pas compte non plus, le cas échéant, du mode de projection utilisé lors du rendu (orthographique ou perspectif) ; si nécessaire, le champ de hauteur peut être linéarisé dans une phase de prétraitement.

Notons $\Omega = \omega$ l'ensemble des k directions également réparties autour de p. Le voisinage $\mathcal{V}_p$ d'un pixel p est défini dans l'espace image par les pixels se trouvant à une distance $n.d$ de p sur chacune des directions ω, d désignant la distance minimale du voisinage (norme euclidienne dans l'espace image), et n un entier ($n \in [1:N]$). Ce voisinage est illustré sur la Fig.4.5 (a). Il correspond simplement à un échantillonnage régulier des cercles de rayon $n.d$ centrés en p. Sauf mention explicite, les résultats exposés plus loin utilisent $k = 8$ et $n = 1$.

Enfin, nous appellerons $I(p)$ la valeur d'ombrage du pixel p. Il s'agit d'une valeur variant entre 0 et 1, du plus sombre au plus clair.

2.2.2 Calcul de la valeur d'ombrage

Le principe de base de l'algorithme est illustré sur la Fig.4.5(b). On considère pour chaque pixel p de hauteur z_p le dôme correspondant à la demi-sphère centrée en p et délimitée par le plan horizontal ($z = z_p$). L'ombrage en p est alors fonction de la quantité de ce dôme visible en p, ou, réciproquement, il est inversement déterminé par la quantité de ce dôme masquée par les voisins de p. Autrement dit, un voisin q de p réduira l'éclairage en p s'il se trouve plus haut que p, à savoir $z_q > z_p$. La quantité de ciel masqué est déterminée pour chaque rayon ω, elle est le maximum de ciel masqué pour chaque voisin

q appartenant à ω. Ces valeurs sont ensuite intégrées sur l'ensemble des rayons, pour déterminer la quantité totale du dôme de vue visible pour p.

Si l'on note $s(p,q)$ la portion de ciel de p masquée par q et $\mathcal{S}(p)$ la portion de ciel de p masquée par l'ensemble de son voisinage, les relations décrites précédemment s'expriment :

$$\mathcal{S}(p) = \sum_{\omega \in \Omega} \max_{q \in \omega \cap \mathcal{V}_p} s(p,q).$$

Nous avons mentionné plus haut que l'on ne connaissait pas nécessairement la géométrie du champ de hauteurs à éclairer, et partant pas sa métrique. Plus précisément, on ne connaît pas la relation reliant le z normalisé aux hauteurs réelles au regard des dimensions de l'image, le calcul du dôme visible est donc une approximation de la réalité, et n'a pas de signification physique réaliste. La fonction $s(p,q)$ peut donc être définie de différentes manières, tant qu'elle répond à deux critères principaux : elle doit être positive si $z_q > z_p$, nulle autrement ; et elle doit être croissante en fonction de l'élévation, c'est à dire l'angle formé par le sol en p et la direction $\vec{pq}$. Dans notre implémentation, nous avons choisi

$$s(p,q) = \left[\frac{z_q - z_p}{d_{pq}}\right]^+$$

où d_{pq} désigne la distance entre p et q dans le plan image, et $[x]^+ = max(x,0)$.

Pour calculer la quantité de lumière visible en p, il nous faut maintenant inverser S et obtenir une valeur normalisée. La valeur maximale de S est simplement égale à k, ce qui correspond au cas où un pixel d'altitude $z = 0$ est entouré de k voisins d'altitude $z = 1$. Cependant, il s'agit d'un cas extrême, et une inversion dure du type $I(p) = 1 - \mathcal{S}(p)/k$ générera des images beaucoup trop claires. Nous préférons donc utiliser la transformation plus progressive :

$$I(p) = \exp\left(-A\mathcal{S}(p)\right)$$

Le facteur A permet de d'amplifier ou réduire simplement l'intensité de l'ombrage généré.

La figure 4.6 illustre le résultat de l'algorithme, ainsi que l'effet produit par différentes fonctions s. Un calcul d'élévation par un angle factice ($s(p,q) = [\arctan(z_q - z_p)/d_{pq}]^+$) produit un résultat très semblable à l'implémentation de base, alors qu'une modification non linéaire des différences de hauteur (passage à la racine carrée (c) et au carré (d)) influence fortement les contrastes générés par les différences locales de hauteur.

2.2.3 Implémentation

D'un point de vue pratique, tous les calculs mentionnés ici et ci-après sont réalisés sur la carte graphique, en utilisant le langage de programmation de shaders *GLSL* (*OpenGL Shading Language*). La scène 3D à ombrer est tout d'abord rendue sous OpenGL hors écran en utilisant des *FrameBuffer Objects* (*FBO*), ce qui permet d'enregistrer les images de couleurs et de profondeur dans des textures 2D. Pour calculer l'ombrage, la texture correspondant à la carte de profondeur est redessinée à l'écran dans un rectangle de taille identique, en activant les shaders de l'EDL. On obtient ainsi en sortie la texture d'ombrage, qui peut être superposée à l'image de couleurs au moment du rendu final (par exemple en multipliant la carte de couleur par la valeur d'ombrage en chaque pixel). Ce pipeline de rendu, classique en ombrage différé, est illustré sur la figure 4.7.

2.3 Approche multirésolution

2.3.1 Principe

Tel que décrit précédemment, l'EDL est limité à un voisinage relativement restreint dans l'espace image. Seuls les pixels très proches contribuent à l'ombrage, à moins de considérer un voisinage plus

FIG. 4.6: Résultat produit par l'EDL en monorésolution, et pour différents calculs de de $s(p,q)$. (a) Fonction de base : $s(p,q) = [(z_q - z_p)/d_{pq}]^+$. (b) Calcul d'un pseudo-angle, le résultat obtenu est quasiment identique : $s(p,q) = [\arctan(z_q - z_p)/d_{pq}]^+$. (c) $s(p,q) = \sqrt{[z_q - z_p]^+}/d_{pq}$. (d) $s(p,q) = ([z_q - z_p]^+)^2/d_{pq}$. Les modifications non linéaires du champ de hauteur en (c) et (d) modifient fortement les contrastes dans l'image en fonction des différences locales de hauteur.

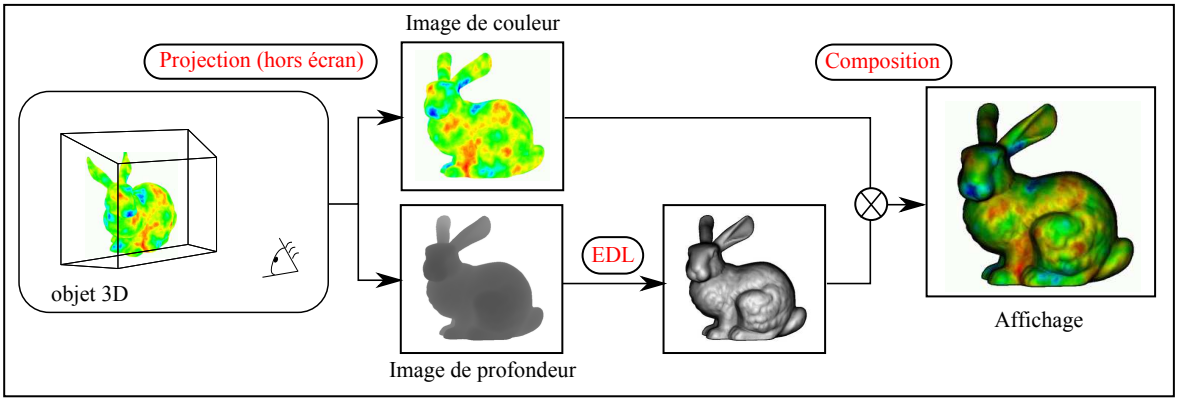


FIG. 4.7: Description de l'architecture de rendu intégrant l'EDL. La scène 3D (surface colorée à gauche) est d'abord projetée en utilisant un rendu OpenGL hors écran. Les images de couleur et de profondeur sont enregistrées dans des texture 2D. L'image de profondeur est utilisée pour calculer l'ombrage en appliquant le shader correpondant à l'EDL, et le résultat est utilisé pour moduler l'image de couleur dessinée à l'écran (par exemple en multipliant pour chaque pixel les composantes *RGB* par la valeur d'ombrage calculée).

étendu (nombre de voisins par rayon, n) ou un rayon de voisinage plus important (d). La première solution augmente rapidement le nombre de calculs à réaliser, et en plus épaissit d'autant les silhouettes générées par la fonction d'ombrage. En effet, une discontinuité de hauteur sature la fonction d'ombrage, ce qui créé une silhouette à l'image. D'après les formules utilisées, l'ombrage le long d'un rayon ω est déterminé par la contribution maximale des pixels inclus dans le voisinage le long de ω. Donc, si l'on considère des pixels distants de $n.d$, on risque le long d'une discontinuité de hauteur de créer une silhouette de même épaisseur, masquant localement des détails. L'augmentation du rayon de voisinage produira le même effet, en faussant de plus des contributions liées aux hautes fréquences du champ de hauteur dans l'espace image.

Pour prendre en compte les contributions de pixels plus lointains tout en limitant les temps de calcul et les artefacts, nous proposons de calculer l'ombrage à différentes résolutions d'image. Cette idée est classique en traitement d'images, la construction de pyramides d'images filtrées et sous-échantillonnées permettant d'effectuer des traitements à différentes échelles, mimant ainsi le fonctionnement de la rétine [AAB+84]. Dans notre cas, nous disposons initialement d'une image de profondeur de taille (w,h), obtenue après projection de la scène à ombrer. Pour obtenir la $i^{ème}$ résolution de l'EDL ($i \geq 1$), nous réalisons les calculs décrits en section 2.2.2 en dessinant l'image de profondeur sur une fenêtre de tailles respectives $(w/2^{i-1}, h/2^{i-1})$. L'image I_1 correspond ainsi à l'ombrage en pleine résolution, I_2 à l'ombrage de résolution divisée par 2, et ainsi de suite. L'ombrage total I est ensuite obtenu par une pondération de ces différentes résolutions :

$$I = \sum_i \alpha_i I_i$$

avec $\sum_i \alpha_i = 1$. En pratique, nous nous contentons de calculer les trois premiers niveaux. La figure 4.8 illustre le principe de cette approche multirésolution.

L'intérêt de cette démarche réside dans la diminution importante du nombre de calculs à réaliser. En effet, en divisant la taille de l'image par 2, le nombre de pixels à traiter est divisé par 4, et le nombre de calculs est ainsi réduit d'autant. Par ailleurs, l'algorithme étant implémenté sur le GPU, nous bénéficions des accélérations matérielles relatives au filtrage bilinéaire des textures, que ce soit lors de la projection sur une résolution inférieure (sous-échantillonnage) ou lors de la composition des différents niveaux (sur-échantillonnage).

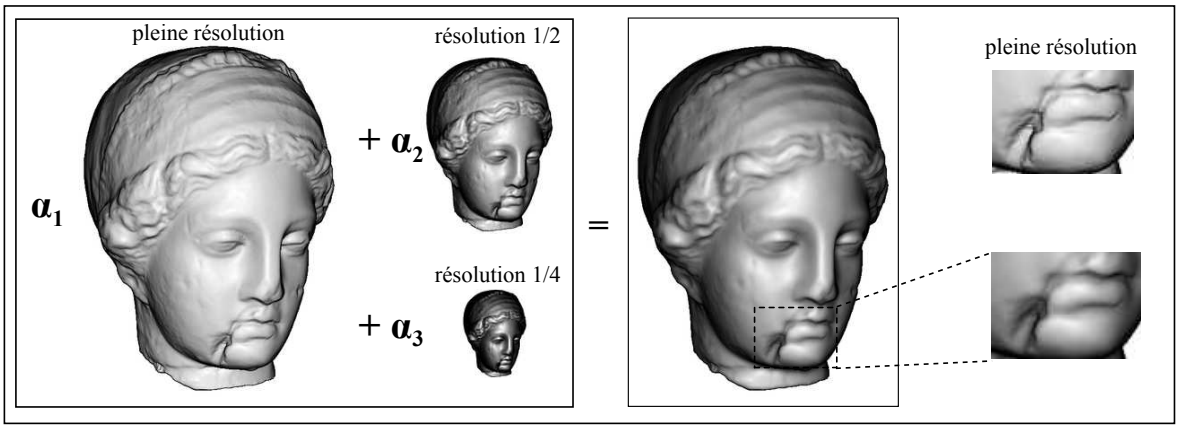


FIG. 4.8: Pour intégrer les contributions de pixels plus lointains, le calcul d'ombrage est réalisé à différentes résolutions, l'ombrage final étant obtenu en pondérant ces différents résultats. Dans cet exemple, cela permet par exemple de renforcer visuellement l'impression de creux au bord de la bouche.

2.3.2 Lissage par filtre bilatéral croisé

Le sur-échantillonnage réalisé au moment de l'intégration des résolutions inférieures à l'image finale, ajouté à la nature du voisinage utilisé pour l'EDL, génère un crénelage de l'image qui peut s'avérer perturbant. Celui-ci est particulièrement perceptible au niveau des discontinuités de hauteur, où le contraste est localement plus élevé. Pour éviter ce désagrément, il est nécessaire de lisser les images de résolution plus basse, tout en conservant les hautes fréquences calculées à pleine résolution (voir figure 4.9 (a)).

La solution la plus simple consiste à appliquer un flou gaussien à I_2 et I_3, technique classique en traitement d'image. La convolution par un noyau gaussien, réalisée dans un voisinage $\wp$ du pixel p, s'écrit

$$\widetilde{I}(p) = \frac{1}{n}\sum_{q\in\wp} I(p)g(p-q,\sigma_g) \qquad \text{avec} \qquad g(p-q,\sigma_g) = e^{-\frac{(x_p-x_q)^2+(y_p-y_q)^2}{2\sigma_g^2}}$$

où g désigne la gaussienne d'écart type σ_g et n le facteur de normalisation : $n = \sum_{q\in\wp} g(p-q,\sigma_g)$. D'un point de vue implémentation, la séparabilité de la gaussienne permet d'appliquer ce filtre sur le GPU en deux passes successives, horizontale puis verticale, ce qui pour un voisinage carré de taille N fait passer le nombre d'opérations à réaliser d'un facteur N^2 à un facteur $2N$ sans perte de précision.

Bien que rapide et efficace, cette solution n'est pas optimale car elle va introduire des bavures au niveau des discontinuités en profondeur (silhouettes). En effet, imaginons deux pixels voisins dans l'espace image, mais situés de part et d'autre d'une silhouette ; si ceux-ci présentent des valeurs d'ombrage très différentes, une simple convolution par un filtre gaussien va rapprocher leurs intensités finales respectives, ce qui conduit à réduire le contraste et à rendre plus difficile la perception des différences de profondeur dans la scène rendue.

Une solution à ce problème consiste à remplacer le flou gaussien par un filtre bilatéral croisé [ED04]. Pour lisser des images 2D tout en préservant les discontinuités (arêtes), [TM98] et [SB97] proposent de moduler le noyau gaussien appliqué aux couleurs par la différence d'intensité entre les pixels. Le filtre proposé, appelé filtre bilatéral par [TM98], permet d'éliminer le bruit au niveau de régions relativement homogènes tout en évitant les phénomènes de halos au niveau des discontinuités [DD02]. S'intéressant à la fusion de deux photographies prises sous des conditions d'éclairage différentes (avec et sans flash), [ED04] et [PSA+04] sont confrontés à un rapport signal sur bruit proche de 1 dans l'image la plus sombre, rendant impossible l'utilisation directe d'un filtre bilatéral ; pour contourner cet écueil, ils suggèrent donc d'utiliser un filtre bilatéral pour lequel la modulation du noyau gaussien est réalisée au moyen l'intensité de l'image prise avec flash (d'où le nom de filtre bilatéral croisé).

Dans le cas présent, nous souhaitons lisser les images I_2 et I_3 en évitant les bavures le long des discontinuités de hauteur. Nous proposons donc d'utiliser un filtre bilatéral croisé, le noyau gaussien de l'espace image étant modulé par une gaussienne relative aux différences de hauteurs attachées aux pixels dans l'image de profondeur initiale :

$$\widetilde{I}(p) = \tfrac{1}{n(p)}\textstyle\sum_{q\in\wp} I(p)g(p-q,\sigma_g)f(p-q,\sigma_f) \quad / \quad \begin{cases} g(p-q,\sigma_g) & = & \exp\left(-\frac{(x_p-x_q)^2+(y_p-y_q)^2}{2\sigma_g^2}\right) \\ f(p-q,\sigma_f) & = & \exp\left(-\frac{(z_p-z_q)^2}{2\sigma_f^2}\right) \end{cases}$$

avec $n(p)$ facteur de normalisation spécifique à chaque pixel :

$$n(p) = \sum_{q\in\wp} g(p-q,\sigma_g)f(p-q,\sigma_f)f(p-q,\sigma_f)$$

La figure 4.9 (b) illustre sur un exemple concret la différence entre un filtrage gaussien simple et l'emploi d'un filtre bilatéral croisé. Notons que, contrairement au noyau gaussien simple, le filtre bilatéral n'est mathématiquement pas séparable. Cependant, s'il est appliqué sur une fenêtre de petite taille, éventuellement de façon récursive, implémenter une séparation donne malgré tout de bons résultats [PvV05, WOG06]. Les résultats que nous montrons ici utilisent la version complète du filtre bilatéral croisé, appliqué sur une fenêtre de taille 5 (par exemple figure 4.9 (a)).

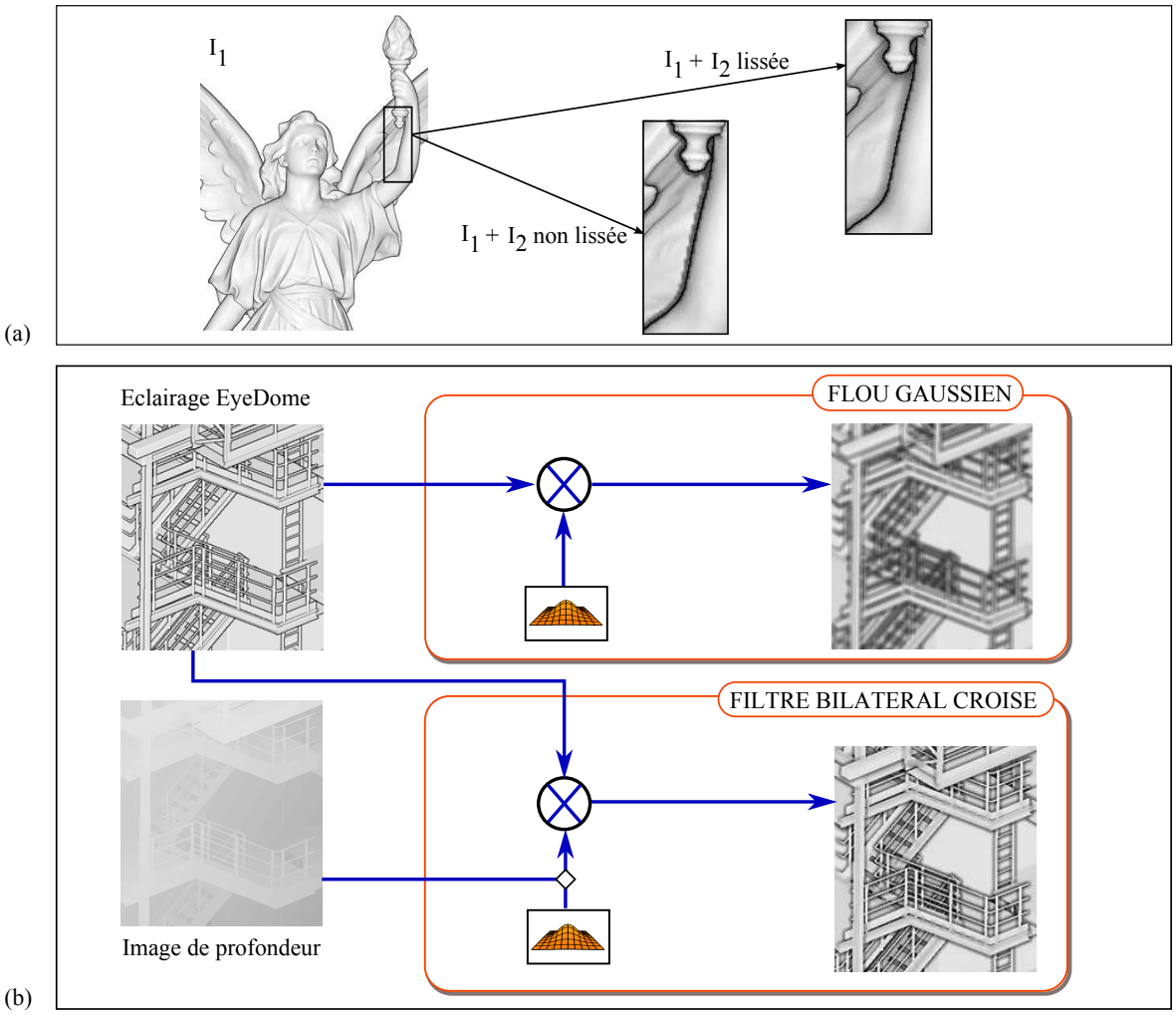


FIG. 4.9: (a) Pour éviter un crénelage résultant du sous-échantillonnage induit par les calculs en résolution inférieure, il est nécessaire d'effectuer un lissage de ces images. (b) L'utilisation d'un filtre gaussien simple (haut) induit des bavures au niveau des discontinuités de hauteur qui marquent la séparation entre les éléments de la scène. Un filtre bilatéral (bas), en faisant intervenir une non linéarité relative aux différences de profondeurs, permet de lisser l'image tout en évitant les bavures au niveau des silhouettes.

2.4 Résultats

2.4.1 Caractéristiques de l'EyeDome Lighting

Pour schématiser, les patterns d'ombrage générés par l'EDL comportent deux composantes dépendant de la nature locale de la géométrie du champ de hauteur obtenu lors de la projection.

Dans les régions de l'image où le champ de hauteur est continu, l'EDL produit un ombrage dont la valeur dépend d'une part de l'orientation de la surface projetée, d'autre part de la courbure locale, même si ces deux variables ne sont pas directement considérées dans les calculs. Ce résultat est illustré sur la figure 4.10. La dépendance de l'ombrage à l'orientation relative de la surface peut se comprendre simplement, en considérant une région localement plane (figure 4.10, images (a) à (c)). Pour un pixel p, le pseudo-dôme de lumière est toujours orienté dans la direction de vue, associée aux z croissants, et délimité par le demi-plan horizontal passant par p. La portion d'éclairage visible

en p sera donc d'autant plus réduite que la surface apparaîtra orthogonale au point de vue. Cette propriété peut être assimilée à celle qui définit un modèle d'ombrage local lambertien, comme on le retrouve dans le modèle de Phong par exemple. Mais en plus de dépendre de l'orientation de la surface relativement au point de vue, l'EDL permet de distinguer des régions de courbure différente (figure 4.10, images (d) et (e)). En effet, la portion du dôme d'éclairage visible depuis un pixel sera naturellement plus importante dans des régions convexes que concaves. Ainsi, pour une courbure identique en valeur absolue, un creux apparaîtra plus sombre qu'une bosse. Il faut noter que ces deux effets étant produits en considérant un voisinage bien défini dans l'image de profondeur, ils ne sont pas absolument liés aux orientations et courbures de la surface avant projection, mais dépendent fortement de la projection ainsi que du niveau de zoom.

La deuxième caractéristique importante de l'EDL se situe au niveau des discontinuités de profondeur. Dans les exemples que nous considérons, celles-ci peuvent correspondre à des silhouettes au sens géométrique du terme, c'est à dire des points de surface en lesquels la normale est orthogonale à l'angle de vue, mais pas nécessairement (bords de surfaces non closes). Pour un pixel se situant au voisinage d'une discontinuité, la quantité d'ombrage est saturée, le pixel est alors complètement noir. Dans l'image de pleine résolution (I_1), cela introduit des contours qui marquent clairement les limites des objets en superposition. Pour les calculs réalisés en résolution inférieure (I_2 et I_3), l'application d'un filtre bilatéral respectant la discontinuité de profondeur introduit des halos d'obscurité. Ceux-ci se retrouvent au niveau de l'objet arrière, c'est à dire partiellement occulté, ce qui introduit un détachement apparent dans l'image et facilite la perception de la profondeur relative dans ces configurations.

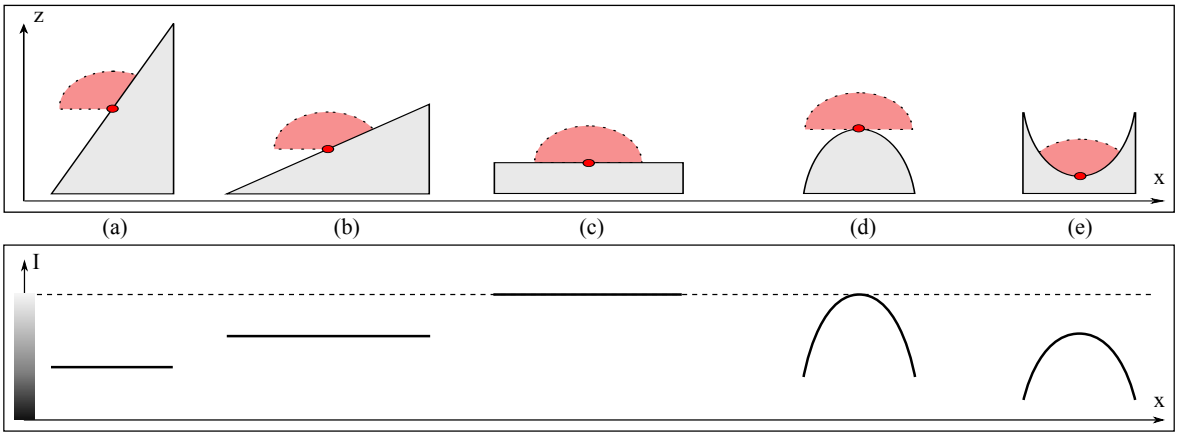


FIG. 4.10: Valeurs d'ombrage résultant de l'éclairage EDL pour différents cas 1D. La ligne du haut représente la configuration locale du champ de hauteur autour du pixel considéré (rouge), la portion de disque illustre la portion du pseudo-dôme de lumière visible depuis le pixel considéré. Le graphe du bas illustre schématiquement les profils d'ombrage calculés dans les différents cas. L'EDL garantit une intensité variant avec la pente de l'objet ((a), (b) et (c)), et apporte une information sur la courbure locale en distinguant les bosses (d) et les creux (e).

2.4.2 Comparaison de l'EDL avec des ombrages de type Phong et ambient occlusion

Afin d'illustrer les propriétés décrites ci-dessus, il est utile de comparer les résultats obtenus avec l'EDL à l'ombrage résultant d'autres modèles. Pour nous placer dans des conditions identiques, nous nous limiterons ici à des algorithmes opérant en espace image sur la seule base des informations de profondeur.

Nous avons d'abord implémenté un modèle de type Phong. Pour cela, il nous faut commencer par calculer la normale $\vec{N}$ pour le pixel considéré, à partir des informations de profondeur. On l'obtient par exemple en normalisant la somme des gradients horizontal et vertical du champ de hauteur. A partir de là, on considère une direction d'éclairage parallèle à la vue, c'est à dire $\vec{L} = (0,0,1)$, et on calcule l'ombrage comme le produit scalaire de la normale et de la direction d'éclairage : $I_{Phong} = \vec{L}.\vec{N}$. Il s'agit de la seule composante diffuse du modèle de Phong, la contribution spéculaire est ignorée car dans cette configuration particulière (colinéarité de la vue et de l'éclairage), les extrema d'intensité qu'elle générerait seraient identiques à ceux de la contribution diffuse.

Nous avons également mis en œuvre une version approchée de l'algorithme d'ambient occlusion en espace image (SSAO [Mit07]). L'idée de l'implémentation a été explicitée plus haut, dans la section 1.2.3 p.110, ainsi que dans les figures 4.3 et 4.4. Pour les images produites ici, nous utilisons un voisinage constitué de 32 points choisis de manière quasi aléatoire (générateur de Sobol) dans une boule de rayon $r = 0.05$ (toutes les coordonnées étant normalisées) centrée sur le pixel (vu comme un point 3D, avec sa hauteur associée). Pour éviter les effets de bande, ce voisinage est modifié d'un pixel à l'autre, en utilisant une réflexion par rapport à un axe choisi aléatoirement en chaque point. Cette méthode génère dans l'image un bruit qui est adouci au moyen d'un filtre bilatéral.

La figure 4.11 illustre les différences entre les différents algorithmes sur deux modèles assez différents, avec en plus une comparaison avec un éclairage classique en espace objet. L'*omotondo* (source : projet européen Aim@Shape) est une surface régulière présentant différentes courbures. Un éclairage de type Phong rend compte des orientations locales, mais des zones de creux (ou vallées) apparaîtront beaucoup trop claires. A contrario, le SSAO permet de faire ressortir ces courbures particulières, mais génère une image moins contrastée où les orientations de surface sont peu évidentes. L'EDL réunit ici les avantages de ces deux méthodes. Le deuxième modèle représente un des bâtiments de l'INRIA Rhône-Alpes. Au contraire des ombrages de type Phong, le SSAO et l'EDL permettent ici de mettre en valeur les discontinuités de profondeur dans l'image, faisant ressortir plus nettement les différents plans, notamment au moyen des halos engendrés par les résolutions inférieures de l'EDL.

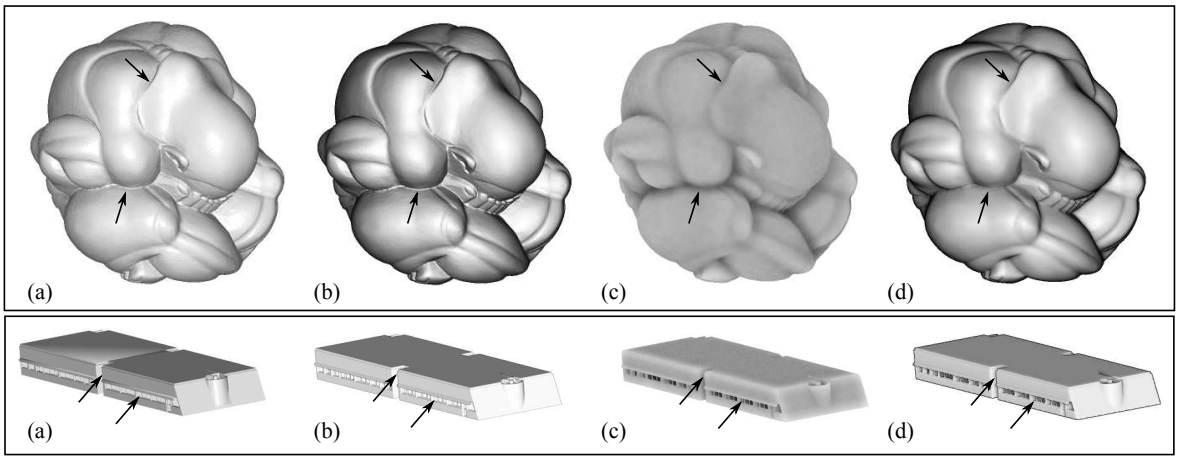


FIG. 4.11: Comparaison des ombrages générés par des méthodes de type Phong en espace objet (a) et image (b), SSAO (c) et EDL (d). Les flèches indiquent des zones de l'image où des différences importantes sont détectables. Pour l'omotondo (en haut), EDL et SSAO font mieux ressortir les creux et vallées locales qu'un ombrage de type Phong, mais le SSAO rend peu compte des orientations de surface. Dans le cadre d'une architecture présentant des angles saillants et différents plans (en bas, bâtiment de l'INRIA, ©Laurence Boissieux), l'EDL génère des halos permettant de mieux séparer les différentes profondeurs superposées au niveau des silhouettes.

2.4.3 Exemples particuliers

Jusqu'à présent, nous avons montré les résultats produits par l'EDL dans des configurations simples, c'est à dire lorsque l'objet à visualiser est une surface triangulée, pour laquelle les normales sont a priori disponibles. Intéressons-nous maintenant aux cas plus délicats d'ensembles discontinus.

Ces objets discontinus peuvent représenter le support même de la simulation numérique. C'est le cas notamment des simulations particulaires, que l'on retrouve parfois en astrophysique ou hydrodynamique. Des points particuliers peuvent également être introduits pour réaliser certains calculs (points de Gauss), et les résultats devront être affichés sur cet ensemble de points déconnecté de la géométrie initiale. Pour gérer ces ensembles de points, des méthodes de visualisation particulières doivent être mises au point. Ainsi, une version séparée du logiciel libre *Paraview*, baptisée *pv-meshless*, a été développée pour proposer un ensemble de routines de posttraitement et d'affichage spécifiquement adaptées aux simulations particulaires. Au sein d'EDF, un visualiseur a été implémenté spécifiquement pour afficher les résultats de calculs aux points de Gauss. Pour accélérer l'affichage, une technique désormais courante consiste à utiliser des sprites, qui sont en fait des textures 2D parallèles à l'écran, plaquées sur un rectangle et représentant par exemple l'image d'une petite boule. Ces sprites sont aujourd'hui efficacement implémentés dans les cartes graphiques, et permettent ainsi de visualiser interactivement des millions de points. Cependant, ce sont des imposteurs 2D, et l'ombrage d'une telle scène, bien que nécessaire pour bien appréhender l'organisation des données (voir l'étude de Gribble et Parker [GP06]), demeure problématique. La figure 4.12 montre un champ de température affiché sur un sous ensemble du volume au moyen de sprites. Un rendu brut est relativement confus, et ne permet que difficilement de se faire une idée de l'organisation spatiale des points en dépit des nombreux indices d'occultation présents dans l'image. Un éclairage Phong n'apporte aucune information, en raison de sa localité et de la structure plane des sprites. L'EDL, en revanche, facilite la séparation des éléments en avant et en arrière plan, et donne également des indications quant à la présence de plans de points (partie gauche de l'image), cela notamment grâce à l'approche multirésolution qui permet d'étendre dans l'image le voisinage de chaque point.

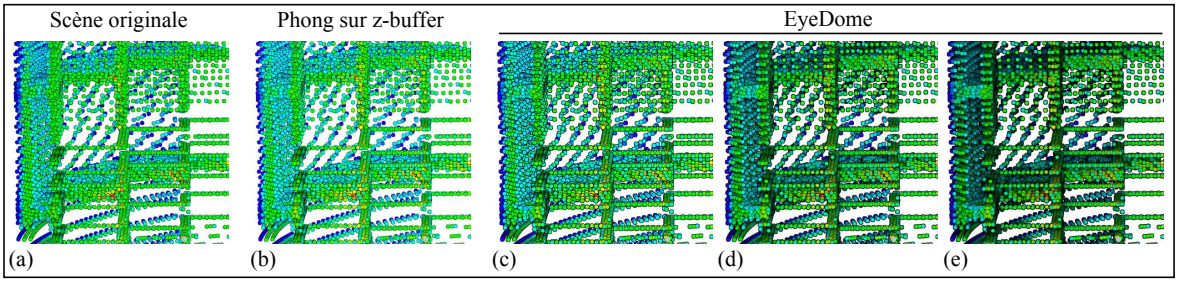


FIG. 4.12: Mise en évidence des relations de profondeur dans une scène de visualisation de points. Le champ scalaire de température est représenté sur un sous-ensemble de nœuds du maillage au moyen de sprites, qui sont des aplats orthogonaux à la direction de vue. Une représentation brute permet difficilement d'appréhender les relations spatiales entre sous-structures (a), et un éclairage de type Phong basé sur l'image de profondeur (b) n'apporte pas d'informations. L'algorithme d'EDL, en considérant différentes résolutions dans l'image, permet progressivement de faire ressortir les différentes structures ((c),(d) et (e), correspondant respectivement à I_1, $(I_1+I_2)/2$ et $(I_1+I_2+I_3)/3$).

Suite à une simulation en mécanique des fluides, il est parfois utile de visualiser la structure de l'écoulement. Pour cela, une méthode classique consiste à extraire certaines lignes de courant, qui sont définies par la tangence au champ de vitesse instantané. L'extraction de ces lignes constitue un problème en soi, ainsi que leur rendu. Il est possible, pour simplifier la visualisation, de construire des tubes triangulés autour de ces lignes. Cette technique simplifie l'utilisation d'un ombrage, mais

elle nécessite des calculs supplémentaires ainsi que plus d'espace mémoire. Des méthodes reposant sur des rendus hybrides (triangles et sprites) ont également été proposées [SGS05, MSE+06], mais elles ne sont pas nécessairement très simples à mettre en œuvre. La figure 4.13 montre qu'en utilisant des méthodes d'ombrage en espace image, il est possible de rendre simplement des lignes épaisses (*GL_LINES*) pour un résultat très satisfaisant. Les approximations réalisées par une telle approche peuvent générer quelques artefacts visuels, mais ceux-ci demeurent à notre avis tout à fait acceptables au regard de l'interactivité apportée par ces méthodes.

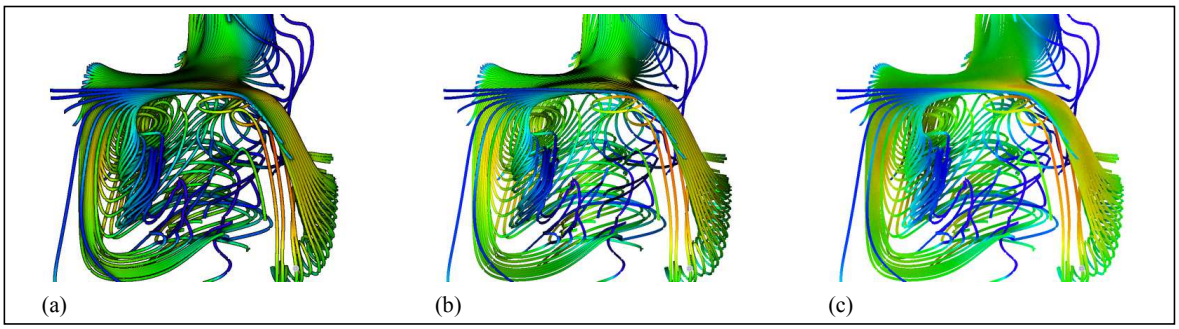


FIG. 4.13: Nécessitant seulement une information de profondeur, l'EDL peut être appliqué directement pour éclairer un ensemble de lignes de courant représentées seulement par des lignes épaisses (a). Il produit à la fois un éclairage dépendant des inclinaisons, comparable en cela à un éclairage de type Phong sur GPU (b), et d'assombrir les structures d'arrière plan pour lesquelles le dôme de lumière est partiellement occulté, comme c'est la cas pour le SSAO (c).

Pour terminer, nous évoquerons le cas particulier de la visualisation de très gros nuages de points. Le développement de scanners laser de plus en plus rapides et précis conduit à utiliser ces outils par exemple pour effectuer des relevés de mesures dans de grands ensembles industriels. Cela permet de disposer de modèles tel-que-construit de l'environnement, qui peuvent par la suite être consolidés (recalage de différentes mesures, élimination du bruit) et éventuellement maillés. Les normales au niveau des points mesurés ne sont pas toujours disponibles, ou peuvent être très bruitées, ce qui nécessite alors des prétraitements éventuellement très lourds pour reconstituer cette information. Pour des nuages de points très denses, comme dans la figure 4.14 (ici plus de 29 millions de points), les techniques d'ombrage reposant sur l'image de profondeur peuvent ainsi s'avérer très utiles pour explorer interactivement les données brutes. On voit sur cette image que l'EDL apporte de nombreuses informations utiles à la compréhension de la structure spatiale complexe et incomplètement représentée de cet intérieur de centrale.

2.4.4 Performances : avantages et inconvénients

Le coût de l'EDL en terme de temps de calcul est délicat à quantifier, tant il est lié aux caractéristiques techniques de la carte graphique utilisée. Pour commencer, précisons que les algorithmes d'ombrage différé comme l'EDL connaissent un développement important depuis l'apparition des architectures graphiques unifiées. Avant celles-ci, une carte graphique était constituée d'unités de calculs dont une portion fixe était réservée aux calculs sur les sommets (*vertex shaders*), l'autre partie étant spécialisée dans les calculs sur les pixels (*pixel shaders*). Les nouvelles générations de cartes graphiques commercialisées après 2007 (apparition de la G80 chez NVIDIA) ont fait sauter cette séparation, une carte étant dorénavant constituée de processeurs de flux alloués dynamiquement aux calculs sur les sommets ou les pixels en fonction des besoins. Pour un rendu en ombrage différé, une fois la géométrie projetée, toutes ces unités de calculs peuvent être allouées sans restrictions au calcul de l'éclairage par pixel, ce qui accélère le traitement. Par ailleurs, il convient de noter que

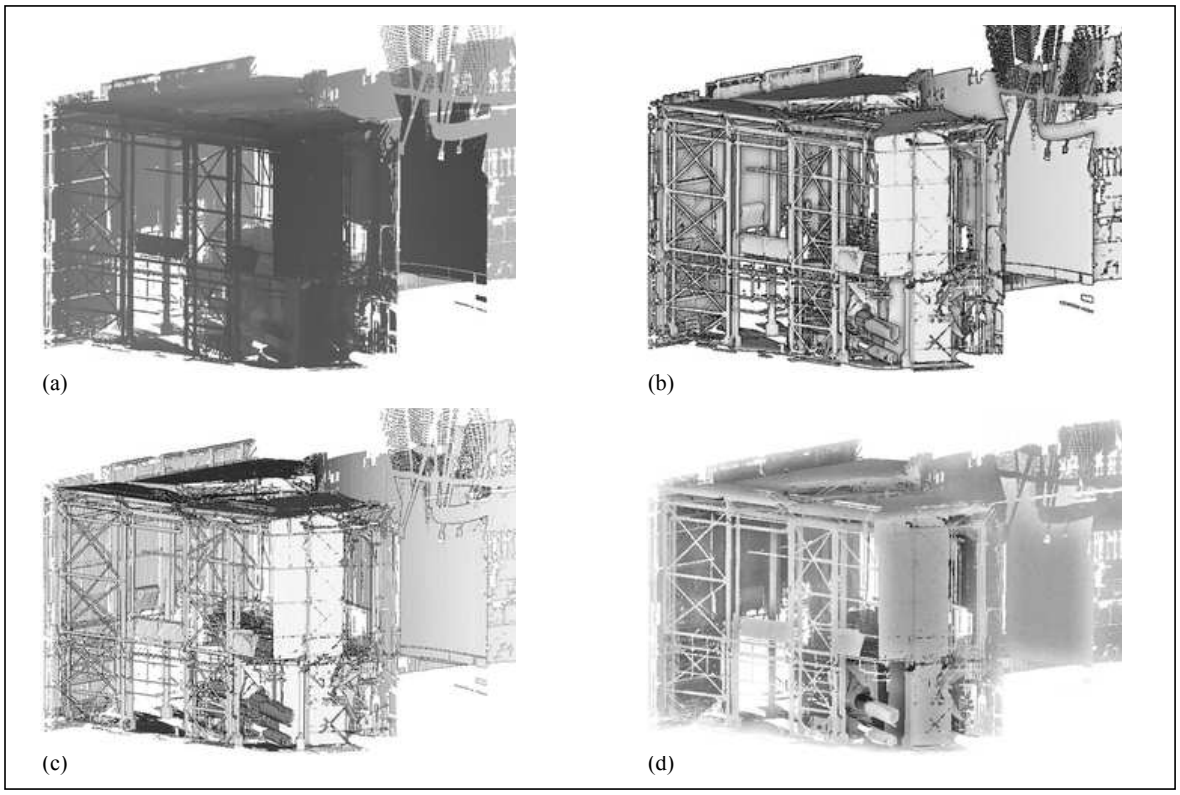


FIG. 4.14: L'EDL peut également être utilisé pour visualiser un nuage de points dense (intérieur de centrale, 29 millions de points) pour lequel on ne dispose pas nécessairement des informations de normales. Sur la base de l'image de profondeur (a), l'EDL (b) facilite la détection de discontinuités dans le nuage, ce qui est plus délicat avec un Phong sur GPU (c), tout en conservant les hautes fréquences de surface, contrairement au SSAO (d). Par contre, pour des points isolés (en haut à droite de l'image), l'EDL génère des silhouettes qui peuvent obscurcir un peu trop l'image.

l'évolution actuelle des cartes graphiques tend vers une multiplication du nombre de processeurs de flux (plus de 300 sur les dernières cartes), les constructeurs cherchant également à promouvoir l'utilisation des cartes en tant que puissants calculateurs parallèles (API *CUDA* de NVIDIA par exemple).

Pour un algorithme en espace image comme l'EDL, le temps de calcul est principalement fonction du nombre d'opérations par pixel ainsi que de la taille de la fenêtre de rendu (nombre de pixels). Il ne dépend en revanche pas de la taille de la géométrie projetée. Les calculs que nous mentionnons ici ont été réalisés sur un ordinateur portable équipé d'une carte NVIDIA GeForce8600M (32 processeurs). Pour indication, un test réalisé avec une première version de l'algorithme donnait des temps de calcul en moyenne 4 à 5 fois plus réduits sur une GeForce8800 GTX (128 processeurs) que sur la GeForce8600M. Le tableau 4.1 résume les performances obtenues pour l'EDL et notre implémentation du SSAO de [Mit07], en nombre d'images par seconde (moyenne sur 500 images). Le calcul de l'EDL sur une image de taille 800x800 prend environ 27 ms, avec respectivement 5, 2.2 et 1.9 ms pour le calcul des résolutions I_1, I_2 et I_3, 16 ms pour le filtrage bilatéral (réalisé sur une fenêtre de taille 5) et 2.1 ms pour la composition finale des ombrages aux différentes résolutions. Logiquement, l'opération la plus pénalisante est le lissage des moyenne et basse résolutions (I_2 et I_3), le filtre bilatéral croisé souffrant de sa non-séparabilité. Si ce point nuit effectivement à l'interactivité du rendu, plusieurs solutions peuvent être adoptées. La première consiste à utiliser l'approximation séparée du filtre proposée par [PvV05]. Egalement, il est possible de n'appliquer le filtre complet que

Taille de la fenêtre d'affichage	Algorithme d'ombrage	Lissage de l'image	Vitesse (FPS)
800 x 800	EDL	non	89
		gaussien	57
		bilatéral	38
	SSAO (32 voisins)	non	31
	SSAO (32 voisins)	bilatéral	15
	SSAO (16 voisins)	non	59
	SSAO (16 voisins)	bilatéral	20
	Phong GPU	-	217
1024 x 800	EDL	non	75
		gaussien	50
		bilatéral	30
	SSAO (32 voisins)	non	11
	SSAO (32 voisins)	bilatéral	8
	SSAO (16 voisins)	non	21
	SSAO (16 voisins)	bilatéral	12

TAB. 4.1: Comparaison des vitesses d'exécution de l'EDL, du SSAO et de Phong sur GPU, en fonction de la taille d'image et de certains paramètres des algorithmes. Les résultats sont exprimés en images par secondes (FPS), moyennés sur 500 images. Les calculs ont été réalisés sur une carte NVIDIA GeForce 8600M (32 processeurs de flux).

lorsque la scène est immobile, et de le remplacer sinon par le filtre gaussien séparé, la différence étant vraisemblablement peu perceptible lors du mouvement d'objets à l'écran (c'est le principe utilisé pour les LOD dépendant de la vitesse par exemple).

3 Extension de l'EyeDome Lighting

3.1 EDL et lumière directionnelle

Le principe de l'algorithme présenté précédemment repose sur la métaphore d'un pseudo-dôme de lumière se situant pour chaque pixel dans la direction de vue. Cependant, modifier la direction d'éclairage lors de l'analyse d'un objet 3D peut s'avérer particulièrement utile pour mieux en comprendre la structure spatiale. Dans cette section, nous présentons l'intérêt perceptif que représente le contrôle de la direction d'éclairage, puis nous proposons une modification de l'algorithme permettant de modifier la direction apparente de l'EyeDome.

3.1.1 Justifications perceptives : hypothèse de l'éclairage par le haut et intérêt d'un éclairage dynamique pour lever les ambiguïtés

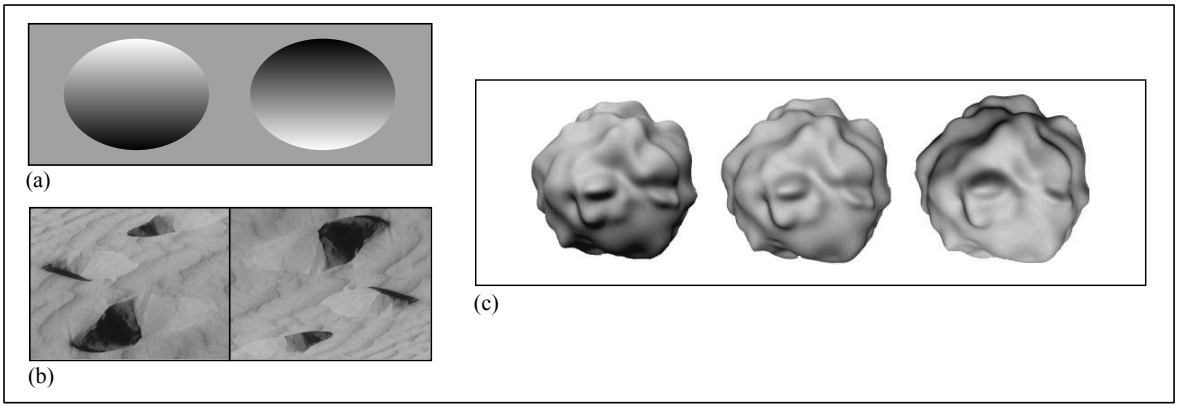


FIG. 4.15: (a) Illusion du cratère : en retournant une image, le même gradient de luminance peut être interprété préférentiellement comme une figure convexe à gauche et concave à droite, ce qui suppose de faire l'hypothèse d'une source de lumière en haut de l'image. (b) En utilisant ces figures pour lesquelles la direction de l'éclairage varie le long d'une verticale, O'Shea et al. [OBA08] observent une meilleure perception de la surface chez des sujets humains quand l'éclairage est positionné légèrement au dessus de l'axe de vue.

L'exploration de surfaces complexes peut nécessiter de modifier de façon interactive la direction d'éclairage, afin de révéler certains détails peu exposés par une lumière trop lointaine. Récemment, Caniard et Fleming [CF07] ont mené une étude dans laquelle ils ont mis en évidence la difficulté qu'il y a à reconstruire mentalement et apparier des surfaces lisses lorsque la direction d'un simple éclairage de type Phong varie. Mais au delà de problèmes liés à un manque de contraste dans des zones mal éclairées, il faut également tenir compte de la façon dont le système visuel humain peut être biaisé dans sa perception de la forme à partir de l'indice d'éclairage.

En 1786, Rittenhouse [Rit86] explique un phénomène observé quelques années plus tôt : les parties concaves d'un sceau observé à travers un microscope étaient perçues convexes, et réciproquement. En fait, Rittenhouse remarque que l'image de l'objet étant inversée dans l'appareil optique, les ombres se trouvent mal positionnées par rapport à la source de lumière qui est supposée la même que dans le monde environnant, ce qui génère une inversion de la perception du relief.

Plus précisément, le système visuel semble s'appuyer sur l'hypothèse que dans le monde, et aussi dans les images, la lumière vient d'en haut (hypothèse souvent désignée dans la littérature par le terme de *light from above*). Ce phénomène a été mis en évidence par de nombreuses études reposant sur des objets et des tâches expérimentales variées. En utilisant des ellipses avec un gradient de luminance (Fig.4.15(a)) Ramachandran [Ram88] a observé que l'on pouvait extraire une surface avec une simple information d'ombrage, et qu'il existait en outre une préférence forte pour un éclairage venant d'en haut. Avec les mêmes stimuli, Sun et Perona [SP98] ont mené une tâche de recherche visuelle en faisant varier la direction du gradient ; leurs résultats montrent qu'une cible dont le gradient est inversé par rapport aux distracteurs environnants est plus efficacement retrouvée lorsque les distracteurs semblent être éclairés par le haut. Avec des objets un peu plus complexes mais toujours une méthodologie de recherche visuelle, [OCS05] ont montré d'une part que la direction d'éclairage n'est pas extraite globalement, c'est à dire qu'il n'y a pas de phénomène clair de pop-out pour une cible éclairée différemment, mais qu'il existe cependant une asymétrie dans le traitement de la direction principale d'éclairage, avec une préférence pour une direction proche de la verticale. Mamassian et Goutcher [MG01] ont retrouvé un résultat similaire avec un dessin mêlant des traits noirs et blancs

représentant une surface bosselée, les sujets devant décrire la forme de l'objet dont l'inclinaison variait d'un essai à l'autre. Plus récemment, O'Shea et al. [OBA08] ont employé un stimulus plus proche de la réalité quotidienne en informatique graphique et visualisation scientifique, à savoir une sphère déformée avec un éclairage de type Phong dont la direction était régulièrement décalée le long de la verticale (Fig.4.15(b)) ; les sujets à qui il était demandé de déterminer l'orientation de la surface en différents points de l'objet au moyen d'une petite jauge (système inspiré de [KvDK92]) ont été plus précis lorsque la direction d'éclairage était positionnée légèrement au dessus du point de vue. Cette dernière étude est particulièrement intéressante en ce qui concerne une application en visualisation, car elle apporte une définition plus pratique (en terme de slant et tilt) de la notion de lumière « venant d'en haut », l'espace physique du champ lumineux supposé étant laissé plutôt abstrait dans les premières études. Notons également que plusieurs études parmi lesquelles [SP98, MG01, GdMM07], et dans une moindre mesure [MBW04], ont également mis en évidence un biais latéral, à savoir une préférence pour une lumière venant d'en haut à gauche.

Cette préférence du système visuel pour une direction d'éclairage particulière est exploitée depuis longtemps dans la peinture, et a servi de base à des outils visant à améliorer l'éclairage de scènes 3D (par exemple LightKit pour VTK ([HM03]). Dans le cas de l'algorithme non photoréaliste EDL, il peut être intéressant d'introduire une modification des calculs permettant de simuler un éclairage directionnel, même si le principe initial correspond à une lumière frontale.

3.1.2 Lumière directionnelle : le TorchDome Lighting (TDL)

Pour modifier l'orientation du dôme local d'éclairage, nous proposons une solution qui repose sur une approximation géométrique simple. Celle-ci donne de bons résultats pour des angles raisonnablement proches du point de vue.

Considérons une direction d'éclairage quelconque $\vec{L}$. Pour un pixel p donné, définissons le plan (P_p) passant par p et de normale $-\vec{L}$. Avec $p = (x_p, y_p, z_p)$ et $\vec{L} = (x_L, y_L, z_L)$, ce plan est simplement décrit par son équation cartésienne :

$$(\mathrm{P_p}) = \{M(X,Y,Z)/ - x_L X - y_L Y - z_L Z + x_L x_p + y_L y_p + z_L z_p = 0\}.$$

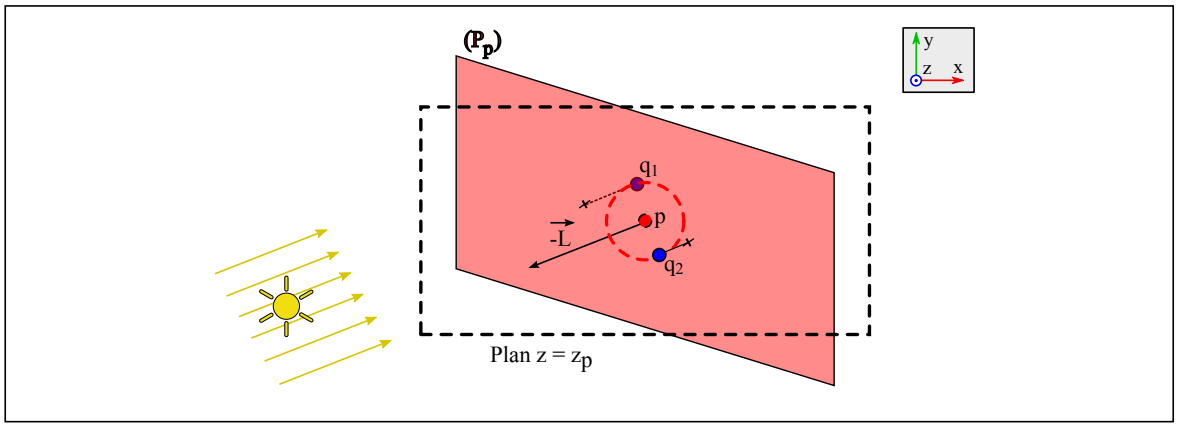
Un calcul exact de la valeur d'ombrage suivant l'algorithme de l'EDL mais dans la direction de $\vec{L}$ est complexe et pas nécessairement réalisable. Dans un premier temps il faut définir les points appartenant au voisinage de p sur le plan (P_p). Puis, pour chacun, il convient de rechercher le point de la scène correspondant, ce qui doit être fait en lançant un rayon orthogonal à (P_p) et passant par ce point. Cette recherche, qui doit être réalisée dans l'image de profondeur, est coûteuse, puisqu'il s'agit de parcourir un rayon dans l'image jusqu'à ce que celui-ci intercepte la surface projetée ou sorte du cadre.

La solution que nous proposons est une approximation de la solution correcte qui est aussi rapide que le calcul basique d'EDL et qui donne des résultats visuellement convaincants au moins pour des angles d'éclairage pas trop rasants (jusqu'à environ 50 ° d'écart par rapport à la verticale définie par le point de vue). Le principe en est illustré par la figure 4.16 (a). Le voisinage $\mathcal{V}_p$ utilisé pour le calcul de la fonction d'ombrage est défini dans l'image exactement de la même manière que pour le cas de lumière frontale (section 2.2.1). Pour un point $q \in \mathcal{V}_p$, on calcule la distance signée de q à (P_p), à savoir

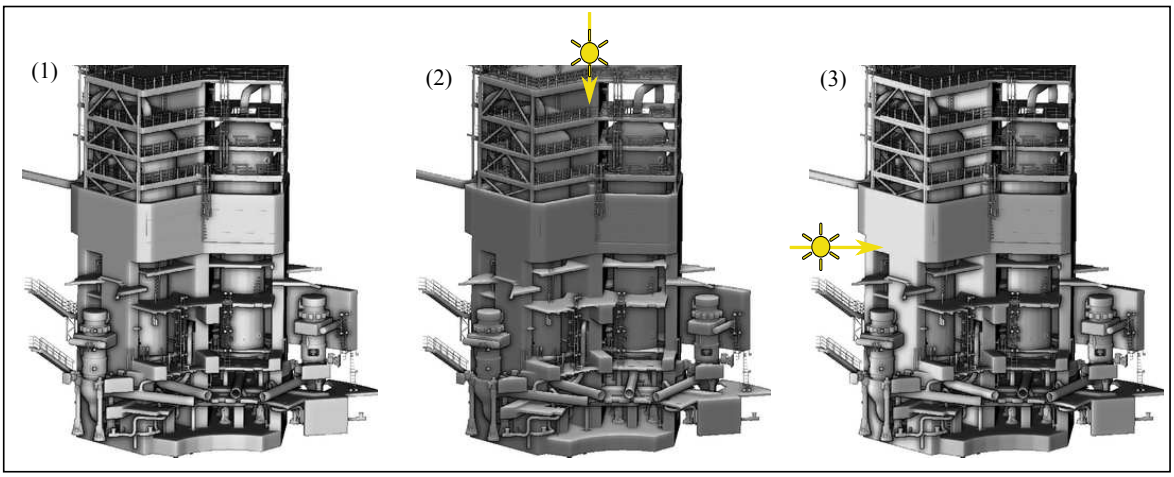
$$\delta(q, \mathrm{P_p}) = -x_L x_q - y_L y_q - z_L z_q + x_L x_p + y_L y_p + z_L z_p\}.$$

Cette distance au plan (P_p) est alors utilisée pour calculer l'éclairage au point p, en remplacement de la simple différence de hauteur $(z_p - z_q)$, la distance image entre p et q étant définie à l'identique :

$$s^+(p,q) = \left[\frac{\delta(q, \mathrm{P_p})}{d_{p,q}}\right]^+$$



(a)



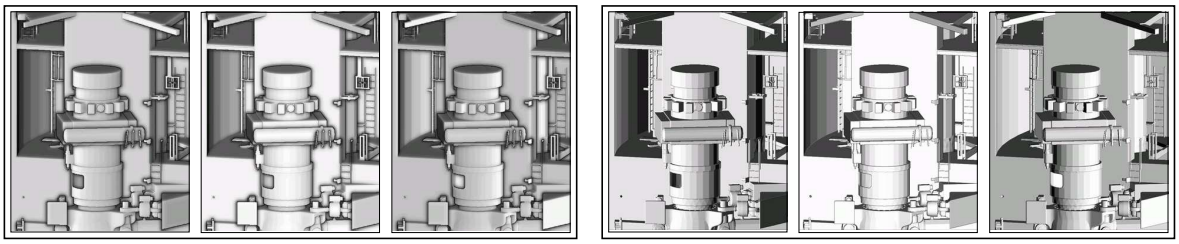
(b)

FIG. 4.16: (a) Principe de calcul de l'ombrage pour une lumière de direction $\vec{L}$. (b) EDL simple (1) et EDL directionnel (TDL) avec une lumière positionnée vers la gauche (2) et vers le haut (3). Dans cet exemple, le TDL permet d'augmenter le contraste des surfaces orientées vers la gauche ou le haut, et ce sans bouger la géométrie.

.

3.1.3 Résultats

Les résultats produits par le TDL sont illustrés sur la figure 4.16(b). Dans cet exemple de visualisation d'un complexe industriel, modifier la direction d'éclairage permet de mettre en valeur soit les planchers du bâtiments, soit les murs présentant une certaine orientation, sans perdre pour autant l'orientation de la scène. De plus, cela peut servir à améliorer le contraste dans des zones trop sombres, qui ne pourraient être observées frontalement en raison des occultations engendrées par la complexité de la structure. La figure 4.17 compare les résultats produits avec le TDL et avec l'ombrage de type Phong sur GPU lorsque l'on change la direction d'éclairage. Une des caractéristiques essentielles de l'EDL, à savoir les halos permettant de séparer rapidement des objets en occultation partielle dans la direction de vue, est clairement préservée avec le TDL.



(a) (b)

FIG. 4.17: Eclairage directionnel sur une scène représentant l'intérieur d'une centrale, avec le TDL (a) et du Phong sur GPU (b). L'EDL préserve les halos dans la direction de vue, quelle que soit la direction d'éclairage choisie par l'utilisateur. Il est ainsi possible d'améliorer le contraste dans certaines zones de l'image tout en conservant la facilitation visuelle de séparation en profondeur produite par l'EDL.

3.1.4 Remarques et critiques

Comme pour l'EDL frontal, tous les calculs sont réalisés entièrement en espace image, en négligeant toute information de géométrie différentielle comme la normale à la surface ou le tenseur de courbure. En raison de la projection explicitée par le modèle de caméra, la direction de la lumière que nous avons introduite n'a pas de signification physique directe. Ainsi, une élévation de 20 ° par rapport au point de vue avec un EDL directionnel ne donnera pas la même impression qu'une même élévation définie pour un éclairage Phong dans l'espace objet. Par conséquent, on ne peut ici prétendre générer une perception exacte de l'orientation absolue des surfaces par rapport à l'observateur, mais simplement faciliter la compréhension des variations de la surface. De plus, l'approche que nous proposons fonctionne dans la limite d'angles pas trop importants (notamment il est délicat de s'écarter de la direction de vue selon les deux axes).

Une propriété intéressante de l'implémentation proposée du TDL consiste en la conservation des silhouettes et halos générés par les différences de profondeur dans la scène (voir Fig.4.17). La stabilité de ces marqueurs sombres pourrait être perçue en contradiction avec le mouvement apparent de l'éclairage lorsque l'on change sa direction, induisant une certaine gêne pour l'utilisateur, mais il semble bien que ce ne soit pas le cas à l'usage.

3.2 Focus stylisé

Lors du rendu de scènes complexes, la surabondance de détails dans l'image risque de perturber la visualisation, en attirant le regard de l'utilisateur dans différentes régions à la fois. Afin de faciliter l'exploration des données, il peut s'avérer intéressant de définir une région dans l'image en dehors de laquelle la saillance induite par l'ombrage est dégradée. Cette idée s'inspire de travaux en rendu non photoréaliste visant à guider l'attention de l'observateur tout en apportant une représentation simplifiée, esquissée, en dehors des zones jugées importantes. On parle alors de *focus stylisé*. Kosara et al. [KMH+02] proposent par exemple d'utiliser un effet de profondeur de champ pour flouter les zones non sélectionnées, de manière à attirer le regard vers les régions sémantiquement saillantes. DeCarlo et Santella [DS02] présentent un système d'abstraction de photographies préservant plus de détails dans les régions où le regard de l'utilisateur se fixe lors d'une phase exploratoire. Cole et al. [CDF+06] étendent cette approche à des scènes 3D, leur focus stylisé permettant de varier la densité et la qualité des traits, la saturation des couleurs et le contraste dans l'image. Ils mènent par ailleurs une expérience avec un capteur de regard, qui démontre que leur technique attire effectivement plus l'attention ouverte des sujets vers les zones de focus qu'un rendu complet.

Nous proposons ici une extension très simple de l'EDL (ou du TDL) visant à produire un effet similaire. L'objectif est de réduire le contraste généré par l'ombrage en dehors de certaines régions,

afin que l'utilisateur puisse se concentrer sur une sous-partie du volume, voire attirer simplement l'attention de ses interlocuteurs sur une région particulièrement intéressante. Pour cela, nous définissons deux focus, le premier en profondeur ($Z_f \in [0:1]$, le second dans l'espace image ($(X_f, Y_f) \in [0:1]^2$). D'un point de vue pratique, ces valeurs peuvent être déterminées manuellement, en fonction d'objet particulier, ou en utilisant le pointeur de la souris. Nous définissons alors deux fonctions d'importance centrées sur ces focus, $Imp^z(Z_f, z)$ et $Imp^{img}(X_f, Y_f, x, y)$, qui modulent l'effet de l'ombrage en un pixel p de profondeur associée z et de coordonnées image (x, y) :

$$I(p) = \exp\left(-A\mathcal{S}(p) Imp^z(Z_f, z) Imp^{img}(X_f, Y_f, x, y)\right)$$

Les fonctions d'importance peuvent être définies de diverses manières, la seule contrainte étant la présence d'un maximum global en Z_f et en (X_f, Y_f). Dans notre implémentation, nous avons utilisé des gaussiennes centrées en ces valeurs et d'écart-types respectifs σ_z et σ_{img} (voir figure 4.18(a)) :

$$\begin{aligned} Imp^z(Z_f, z) &= \alpha_z + (1-\alpha_z)\exp\left(-\frac{(Z_f - z)^2}{2\sigma_z^2}\right) \\ Imp^{img}(X_f, Y_f, x, y) &= \alpha_{img} + (1-\alpha_{img})\exp\left(-\frac{(X_f - x)^2 + (Y_f - y)^2}{2\sigma_{img}^2}\right) \end{aligned}$$

Les paramètres α_z et α_{img} définissent implicitement le contraste minimal loin du point de focus. Une valeur nulle entraîne une extinction de l'effet d'ombrage à distance (pixels blancs), alors qu'une valeur de 1 annule l'effet de focus.

Les résultats produits par ces équations sont illustrés sur deux modèles (figure 4.18, (b) et (c)). L'exemple de la statue de Bouddha montre l'effet produit par la variation de σ_{img}, la région de focus étant plus ou moins resserrée. Les images de complexe industriel mettent en valeur l'utilisation séparée et conjointe des deux focus.

La modulation de l'ombrage que nous proposons ici offre un moyen simple et efficace pour réduire le contraste dans certaines régions de l'image. En étant intégrée directement dans le calcul d'ombrage, son coût est relativement faible. La réduction de contraste produite est implicite, dans la mesure où elle correspond en fait à une extinction progressive de l'ombrage à distance du focus. Si l'image d'ombrage est par la suite superposée à une carte de couleur (voir figure 4.7), les pixels éloignés du focus conservent a priori leur couleur d'origine. Pour obtenir un effet plus complet, il serait possible de conserver les valeurs d'importance, par exemple dans les composantes *GB* de la texture utilisée pour enregistrer l'ombrage, afin de les appliquer également à une modulation de la carte de couleur.

4 Conclusion et perspectives

L'EyeDome Lighting constitue un nouvel algorithme d'ombrage destiné à faciliter la perception des formes et des relations spatiales dans des scènes tridimensionnelles arbitrairement complexes. Il a été élaboré en réponse aux attentes énoncées dans la section 1.4, à savoir :

- *Guidage perceptif :* s'il semble délicat d'établir une vision globale de notre perception par l'ombrage, un certain nombre de principes simples peuvent être dégagés. Parmi ceux-ci, l'EDL prend en compte, au moins localement en espace image, le lien établi par le SVH entre zone sombre et profondeur apparente (*dark means deep* [LB00]). De plus, les halos, produits par les traitements en résolutions inférieures, facilitent la séparation en profondeur des différents objets ainsi que la perception des contours occultants, dont l'importance pour l'extraction de la forme à partir de l'ombrage a été montrée expérimentalement [TR89, RT90].

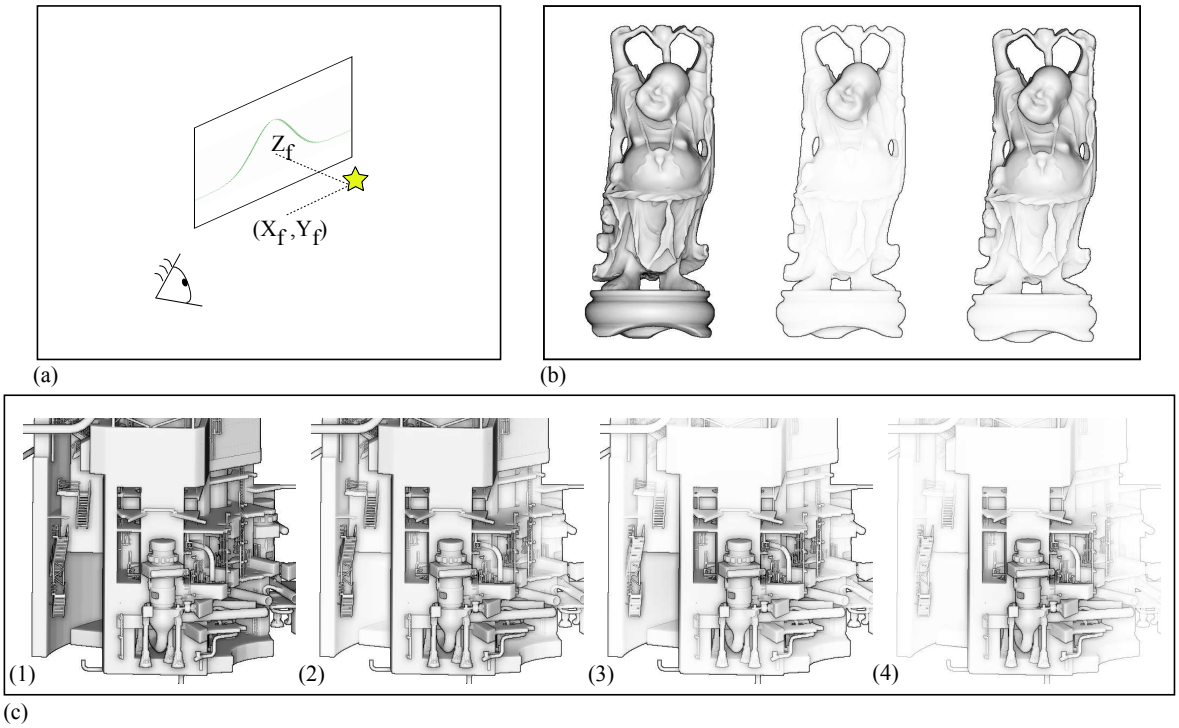


FIG. 4.18: (a) Le focus relatif à l'objet étoilé est défini par une gaussienne relative à la profondeur et une gaussienne en espace image qui modifient la valeur de l'ombrage. Plus un pixel se trouve éloigné du centre de focus, plus son ombrage est réduit, il apparaîtra donc d'autant plus clair. (b) Exemple de focus image centré sur la tête de la statue de bouddha (de gauche à droite : pas de focus, $\sigma_{img} = 0.05$, $\sigma_{img} = 0.2$). (c) Le focus est défini en pointant la souris sur la structure cylindrique au premier plan. Une modulation de l'ombrage obtenu par l'EDL (1) est alors définie en profondeur seulement (2), en espace image seulement (3) ou à la fois en profondeur et dans l'image (4).

- *Généralité :* l'EDL est calculé à partir des seules informations de profondeur enregistrées au moment de la projection de la scène. Comme nous l'avons illustré sur plusieurs exemples, il peut ainsi être appliqué à des scènes comportant des objets aussi variés que des surfaces triangulées, des lignes épaisses, des point sprites.
- *Interactivité :* le nombre restreint d'opérations réalisées sur le GPU et l'approche multi-résolution permettent un rendu très rapide, tout du moins avec des cartes graphiques récentes bénéficiant d'une architecture unifiée. De plus, l'algorithme ne requiert aucun prétraitement des données à éclairer, le goulot d'étranglement principal se situe ainsi au niveau de la taille de l'affichage (nombre de pixels).
- *Directionnalité :* modifier facilement la direction de l'éclairage permet d'augmenter le contraste dans des régions autrement sombres, et ce en évitant de bouger la scène entière, ce qui peut être lourd voire impossible si certains objets viennent alors cacher la région à observer. Nous avons proposé une extension à l'EDL, appelée TorchDome Lighting, qui répond à ce besoin. N'utilisant toujours que l'image de profondeur, le rééclairage peut être réalisé dynamiquement, sans besoin de redessiner toute la scène.

Pour conclure ce chapitre, nous allons maintenant essayer de replacer l'EDL dans un contexte un peu plus général de visualisation et de perception.

4.1 Limites de l'EDL

Nous avons jusqu'à présent insisté sur les avantages principaux de notre algorithme, essayons maintenant d'en définir les limites, tant liées à l'implémentation particulière proposée qu'aux idées sur lesquelles elle repose.

En premier lieu, remarquons qu'en dépit de l'utilisation de plusieurs résolutions d'image, l'EDL demeure très localisé, sa portée maximale étant de l'ordre de la dizaine de pixels. En conséquence, une concavité étendue ne pourra pas être obscurcie, ce qui peut générer une illusion perceptive. Pour pallier cette difficulté, il est possible de calculer également une approximation image d'ambient occlusion (SSAO [Mit07]) dont le rayon d'action plus important permettra de rendre compte des basses fréquences spatiales des variations de surface. Les deux ombrages peuvent alors être multipliés pour produire une image plus lisible. Il faut cependant tenir compte du temps de calcul supplémentaire introduit ici.

Une difficulté inhérente à l'EDL tient au réglage du facteur d'amplification de l'ombrage. En effet, l'EDL n'ayant pas de support physique, il faut déterminer ce facteur afin d'obtenir une luminance moyenne et un contraste acceptables. Pratiquement, il nous est apparu que la valeur par défaut convient à beaucoup de situations, mais des problèmes peuvent survenir notamment lorsque l'on modifie le zoom de la scène (voir figure 4.19). En effet, le voisinage utilisé pour le calcul de la valeur d'ombrage est défini en espace image. Si un objet est vu de très loin, les voisins d'un pixel sont susceptibles de représenter des profondeurs très différentes, ce qui conduira à un assombrissement de l'image. Au contraire, si l'on regarde ce même objet de près, les variations de profondeurs observées dans l'image seront beaucoup plus restreintes, d'où un ombrage relativement faible et une image trop claire. Dans la version actuelle de l'algorithme, il est possible de régler manuellement le facteur d'amplification de l'EDL, afin d'obtenir un niveau moyen de luminosité acceptable. Cependant, l'exploration de données nouvelles en visualisation scientifique s'accompagne souvent de variations continues du niveau de zoom, l'observateur pouvant se concentrer alternativement sur des phénomènes étendus dans l'espace ou plutôt localisés. Pour faciliter l'utilisation de l'EDL dans ce cadre, il serait donc particulièrement utile d'intégrer un contrôle automatique de l'image produite. Le niveau d'amplification de l'EDL pourrait par exemple être déterminé en fonction du contraste dans l'image résultante, calculé rapidement par une méthode pyramidale implémentée sur le GPU, ou varier continûment en fonction du niveau de zoom. Plus généralement, de nombreux paramètres de l'EDL peuvent être réglés par l'utilisateur, rendant a priori plus complexe sa mise en œuvre (taille du voisinage, pondération des 3 résolutions...). Cependant, des configurations génériques peuvent être proposées, et dans la pratique les paramètres utilisés pour produire la plupart des images de ce chapitre donnent un résultat convaincant.

L'interactivité de l'EDL est garantie par son implémentation sur GPU et par l'évolution actuelle des cartes graphiques. Cependant, les temps de calcul sont fortement liés à la taille de l'affichage, et une application sur un écran géant (mur d'images de plusieurs millions de pixels) peut s'avérer problématique. Notons cependant que dans de tels cas, il est d'usage d'utiliser un cluster graphique intégrant plusieurs cartes gérant en parallèle différentes sous-parties de l'image globale. La difficulté du portage de l'EDL se situera alors au niveau des pixels de bord de chacune de ces sous-parties, puisqu'il est nécessaire de disposer d'informations de voisinage pouvant se trouver gérées par une autre carte.

Enfin, si l'EDL peut être appliqué sur des entités géométriques diverses, il est des cas où son utilisation se révèle non concluante, voire impossible. Par exemple, Luft et al. [LCD06] utilisent leur technique d'*unsharp masking* de l'image de profondeur pour ajouter des halos à des photographies pour lesquelles les disparités entre deux prises de vues, et donc implicitement les profondeurs relatives, ont été engistrées. Mais ces cartes s'avèrent très discrétisées, avec des plateaux de profondeur

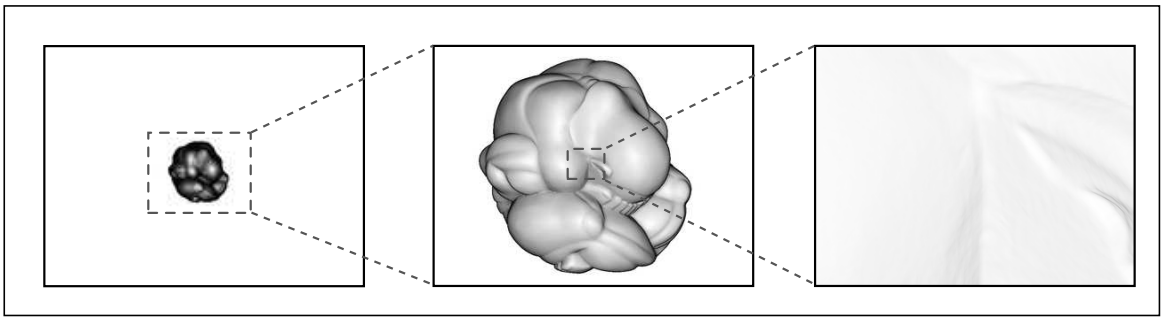


FIG. 4.19: Le voisinage utilisé pour calculer l'ombrage étant défini purement en espace image, le niveau de zoom modifie directement le contraste dans l'image. En effet, un objet vu de loin présentera des variations de hauteur bien plus importantes sur quelques pixels que s'il est vu de très près. D'après la formulation de l'EDL, il en résultera une image plus sombre dans le premier cas. L'amplification de l'ombrage peut être réglée manuellement pour obtenir le contraste désiré, mais une prise en compte automatique du niveau de zoom faciliterait l'utilisation de la technique.

constante, et l'EDL appliqué à ces images ne nous a pas fourni de résultats convaincants. Par ailleurs, l'EDL requiert l'existence d'une profondeur unique par pixel, ce qui équivaut à une représentation opaque. Dès lors, son application à l'ombrage de surfaces transparentes superposées, voire à des rendus de type volumiques ne paraît pas envisageable, du moins sous sa forme actuelle.

4.1.1 Comparaisons avec d'autres méthodes d'ombrage

Précédemment, nous avons mis en relation les résultats produits par l'EyeDome Lighting avec deux autres techniques d'ombrage basées sur les seules informations de profondeur dans l'image, à savoir une extension GPU de l'éclairage de Phong pour lequel les normales sont recalculées dans l'image, et l'ambient occlusion en espace écran. Nous étendons ici cette comparaison à d'autres ombrages mentionnés dans la première partie de ce chapitre. Le but n'est pas de tirer des conclusions définitives, car un tel exercice est subjectif tant il dépend des objets rendus et de la tâche à réaliser, mais simplement de mettre en relief les différences entre les diverses approches. La mise en place de méthodes objectives de comparaison d'algorithmes d'éclairage est un problème complexe. Si elle est envisageable pour des techniques accélérées de lancers de rayons (benchmark *BART* par exemple [LAM01]) qui visent à produire un rendu photoréaliste, elle paraît beaucoup plus délicate lorsque l'on considère des algorithmes non photoréalistes.

La figure 4.20 montre l'éclairage d'un complexe industriel, l'exploration de cette scène servant à préparer des opérations de maintenance. L'utilisation d'un ombrage classique de type Phong est très problématique, notamment car des panneaux parallèles possèdent la même intensité lumineuse, ce qui rend difficile leur distinction. Un calcul précis d'ambient occlusion, réalisé avec le logiciel d'infographie 3DSMax en 26 minutes, rend mieux compte de la géométrie de la scène ; cependant, les renfoncements du bâtiment apparaissent trop sombres, rendant difficile l'exploration visuelle de ces zones. L'EDL produit ici une image lisible, et pour laquelle les différents plans sont nettement distingués.

Dans la figure 4.21, nous comparons trois approches différentes. L'*exaggerated shading* de Rusinkiewicz et al. [RBD06] nécessite un prétraitement important de la géométrie (ici 1.7 millions de triangles, 500 s de précalculs) pour rendre ensuite une image renforçant les variations locales de relief. La technique d'*unsharp masking* de Luft et al. [LCD06] permet de marquer les discontinuités de profondeur par un traitement image simple, mais elle ne produit un ombrage en tant que tel. Elle renforce simplement une image déjà ombrée obtenue par ailleurs. Au contraire, des techniques telles que

le SSAO et plus particulièrement l'EDL proposent un calcul rapide dont le résultat peut être utilisé tel quel.

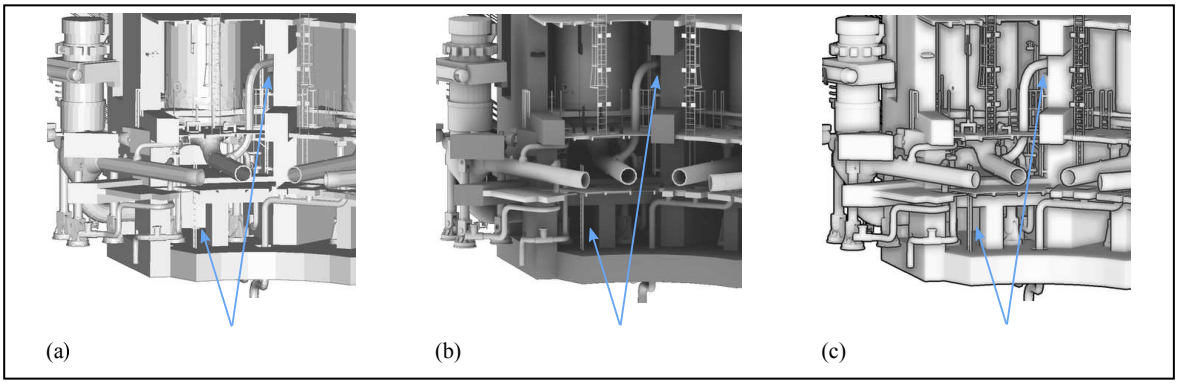


FIG. 4.20: Ombrage d'un complexe industriel. Un rendu correct d'une telle scène est nécessaire pour préparer des opérations de maintenance sur la base de modèles 3D. Ouvriers et ingénieurs doivent pouvoir déterminer simplement et rapidement où se trouvent les éléments critiques (murs et jonctions verticales et horizontales telles que les échelles et les trous pour les tuyaux). Le modèle, composé de 640 000 triangles, est rendu avec un ombrage plat de type Phong (a), un ambient occlusion calculé par pixel sous 3DSMax ((b), 26 minutes de calculs) et avec l'EDL (c). Les flèches bleues pointent sur des régions typiquement mal analysées par les ingénieurs avec les rendus (a) et (b), mais que l'EDL a permis de clarifier rapidement.

4.2 Application à la visualisation chez EDF

Aucune étude utilisateur rigoureuse n'a été menée afin d'évaluer la méthode d'ombrage proposée, mais quelques retours ont été exprimés suite à la présentation d'images produites avec EDL. Ainsi, pour les images de la figure 4.20, l'EDL a reçu la préférence d'experts en sécurité chargés de pointer correctement certains éléments de la scène. Appliqué à de la visualisation scientifique, l'EDL a également reçu un accueil favorable de la part d'ingénieurs spécialisés en neutronique et mécanique des fluides. Evidemment, cela ne présume en rien de la valeur intrinsèque de la méthode, mais permet d'envisager une utilisation concrète dans des applications de visualisation scientifique ou de maintenance, qui donneront lieu à des retours d'expérience plus importants.

A ce sujet, un stage a été réalisé récemment afin de vérifier comment l'EDL peut être intégré dans VTK, un logiciel libre de visualisation scientifique développé par Kitware, et des résultats encourageants ont été obtenus. Le principal obstacle à l'utilisation de techniques en ombrage différé réside pour l'heure dans la rigidité des logiciels de visualisation scientifique, leur architecture n'étant pas pensée pour inclure simplement ces techniques, par exemple sous forme de plugins. EDF et Kitware ont récemment collaboré à la mise en place d'une nouvelle architecture devant faciliter cette intégration. Les travaux de développement sont en cours actuellement.

4.3 Retour à la Perception

Dans une image représentant une scène complexe, de nombreux éléments sont à même de modifier la perception de profondeur induite par l'ombrage.

Le contexte dans lequel une surface apparaît n'est pas neutre. Purves et al. [PWNL04] ont notamment montré qu'il pouvait modifier la perception que l'on a de la luminance d'un élément de surface. Dans le cadre d'un éclairage photoréaliste, certains objets peuvent fournir une indication quant à la

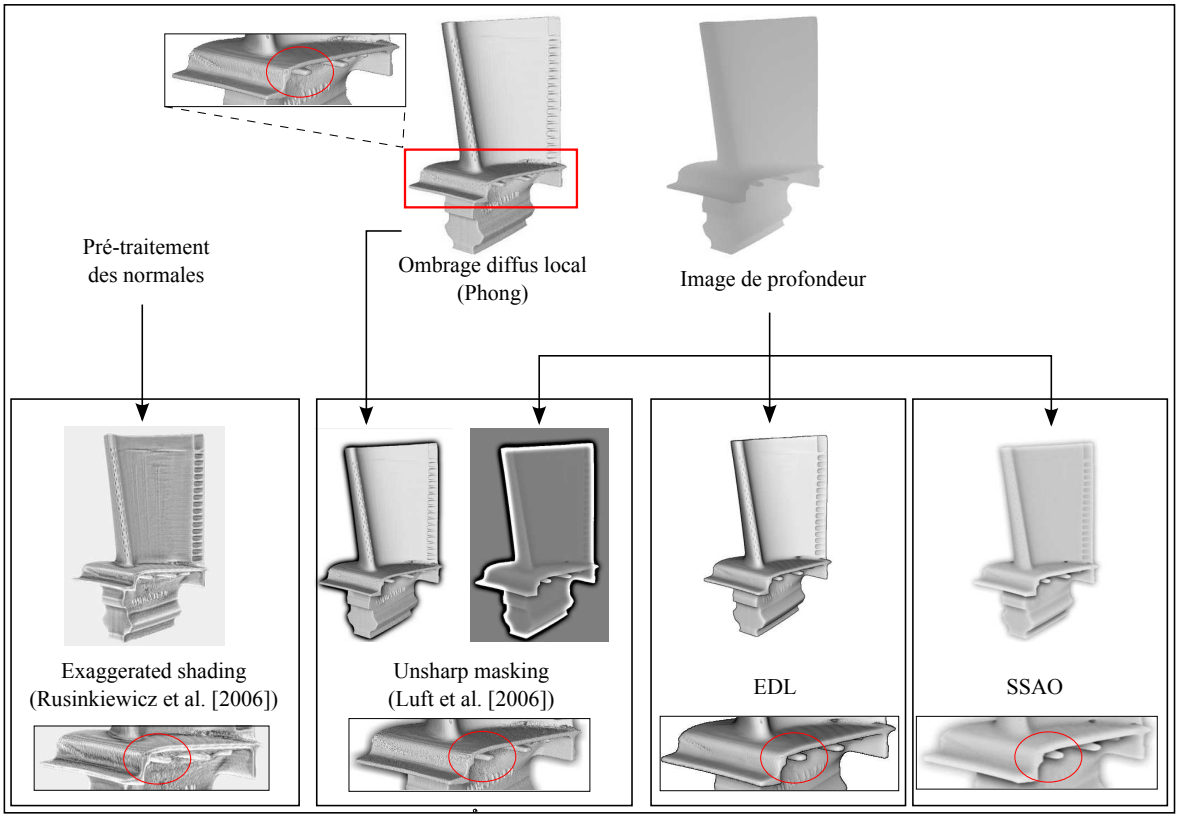


FIG. 4.21: Comparaison de différentes techniques d'ombrage non photoréalistes sur le modèle *blade*. Un zoom sur la géométrie illustre comment ces techniques rendent compte des discontinuités de profondeur. L'*exaggerated shading* [RBD06] produit un éclairage interactif, mais suite à un prétraitement assez lourd. L'*unsharp masking* produit des halos facilitant la perception des relations spatiales, mais doit être intégré à une image déjà ombrée, par exemple avec un éclairage Phong classique. EDL et SSAO génèrent une image nette sur la seule base de l'image de profondeur.

direction principale d'éclairage dans la scène, et modifier en relation la perception des caractéristiques des surfaces dans leur voisinage. Snyder et al. [SDM05] ont ainsi introduit dans une pièce virtuelle une sphère dont la spécularité indiquait une direction d'éclairage différente d'essai à essai, le reste de la scène conservant son éclairage original. Leurs résultats montrent que cet objet supplémentaire influence la perception de l'albédo des différentes surfaces environnantes.

Par ailleurs, en visualisation scientifique, l'emploi d'échelles de couleurs produit des images aux tonalités très variées, sur lesquelles on va rajouter l'information d'ombrage. Or, les indices de textures et de couleurs peuvent modifient de manière non négligeable la façon dont notre système visuel perçoit les formes et la profondeur [RL95, KWYM06], même lorsque l'ombrage seul paraît très informatif.

Cela ne signifie pas qu'il n'est pas important de définir une méthode d'ombrage la plus efficace possible. Simplement, il faut toujours garder à l'esprit que la perception finale des formes et des relations spatiales pourra être perturbée par d'autres indices visuels.

CHAPITRE 5

Focus+Contexte interactif sur GPU avec extrusion conique

En dépit des techniques développées jusqu'à présent, l'analyse et la compréhension de données denses demeurent un problème important en visualisation scientifique. On cherche à afficher le plus d'informations possibles afin d'appréhender le phénomène visualisé, qui est par nature volumique. Cela requiert de superposer des objets pour chaque pixel, et d'être capable, au-delà de l'identification de forme ou de couleur, d'inférer un ordre en profondeur sur les objets reconnus. Il n'existe pas, pour l'instant du moins, de représentations optimales pour ce type d'observations qui sortent du cadre écologique dans lequel notre système visuel s'est construit.

Dans cette tâche, nous sommes confrontés à deux obstacles fondamentaux, à savoir les occultations nécessairement induites par la densité des données volumiques, et la perception de la profondeur et des relations spatiales des structures représentées. Nous avons vu dans le chapitre 3 que le rendu volumique pose en général de nombreux problèmes d'interprétation, même s'il permet en théorie de rendre visible, mais pas nécessairement perceptible, l'ensemble du champ visualisé.

En pratique, l'attention de l'utilisateur engagé dans une visualisation ne se porte à un moment donné que sur un sous-ensemble du volume. Celui-ci peut correspondre à une région particulière de l'espace physique, par exemple une vis pour laquelle on cherche à caractériser le risque de rupture, ou à un sous-ensemble des données, comme des isosurfaces définies par certaines valeurs clés de la variable étudiée. Dans ce cas, il est alors possible de n'afficher que les surfaces ou volumes d'intérêt, ce qui accélère le rendu et règle en partie le problème des occultations. Cependant, une approche aussi radicale engendre une perte de contexte, et il devient difficile voire impossible de situer globalement la zone étudiée, de la mettre en relation avec les autres éléments du volume. Les approches de type Focus+Contexte (F+C) tentent de résoudre ce problème, en proposant des méthodes de rendu différentes pour la région sur laquelle l'attention est portée, le focus, et pour les données environnantes, le contexte. Les techniques de F+C sont très variées, leur finalité peut être simplement de réduire l'information environnante pour accélérer le rendu, d'utiliser de la transparence pour les

régions occultantes du contexte, voire de déformer et grossir les données du focus pour les rendre plus visibles.

Dans ce chapitre, nous proposons une nouvelle méthode de F+C permettant de rendre visible la région de focus quelle que soit sa nature, et de faciliter la distinction des éléments de contexte interposés entre cette région et le point de vue. Pour cela, nous réalisons une découpe ou extrusion conique dans l'image de profondeur de la scène autour des éléments du focus. En effet, une découpe cylindrique, qui consiste à masquer uniquement les éléments directement superposés aux régions d'intérêt, ne fournit aucune indication sur les relations spatiales entre les structures affichées, puisque l'image créée est indépendante de la profondeur du focus. L'implémentation rapide de cette technique sur le GPU par une approche en deux passes est ici notre principale contribution. Pour situer notre travail, nous commençons par présenter certaines des très nombreuses techniques de F+C existantes (section 1). Nous détaillons ensuite l'algorithme proposé (section 2), et montrons comment il s'intègre en pratique dans un processus complet de rendu (section 3) dans lequel la technique d'ombrage présentée dans le chapitre 4 s'applique aisément. Nous terminons par discuter des limites d'une telle approche et des possibles améliorations à apporter (section 4).

1 Techniques variées de Focus+Contexte

1.1 Principes généraux

Les approches de type F+C reposent dans la plupart des cas sur l'utilisation d'une loupe ou lentille définissant l'extension spatiale de la région d'intérêt. Cette lentille peut être définie en espace image (disque, rectangle) ou dans le volume (sphère, boîte). La nature ou les niveaux de détail des données représentées peuvent être sélectionnés explicitement pour les parties internes et externes à cette loupe. Bier et al. [BSP+94] proposent ainsi la métaphore de la lentille magique (*magic lens*), une fenêtre 2D que l'utilisateur peut déplacer pour afficher des informations cachées dans une image ou mettre en valeur une certaine région.

Par ailleurs, les rendus mis en œuvre s'inspirent souvent des techniques utilisées pour les illustrations techniques, pour lesquelles les artistes ne dessinent que les détails essentiels à la compréhension de la structure et du fonctionnement de l'objet, le reste de l'information étant représenté de manière simplifiée, voire omise si elle masque des détails importants. Dans un dessin en écorché, l'enveloppe extérieure de l'objet est partiellement ou intégralement retirée afin de laisser voir les structures internes principales. Eventuellement, les parties occultantes peuvent être ébauchées ou rendues par transparence, au moyen d'une vue fantôme. Une vue éclatée aide à comprendre comment plusieurs éléments de l'ensemble sont agencés, en déplaçant ou en déformant les différentes sous-structures.

1.2 Multirésolutions

Une loupe de focus peut servir en premier lieu à varier le niveau de détails suivant la région de l'espace, ce qui permet d'accélérer le rendu de la scène. Levoy et Whitaker [LW90] présentent une telle approche pour le rendu volumique par lancer de rayons, la densité de rayons par unité de surface décroissant avec la distance à un point d'intérêt. Pour l'affichage de lignes de courant, Fuhrmann et Gröller [FG98] introduisent des loupes 2D et 3D au sein desquelles les densités de lignes affichées sont plus importantes, et Matthaush et al. [MTHG03] augmentent la densité des graines utilisées pour calculer les lignes de courant dans des régions définies par l'utilisateur. Cignoni et al. [CMS94] proposent une *Magic Sphere* dans laquelle l'isosurface est rendue de manière détaillée, alors qu'en dehors une représentation simplifiée est réalisée. Un tel schéma a été développé plus avant pour des grilles régulières par Westermann et al. [WKE99], l'isosurface étant ici progressivement raffinée au

voisinage d'un point choisi par l'utilisateur. Fuhrmann et Gröller [FG98] introduisent des loupes 2D et 3D au sein desquelles les densités de lignes affichées sont plus importantes, et Matthaush et al. [MTHG03] augmentent la densité des graines utilisées pour calculer les lignes de courant dans des régions définies par l'utilisateur. Cignoni et al. [CMS94] proposent une *Magic Sphere* dans laquelle l'isosurface est rendue de manière détaillée, alors qu'en dehors une représentation simplifiée est réalisée. Un tel schéma a été développé plus avant pour des grilles régulières par Westermann et al. [WKE99], l'isosurface étant ici progressivement raffinée au voisinage d'un point choisi par l'utilisateur.

1.3 Focus centré sur les données

Un rendu volumique pour lequel une zone des données est particulièrement opaque peut aussi être considéré comme une méthode de focus+contexte, l'ensemble des données non sélectionnées étant représentées avec une transparence importante afin de rendre compte de zones moins « denses » au voisinage des valeurs d'intérêt. Plusieurs algorithmes de F+C s'appuient ainsi sur une modulation de la fonction de transfert. Viola et al. [VFSG06] utilisent une mesure d'information mutuelle pour définir l'importance de certaines régions, et le rendu volumique est adapté pour chaque pixel de manière à ce que les structures importantes du volume ne soient pas masquées par le contexte environnant. Pour mieux faire ressortir certaines structures principales, Kim et Varshney [KV06] modulent la fonction de transfert par un opérateur de saillance. Une telle approche peut être également employée pour un rendu par nuage de points. Doleisch et al. [DGH03] et Gasser [Gas04] proposent ainsi à l'utilisateur de sélectionner des intervalles d'intérêt sur les histogrammes des données, ce qui permet d'afficher les points correspondants dans l'espace 3D par exemple en couleur alors que le reste apparaît en grisé.

1.4 Déformations et vues éclatées

Pour mieux faire ressortir les régions d'intérêt, certaines techniques modifient la structure spatiale des données, soit en appliquant un grossissement, soit en déformant et et déplaçant les régions environnantes de manière à dégager la vue de l'objet principal.

Pour les rendus volumiques par texture, LaMar et al. [LHJ01] proposent une technique qui permet de grossir une partie de l'image, la transition continue avec les régions rendues à taille normale introduisant une déformation de l'image. Ils discutent notamment des formes de loupe et de région de transition limitant les artefacts visuels. Wang et al. [WZMK05] introduisent la *Magic Volume Lens* : la zone de focus dans l'image agit comme une lentille modifiant la direction des rayons lancés, ce qui assure le grossissement de la zone centrale et une transition douce vers les régions extérieures. Pour des modèles surfaciques, Wang et al. [WLT08] définissent une fonction d'énergie décrivant des contraintes locales et globales, afin de grossir une partie de la surface en limitant au maximum les distorsions et en faisant en sorte que l'objet reste dans le cadre de l'écran. L'inconvénient de toutes ces techniques est l'introduction de déformations géométriques au niveau des transitions, ce qui peut rendre plus difficile la compréhension des relations spatiales entre la région d'intérêt et le contexte environnant.

Afin de mettre la région d'intérêt en évidence et résoudre le problème des occultations, McGuffin et al. [MTB03], Correa et al. [CSC07] et Bruckner et Gröller [BG06] ont développé pour des grilles régulières des techniques inspirées des dessins en éclaté. Ces techniques requièrent une segmentation du volume à partir de laquelle des outils de découpe et de déformation permettent de dévoiler interactivement la région d'intérêt avec un contexte éclaté au voisinage. Bruckner and Gröller utilisent des modèles de force physique pour réaliser la séparation, tandis que Correa et al. proposent

plusieurs interacteurs grâce auxquels ils écartent ou déplient les éléments environnants en fonction de leur distance à la région d'intérêt, le tout étant réalisé sur le GPU.

1.5 Découpes et écorchés

Ces techniques visent à dégager la vue d'objets normalement occultés, en découpant une partie des régions se situant entre elles et l'œil mais sans introduire de déformations dans le rendu.

Pour rendre une scène composée d'objets surfaciques, Diepstraten et al. [DWE03] calculent l'enveloppe convexe des objets intérieurs, et appliquent un masque de stencil permettant de ne pas afficher les surfaces situées devant cette enveloppe. Le principal défaut de la méthode est le calcul de l'enveloppe convexe, qui peut être très coûteux pour de grands modèles. De leur côté, Coffin and Höllerer [CH06] proposent une interface laissant l'utilisateur sélectionner les portions de surface à retirer. Ces deux méthodes sont limitées à une découpe relativement simple de surface masquante.

Li et al. [LRA+07] ont développé une interface très complète permettant de réaliser des illustrations techniques d'objets 3D complexes reposant sur les conventions en usage chez les illustrateurs (figure 5.1 (a)). Les découpes de surfaces occultantes, adaptées aux géométries particulières rencontrées, permettent de dégager la vue des parties importantes, tout en visualisant les différentes couches superposées (coupes obliques ou coniques). L'objectif est de produire des illustrations convaincantes devant faciliter la compréhension de l'organisation spatiale des différents éléments de ces modèles. Leur outil est destiné à faciliter cette tâche pour des utilisateurs connaissant leurs données, anatomiques ou mécaniques, et souhaitant en produire une bonne image de présentation. Cependant, il n'est pas conçu pour explorer des données inconnues, puisqu'il nécessite notamment de préclasser les surfaces rencontrées et de construire pour le point de vue requis un graphe d'occultation qui guidera les coupes. Par ailleurs, ils appliquent des techniques d'ombrage facilitant la perception des coupes et des sauts de profondeur dans la géométrie. Le principe de coupe oblique de Li et al. a par ailleurs été utilisé pour du rendu volumique par Viola et al. [VG05]. Ceux-ci calculent la surface d'extrusion conique autour d'un volume d'intérêt simplement en multipliant la distance dans l'image des pixels à l'empreinte du focus avec la profondeur maximale de cette région. La distance dans l'image est calculée au moyen d'un masque de chanfrein. Viola et al. appliquent ensuite un rendu volumique dépendant de la région traversée. Cependant, avec une telle méthode, si le focus est composé de plusieurs composantes disjointes et situées à des profondeurs différentes, alors l'essentiel du volume risque de se voir creusé inutilement. Burns et al. [BHW+07] proposent également un rendu volumique de type F+C avec une extrusion conique, mais les multiples surfaces d'extrusion semblent précalculées par une méthode non spécifiée.

Enfin, la technique de Krüger et al., baptisée *Clearview* [KSW06], permet de visualiser une surface de focus opaque entourée de plusieurs surfaces de contexte au moyen de vues fantômes (figure 5.1 (b)). Les surfaces de contexte sont partiellement transparentes, des rendus variés facilitant la compréhension de l'ensemble des formes pouvant être sélectionnés via des shaders prédéfinis. S'il est possible de choisir plusieurs surfaces de contexte superposées, il est spécifié que c'est à l'utilisateur de s'assurer qu'elles sont correctement ordonnées en profondeur. Le focus est défini par une sphère centrée sur un point de la surface de focus. L'utilisateur peut déplacer très rapidement le focus dans l'image, mais un changement de point de vue nécessite de recalculer toute la scène, ce qui n'est plus interactif.

1.6 Objectifs

Nous proposons dans ce qui suit une technique d'écorché générique pour laquelle la nature et la géométrie de la zone de focus peuvent être quelconques. L'idée de départ est que l'utilisateur choisit

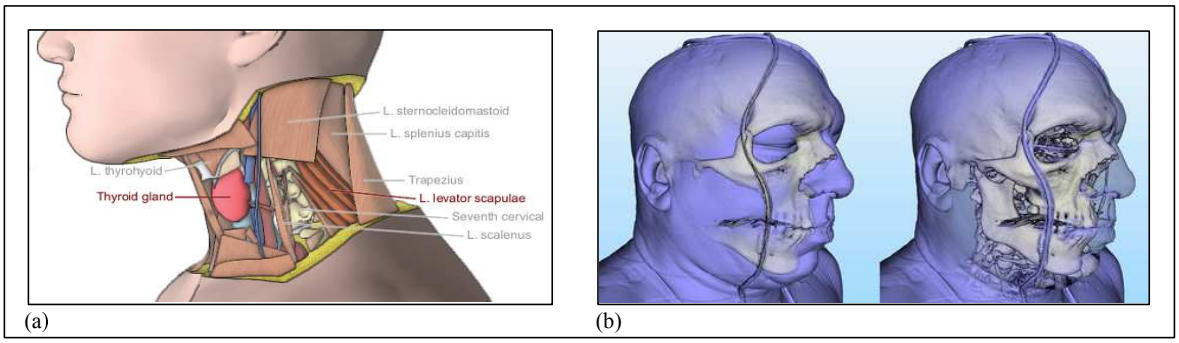


FIG. 5.1: Exemples d'algorithmes produisant des vues en écorché. (a) L'algorithme de Li et al. [LRA+07] utilise un graphe d'occulations et des techniques tirées de l'illustration technique afin de produire une image convaincante. (b) Dans Clearview [KSW06], des vues fantômes permettent de représenter les surfaces occultantes, en modulant leur niveau de transparence par exemple en fonction de la distance à l'œil (image de gauche) ou de la courbure de la surface (image de droite).

un objet dont il souhaite dégager la vue. Cet objet peut être un élément de la géométrie du problème traité (vis, conduit...), un volume englobant (boîte, sphère...) ou une structure extraite des données (isosurfaces, ensemble de lignes de courant tubulées...). La technique implémentée doit permettre de découper le volume situé devant la région d'intérêt, en respectant les contraintes suivantes :

- **Zone de focus quelconque :** notre technique est destinée à dégager la vue d'un objet dont on ne sait rien a priori. Il peut s'agir d'une sphère d'intérêt que l'utilisateur manipule, mais également d'une géométrie calculée à partir des données, comme une isosurface. Dans ce cas, la région d'intérêt est de forme quelconque, pas nécessairement convexe, ni même connexe (voir image 5.11). S'il y a plusieurs composantes, celles-ci peuvent se trouver à des distances diverses de l'œil, ce qui empêche de découper en utilisant simplement une distance dans l'espace image.
- **Découpe autour du focus :** nous cherchons à mettre en place une extrusion conique des surfaces occultantes depuis le focus, de manière à ce que celles-ci soient visibles au moins en partie (voir images 5.7 (a) et (b)).
- **Cohérence au cours du mouvement :** un déplacement continu du point de vue ou des régions de focus (changement de position, modification de la valeur d'intérêt s'il s'agit d'une isosurface) ne doit pas provoquer d'artefacts, il ne doit pas y avoir de saut brutal dans l'image. Ceci vaut notamment pour la surface d'extrusion.
- **Interactivité :** on conserve toujours à l'esprit une exploration des données par un utilisateur qui cherche à comprendre les phénomènes simulés, et qui peut avoir à calculer de nouvelles structures sur lesquelles l'algorithme s'applique. La technique doit donc être aussi rapide que possible, et nécessiter le moins de précalculs possibles, voire aucun.
- **Généralité des données :** dans une visualisation, les objets rencontrés peuvent être de nature très diverses, il peut s'agir de triangles, points, lignes. On ne peut faire l'hypothèse de grilles régulières par exemple, comme c'est le cas dans certains algorithmes.

Comme pour l'EDL présenté au chapitre suivant, une implémentation sur GPU, avec d'éventuelles optimisations, répond aux critères de généralité et d'interactivité. Nous montrons dans ce qui suit comment nous réalisons interactivement l'extrusion conique depuis le volume d'intérêt en creusant l'image de profondeur, et comment cette approche permet d'appliquer des rendus appropriés dans les différentes régions définies par la surface d'extrusion.

2 Algorithme d'extrusion conique sur le GPU

2.1 Partition de l'espace

2.1.1 Définitions et notations

La définition des différents volumes de focus et de contexte est explicitée dans la figure 5.2. Nous considérons un focus de forme quelconque, que nous appelons F. Après projection, celui-ci peut être délimité par sa surface avant et sa surface arrière, respectivement Z^F_{avant} et $Z^F_{arrière}$. L'image de profondeur étant normalisée, on a : $0 < Z^F_{avant}(i,j) < Z^F_{arrière}(i,j) < 1$ pour tout pixel de l'image de coordonnées (i,j) correspondant à un point projeté de F.

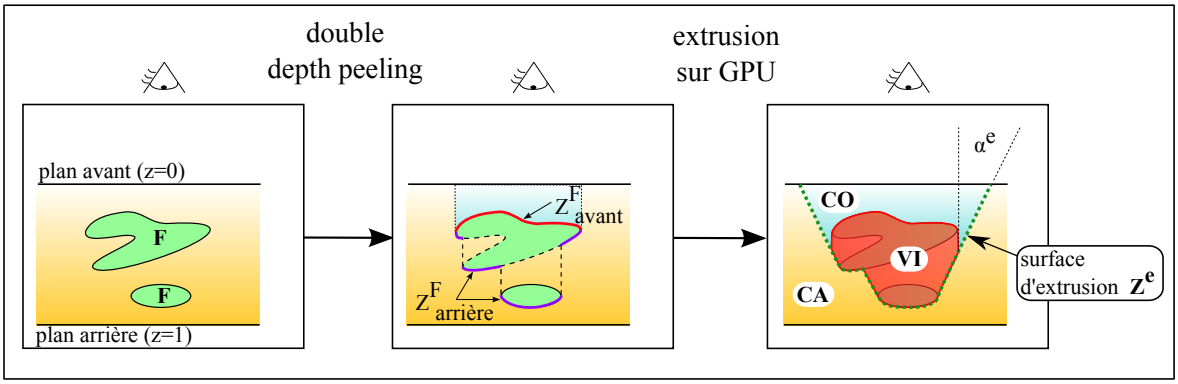


FIG. 5.2: Partitionnement de l'espace en fonction du point de vue. Les surfaces avant et arrière du focus (F) sont calculées par un double depth peeling, et l'image de profondeur est creusée à partir de la surface arrière. La surface d'extrusion ainsi calculée permet de séparer l'espace en trois régions : le volume d'intérêt (VI) autour du focus, le contexte occultant (CO) devant le focus et la surface d'extrusion et le contexte arrière (CA) à l'extérieur de l'extrusion et en arrière plan.

La surface d'extrusion, notée Z^e, correspond à un creusement virtuel de l'image de profondeur réalisé à partir de l'empreinte arrière de F, à savoir $Z^F_{arrière}$. Plus précisément, elle correspond à la surface la plus profonde définie par ce creusement virtuel de l'image de profondeur. Par conséquent, plus les points de Z^e sont distants de F dans l'image, plus faible est leur profondeur. L'extrusion est réalisée en oblique, avec une pente d'angle α^e. La distance maximale dans l'image entre un pixel de F et un pixel de Z^e est appelée D^e. Elle correspond au cas où la surface arrière est située à la profondeur $z = 1$, et est donnée simplement par $D^e = \tan(\alpha^e)$.

Le focus F et la surface d'extrusion définissent un volume d'intérêt noté VI. Le volume situé au dessus de F et de Z^e est appelé contexte occultant (CO), et le volume situé en dessous ou à l'extérieur de la zone de focus et d'extrusion dans l'image contexte arrière (CA). Le contexte occultant correspond à la région de l'espace découpée au-dessus et autour de F, région dans laquelle on ne souhaite pas afficher de données masquant ce qui se trouve en arrière plan. Notons que la définition du volume d'intérêt est plutôt conservative, puisqu'elle peut comprendre des régions de l'espace extérieures au focus si celui-ci est non convexe ou comporte plusieurs éléments disjoints. Par ailleurs, la partition étant définie dans l'image de profondeur, elle est naturellement dépendante du point de vue.

2.1.2 Choix de la surface arrière du focus

La surface d'extrusion pourrait en théorie être calculée aussi bien sur la surface avant du focus que sur sa surface arrière. Pour une image statique, la principale différence résiderait dans la profondeur

effective de cette surface, qui se trouverait plus proche du point de vue dans le premier cas. A contrario, l'utilisation de la surface arrière permet de représenter implicitement l'extension volumique de l'objet focalisé.

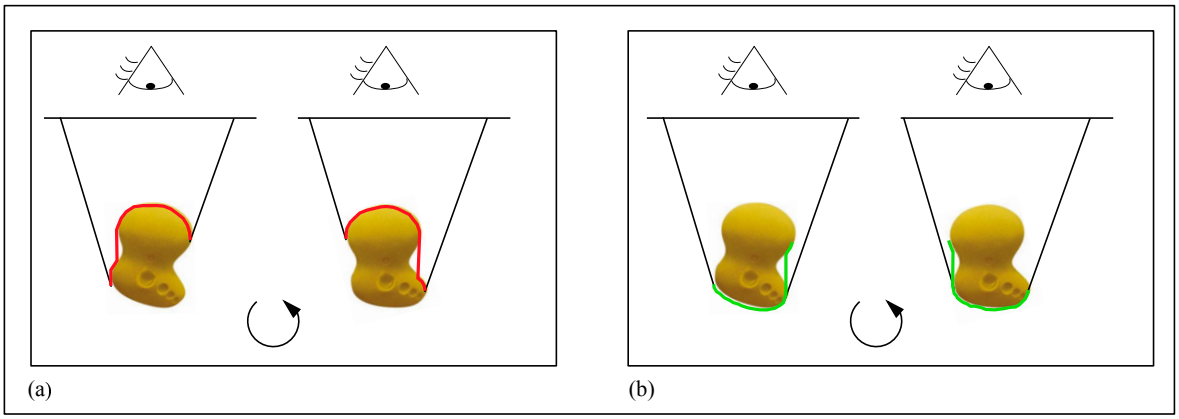


FIG. 5.3: Choix de la surface utilisée pour le calcul d'extrusion conique. Celui-ci est en fait imposé par la définition de la surface d'extrusion comme maximum du creusement de l'image de profondeur. (a) Dans le cas où le calcul repose sur la surface avant du focus (objet jaune), une légère rotation de la scène est susceptible de provoquer une variation brutale de la surface d'extrusion. (b) L'usage de la surface arrière du focus évite ces instabilités et garantit une modification continue de Z^e.

Cependant, lorsque la scène est tournée, la définition de Z^e comme la surface de profondeur maximale obtenue par un creusement d'angle α^e autour du focus est susceptible de provoquer des instabilités si l'on se repose sur la surface avant, comme le montre la figure 5.3. Si le focus est non convexe, une légère rotation du point de vue peut faire varier brusquement ses limites avant et arrière après projection. Dans ce cas, la surface d'extrusion va subir une modification discontinue uniquement si elle est définie relativement à Z^F_{avant}, induisant une brusque variation dans les données affichées.

La technique proposée étant destinée à une exploration interactive pour laquelle l'utilisateur peut modifier le point de vue comme il le souhaite, le calcul de Z^e ne peut donc reposer que sur la surface arrière de la zone de focus.

2.2 Creusement de l'image de profondeur

2.2.1 Formulation

Considérons un pixel quelconque de l'image $P(i,j)$, et un unique pixel de focus $P^f(i',j')$ correspondant à une profondeur « arrière » z^f. Notons d la distance entre ces deux points dans l'image, et d^h et d^v les distances horizontale et verticale. Si l'on utilise la norme L2, on a $d = \sqrt{(d^h)^2 + (d^v)^2} = \sqrt{(i-i')^2 + (j-j')^2}$. Nous reviendrons sur le choix des normes en section 2.2.3.

La profondeur d'extrusion de P générée par P^F est donnée par

$$z^e_{Pf} = z^f - \frac{d}{D^e}, \tag{5.1}$$

en seuillant cette valeur à 0. Si l'on considère maintenant un focus F constitué de plus d'un pixel, la profondeur réelle d'extrusion z^e au niveau de P est définie, d'après ce que nous avons dit précédemment, par

$$z^e = max(z^e_{Pf} / P^f \in \mathrm{F}\ ,\ d(P,P^f) \leq D^e). \tag{5.2}$$

Ainsi formulée, on voit bien que la surface d'extrusion que l'on crée est l'enveloppe maximale en profondeur de toutes les extrusions possibles.

Cependant, un tel calcul est trop coûteux s'il est appliqué simplement sous cette forme. En effet, une implémentation directe demanderait pour chaque pixel de réaliser le calcul pour un voisinage de taille $\Theta(D^{e2})$. Le creusement de l'image de profondeur n'étant qu'une partie du processus de rendu, la contrainte d'interactivité nécessiterait alors de restreindre D^e à des petites valeurs, et donc d'utiliser de faibles angles d'extrusion ne dévoilant que faiblement les différentes surfaces superposées au dessus du focus. Nous montrons dans la section suivante comment accélérer ces calculs, une séparation de l'algorithme en deux passes successives permettant d'atteindre une complexité de $\Theta(D^e)$.

2.2.2 Optimisation en deux passes

Pour commencer, les limites arrière et avant du focus sont recueillies en une seule passe OpenGL en utilisant la méthode proposée par Bavoil et Myers [BM08]. Au moment de dessiner le focus, le test de profondeur est désactivé, et le blending est activé avec une fonction *MAX*. La profondeur z des fragments projetés est écrite comme $(-z, z)$ dans une texture *FLOAT_RG32_NV* initialisée à $(-1., -1.)$, ce qui permet effectivement d'enregistrer les profondeurs minimales et maximales en une seule passe de rendu.

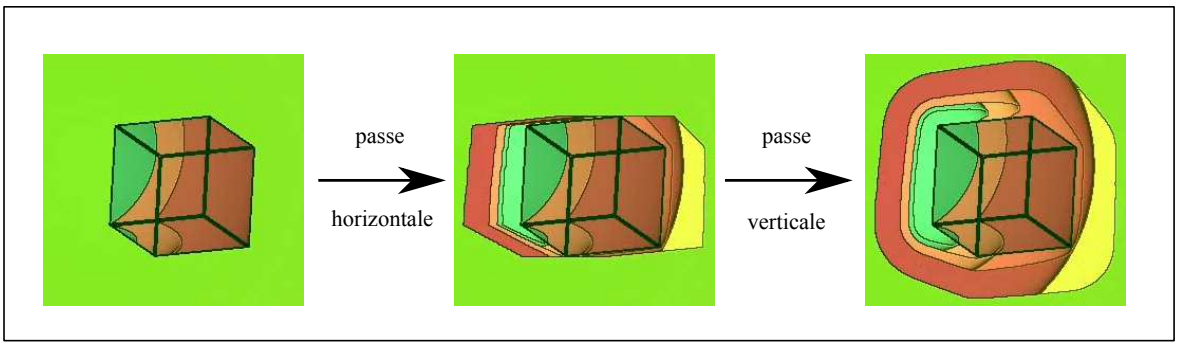
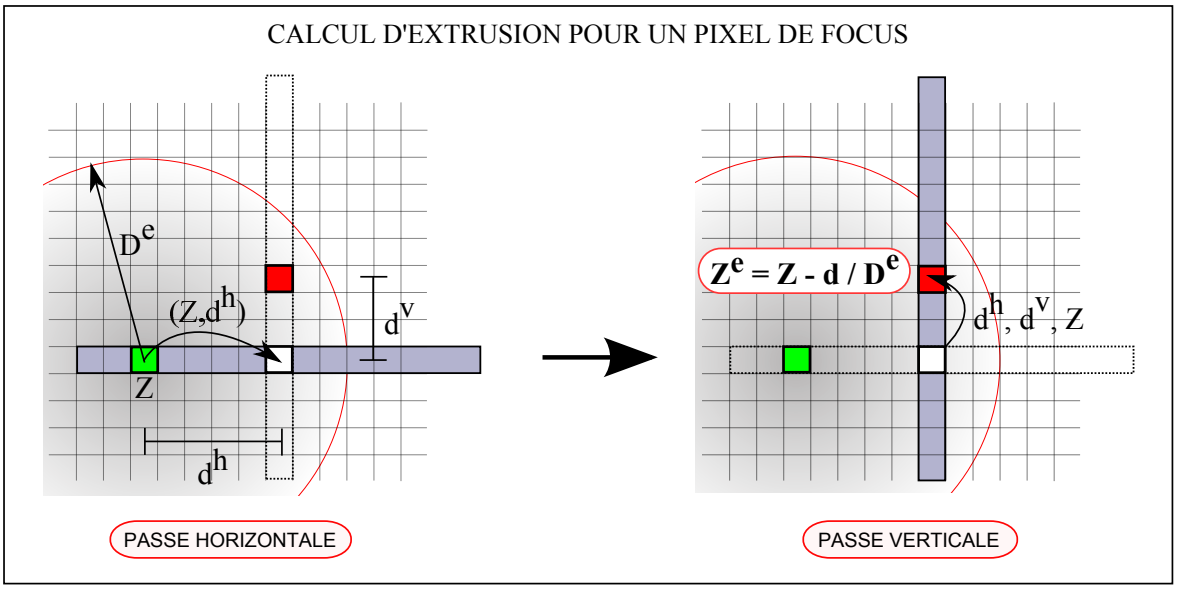


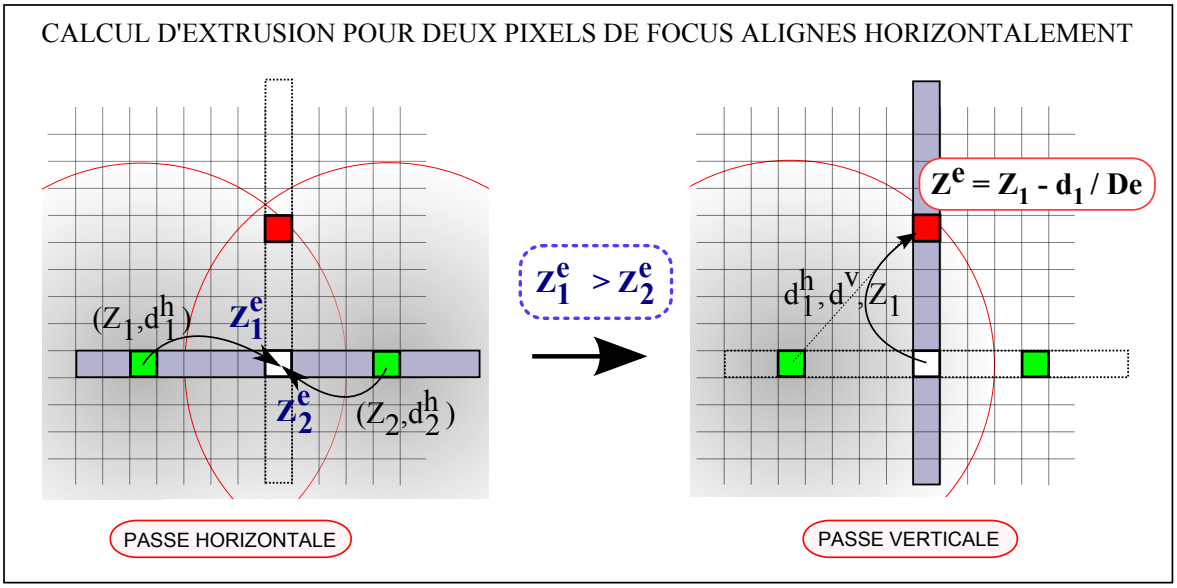
FIG. 5.4: Calcul de la surface d'extrusion conique en deux passes, le focus étant ici défini par un cube (représenté en filaire). L'image de profondeur est creusée horizontalement lors de la première passe, et verticalement au cours de la deuxième passe.

Le calcul de la surface d'extrusion est ensuite divisé en deux passes successives réalisées sur le GPU en GLSL, comme cela est fait en général pour les filtres séparables (une passe horizontale suivie d'une passe verticale, voir figure 5.4). Les résultats intermédiaires et finaux sont enregistrés dans des textures. L'approche proposée est explicitée dans le cas simple où il n'y a qu'un seul pixel de focus sur la figure 5.5 (a). Au cours de la première passe, pour chaque pixel P un voisinage de rayon D^e est analysé. Si un pixel de focus P^f est trouvé, sa profondeur z^f ainsi que sa distance d^h à P sont écrites en P. Au cours de la seconde passe, cette information est lue pour tous les pixels présents dans un voisinage vertical de rayon D^e, et la profondeur d'extrusion est calculée sur la base de la distance verticale au pixel pivot : $z^e = z^f - \frac{d(P,P^f)}{D^e}$, où d dépend de d^h et d^v. Pour que cela fonctionne, les valeurs de profondeur et de distance horizontale dans la texture intermédiaire sont initialisées à 0.

Si la région de focus comprend plus d'un pixel, ce qui est généralement le cas, l'algorithme décrit ci-avant peut encore être appliqué en modifiant juste certains détails. Le principe en est illustré sur la figure 5.5 (b). Pour chaque pixel P de l'image, on calcule lors de la première passe la profondeur d'extrusion engendrée par chacun des pixels du focus inclus dans un voisinage horizontal de rayon D^e. On enregistre alors en P la profondeur z^f et la distance horizontale d^h générant l'extrusion la plus



(a)



(b)

FIG. 5.5: Principe de l'algorithme d'extrusion conique en deux passes. En vert : pixels de focus. En blanc : pixels pivot. En rouge : exemple de pixels pour lesquels la profondeur d'extrusion est calculée. (a) Au cours de la première passe, un voisinage horizontal de rayon D^e est parcouru. La profondeur et la distance dans l'image du pixel de focus trouvé sont reportées dans le pixel courant (pivot). Ces informations permettent dans une deuxième passe de calculer la profondeur d'extrusion au niveau du pixel rouge, en analysant son voisinage vertical. (b) Lors de la passe horizontale, si plusieurs pixels de focus se trouvent dans le voisinage analysé, on reporte les informations du pixel générant la profondeur d'extrusion horizontale la plus importante ($Z_1^e > Z_2^e$).

profonde. Lors de la seconde passe, la profondeur d'extrusion définitive est simplement le maximum des profondeurs calculées relativement à tous les pixels présents dans le voisinage vertical de rayon

D^e :

$$z^E = max(z_j^f - \frac{d(d_j^h, d_j^v)}{D^e}, j \in [-D^e, D^e]). \tag{5.3}$$

La surface d'extrusion calculée par l'algorithme est illustrée concrètement sur la figure 5.6. Pour mieux percevoir son relief, nous avons ici appliqué un texturage correspondant à des lignes de niveau guidées par la profondeur des différents pixels, et ajouté un ombrage local diffus, les normales à la surface d'extrusion étant calculées sur le GPU au moment du rendu en utilisant les profondeurs des pixels voisins.

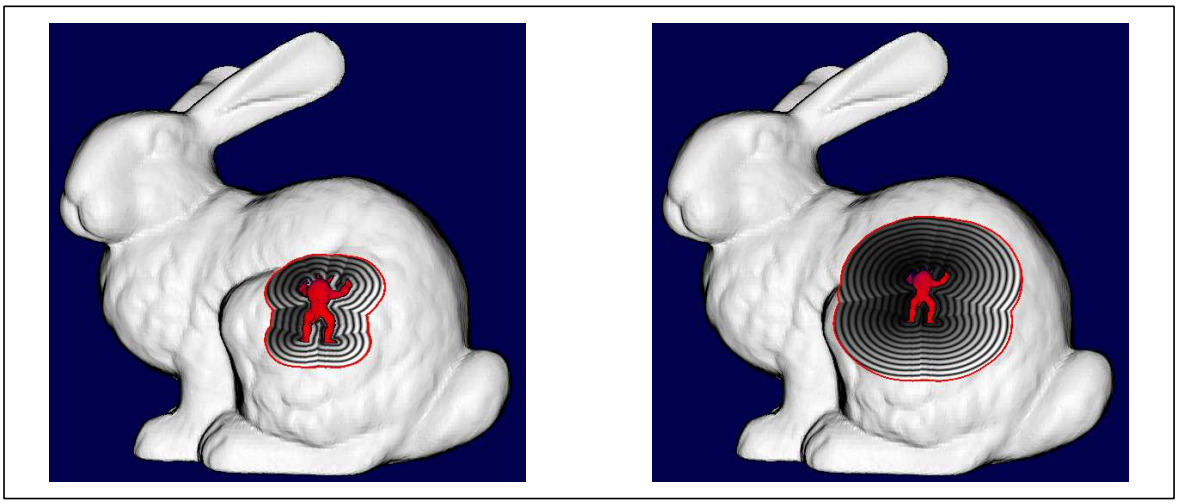


FIG. 5.6: Représentation de la surface d'extrusion conique, au moyen de lignes de niveau en profondeur et d'un ombrage diffus. Ici, le modèle Armadillo sert de focus (en rouge). Il est positionné à l'intérieur du lapin, à deux profondeurs différentes.

2.2.3 Choix de la norme

L'algorithme proposé repose sur l'hypothèse que pour deux pixels de focus P_1^f et P_2^f situés sur une même ligne et de profondeurs respectives z_1^f et z_2^f, une extrusion maximale lors de la première passe implique une extrusion maximale de manière générale. C'est-à-dire que pour un pixel P situé à des distances horizontales respectives d_1^h et d_2^h de P_1^f et P_2^f, et à des distances totales de d_1 et d_2 :

$$\left(z_1^f - \frac{d_1^h}{D^e} > z_2^f - \frac{d_2^h}{D^e}\right) \Longrightarrow \left(z_1^f - \frac{d_1}{D^e} > z_2^f - \frac{d_2}{D^e}\right). \tag{5.4}$$

En réalité, cette affirmation est vraie si l'on considère une distance basée sur la norme L1 dans l'image, mais pas sur la norme L2. Ainsi, pour $d = d^v + d^h$, on peut vérifier aisément que

$$\left(z_1^f - \frac{d_1^h}{D^e} > z_2^f - \frac{d_2^h}{D^e}\right) \Longrightarrow \left(z_1^f - \frac{d_1^h + d_1^v}{D^e} > z_2^f - \frac{d_2^h + d_2^v}{D^e}\right). \tag{5.5}$$

alors que l'assertion suivante est fausse :

$$\left(z_1^f - \frac{d_1^h}{D^e} > z_2^f - \frac{d_2^h}{D^e}\right) \Longrightarrow \left(z_1^f - \frac{\sqrt{(d_1^h)^2 + (d_1^v)^2}}{D^e} > z_2^f - \frac{\sqrt{(d_2^h)^2 + (d_2^v)^2}}{D^e}\right). \tag{5.6}$$

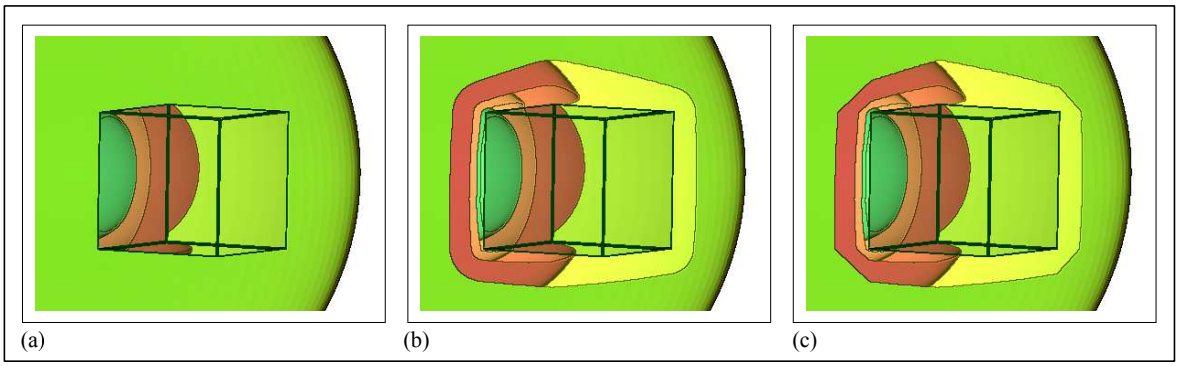


FIG. 5.7: Choix de la norme pour l'extrusion conique. (a) Un creusement cylindrique ne permet pas de situer le focus cubique au sein des sphères concentriques. (b) Une norme L2 produit un résultat satisfaisant avec l'algorithme en deux passes, même si celui-ci repose sur une approximation. (c) Le passage à une norme L1 garantit l'exactitude de l'extrusion, mais génère en contrepartie des intersections anguleuses.

Cependant, en considérant une profondeur appartenant à l'intervalle $[0:1]$, l'approximation réalisée en utilisant malgré tout la norme L2 est acceptable. En effet, les cas problématiques représentent un volume très limité de l'espace des paramètres, et génèrent a priori une erreur de faible amplitude dans le calcul de la surface creusée. L'intégrale volumique correspondant aux cas problématiques a été estimée au moyen d'une simulation de Monte-Carlo, en choisissant aléatoirement $1 < D^e < 150$ (en pixels), $1 < d_1^h, d_2^h, d^v < D^e$ et $(z_1^f, z_2^f) \in [0,1]^2$. Elle représente moins de 1% du volume global de l'espace de ces paramètres.

En conclusion, il est possible d'utiliser la norme L2 pour réaliser le calcul d'extrusion conique en deux passes, et nous n'avons pas noté de discontinuités inattendues dans tous les rendus réalisés avec cet algorithme. Mais si l'on veut être certain d'obtenir une coupe exacte, alors le choix de la norme L1 s'impose. L'emploi de ces deux normes est illustré sur les figures 5.7 (b) et (c).

2.3 Performances

Les temps de calcul de la surface d'extrusion sont principalement déterminés par la taille de l'image et par la distance maximale d'extrusion D^e, qui déterminent linéairement le nombre d'accès textures et d'opérations durant les deux passes. Pour réduire ce coût potentiellement important, il est possible de réaliser l'extrusion à une résolution différente, en utilisant des fenêtres virtuelles de rendu (FBO, *FrameBuffer Object*) plus petites sur lesquelles on projette la double image de profondeur obtenue initialement. En calculant par exemple Z^e sur une image deux fois plus petite dans chaque dimension, on divise par quatre le nombre d'opérations tout en multipliant par deux la distance maximale d'extrusion effective lorsque la surface calculée est utilisée par la suite dans l'image de taille complète, au moyen d'une interpolation linéaire de texture.

Évidemment, le recours à des résolutions inférieures est moins précis et susceptible de générer quelques artefacts dans l'image, comme le montre la figure 5.8. Cependant, une telle approche peut être mise en œuvre dans des situations dynamiques, par exemple lorsque l'utilisateur tourne la scène ou modifie la région focalisée, la surface d'extrusion exacte étant alors calculée lorsque plus rien ne bouge. Une telle variation des niveaux de détails en fonction du mouvement a été proposée notamment pour les rendus de surface par de nombreux auteurs [FS93, Red, YPG01], et est motivée par le fait que le système visuel humain perçoit moins de détails dans les objets en mouvement.

Le tableau 5.1 indique les temps de calcul de la surface d'extrusion pour différentes tailles

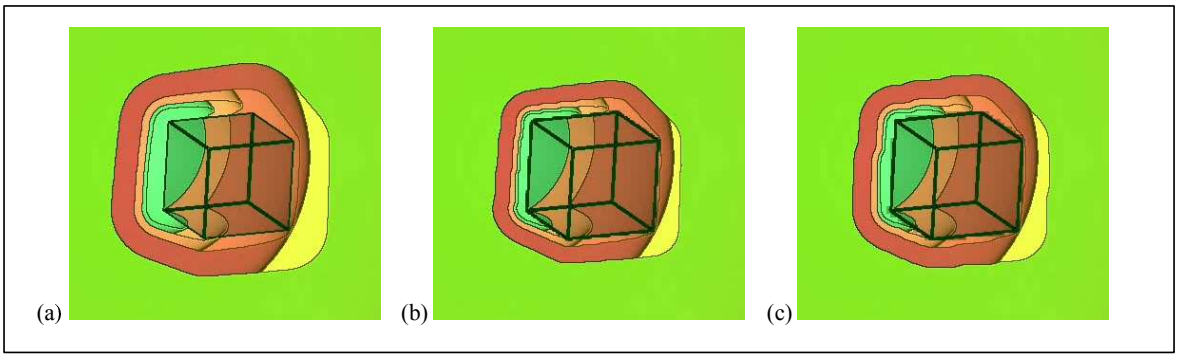


FIG. 5.8: Réalisation du calcul d'extrusion à différentes résolutions. L'extraction de la surface d'extrusion peut s'avérer trop coûteuse si elle est réalisée sur l'image de profondeur en pleine résolution (figure (a), $D^e = 150$). Pour accélérer cette étape, il est possible d'utiliser une projection sur une fenêtre de taille divisée par deux (figure (b), $D^e = 50$) ou quatre (figure (c), $D^e = 30$). Le nombre de calculs à réaliser est alors fortement réduit, d'autant plus que la taille du voisinage à explorer dans l'image est alors divisé d'autant à angle d'extrusion constant. Cependant, l'interpolation linéaire réalisée au moment d'utiliser une image de résolution inférieure pour le rendu final introduit quelques artefacts visibles au niveau des bords découpés.

Taille de fenêtre : 1024x1024		Taille de fenêtre : 512x512			Taille de fenêtre : 256x256		
D^e	FPS	D^e	D^{e*}	FPS	D^e	D^{e*}	FPS
1	640	1	2	820	1	4	1560
25	59	25	50	82	14	56	320
50	30	50	100	44	19	76	220
75	20	75	150	31	25	100	175
100	14	100	200	23	27	108	124
150	9	150	300	16	50	200	96

TAB. 5.1: Performances de l'algorithme d'extrusion conique sur GPU. Les résultats sont donnés en images par seconde (FPS) pour différentes amplitudes d'extrusion et plusieurs résolutions (la carte de profondeur, de taille 1024*1024, est alors projetée sur une fenêtre de taille 512x512 ou 256x256). Pour les résolutions inférieures, l'amplitude d'extrusion D^e (en pixels) est équivalente à une extrusion de taille D^{e*} quand la carte résultante est utilisée dans la fenêtre de résolution fine.

de fenêtre et distances maximales. Ces temps ont été obtenus avec une carte NVIDIA GeForce 8800GTX, pour un focus cubique dont la projection occupait approximativement $1/16^{ème}$ de l'image. Dans l'implémentation actuelle, les calculs des deux passes sont réalisés pour tous les pixels de l'image, ce qui fait que les performances moyennes obtenues avec un focus cubique deux fois plus gros sont quasi identiques. Un précalcul permettant de déterminer dans l'image l'extension maximale des régions d'extrusion pourrait permettre de réduire les temps de calcul pour des focus peu étendus, à condition d'être rapide à réaliser.

3 Rendu complet en écorché

3.1 Architecture générale

Nous montrons dans cette section comment le creusement conique explicité précédemment s'intègre dans une architecture de rendu de visualisation générique entièrement portée sur GPU. Cette architecture est aussi flexible que possible. L'utilisateur peut choisir simplement les éléments appartenant aux différentes régions de la partition de l'espace, et les rendus à appliquer pour chaque région, ces derniers pouvant être enrichis facilement au moyen de shaders ajoutés à la chaîne de traitement. Par contre, il n'a pas à se soucier d'un quelconque ordonnancement des structures dans la profondeur.

L'architecture complète est illustrée sur la figure 5.9. Elle peut être décomposée en trois étapes principales :

- **Calcul du volume d'intérêt (en bleu) :** l'objet définissant le focus est rendu une première fois afin d'extraire ses surfaces avant et arrière en fonction du point de vue, le tout en une seule passe grâce au *double depth peeling*. La surface d'extrusion est ensuite calculée comme décrit dans la section 2. Les limites du volume d'intérêt sont conservées dans une texture *RGB*, pour chaque pixel sous la forme $\mathrm{Z_{VI}} = (Z^e, Z^F_{avant}, Z^F_{arrière})$. Pour les pixels extérieurs à la projection de F, on enregistre $\mathrm{Z_{VI}} = (Z^e, 1, 0)$. Comme nous l'avons mentionné auparavant, l'objet définissant le focus peut être de nature quelconque, défini par la géométrie du problème ou extrait des données.
- **Rendu dans la partition (en orange) :** les données sont rendues hors-écran, séparément dans chaque région de la partition spatiale (volume d'intérêt, contexte occultant, contexte arrière). Un tag associé à chaque structure est utilisé pour préciser dans quelle partie de l'espace celle-ci doit être rendue. Un shader unique permet de gérer les rendus des trois régions, et le tag permet de définir si l'élément doit être ou non affiché en comparant sa profondeur après projection à la carte $\mathrm{Z_{VI}}$.
- **Composition (en rouge) :** l'étape précédente a permis d'obtenir les images correspondant aux trois régions de l'espace qui peuvent être maintenant superposées. La carte $\mathrm{Z_{VI}}$ contient l'information requise pour déterminer la nature du pixel en fonction de sa position dans l'image : celui-ci peut correspondre à un point du focus, à un point de l'extrusion conique qui l'entoure, ou à un point extérieur pour lequel seul le contexte arrière est affiché. Les occultations sont gérées ici au moyen de vues fantômes, et différentes techniques permettant de moduler la transparence peuvent être appliquées. Il est possible également d'ajouter des indices facilitant la perception de la distance entre le focus et l'arrière plan. Ces aspects sont traités dans la section 3.2.

Toutes les opérations décrites ont été implémentées en OpenGL, en utilisant un rendu hors écran reposant sur des FrameBuffer Objects (FBO). Le rendu de surfaces et de lignes est réalisé avec un shader commun, tandis que des shaders dérivés peuvent être implémentés pour des objets spécifiques comme *point sprites*.

Par ailleurs, les trois images contiennent chacune une information de profondeur recueillie au moment de la projection. Il est donc possible d'ajouter des effets supplémentaires, comme le calcul de silhouettes, de halos ou de perspective atmosphérique. Ainsi, toutes les images issues du programme présentées dans ce chapitre ont été ombrées au moyen d'une version simplifiée de l'algorithme EDL présenté dans le chapitre 4.

3.2 Gestion des occultations et distance à l'arrière plan

Au moment de la composition des trois images, pour lesquelles on dispose conjointement des informations de couleur et de profondeur, et la carte $\mathrm{Z_{VI}}$ qui définit le partitionnement dans l'image et

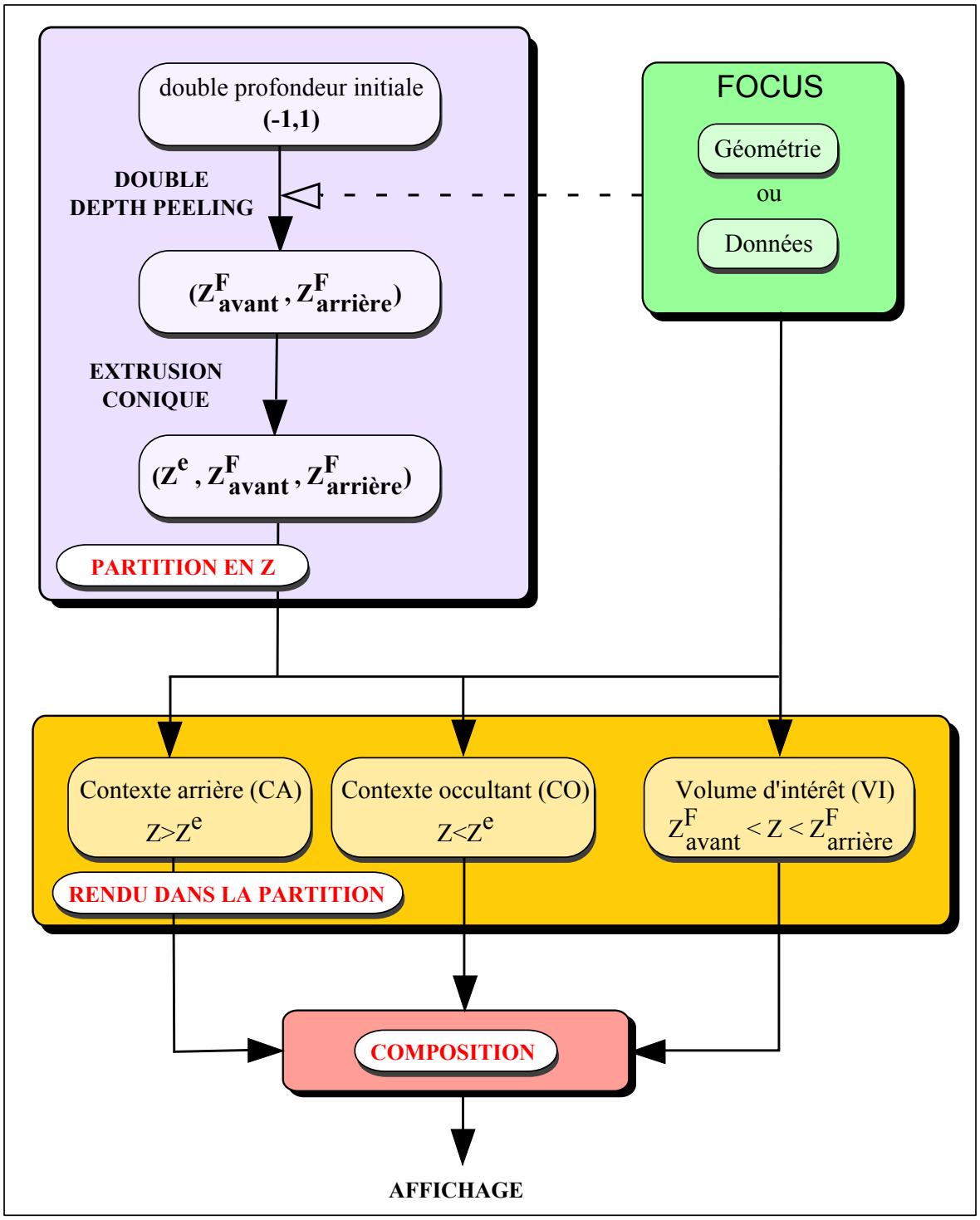


FIG. 5.9: Architecture complète de rendu avec partionnement en profondeur et extrusion conique.

en profondeur, il est possible d'ajouter des effets destinés à faciliter la compréhension de la structure spatiales des données affichées dans la zone de l'image correspondant au focus et à l'extrusion qui l'entoure. Des effets différents peuvent être appliqués au contexte occultant et aux éléments situés en arrière plan.

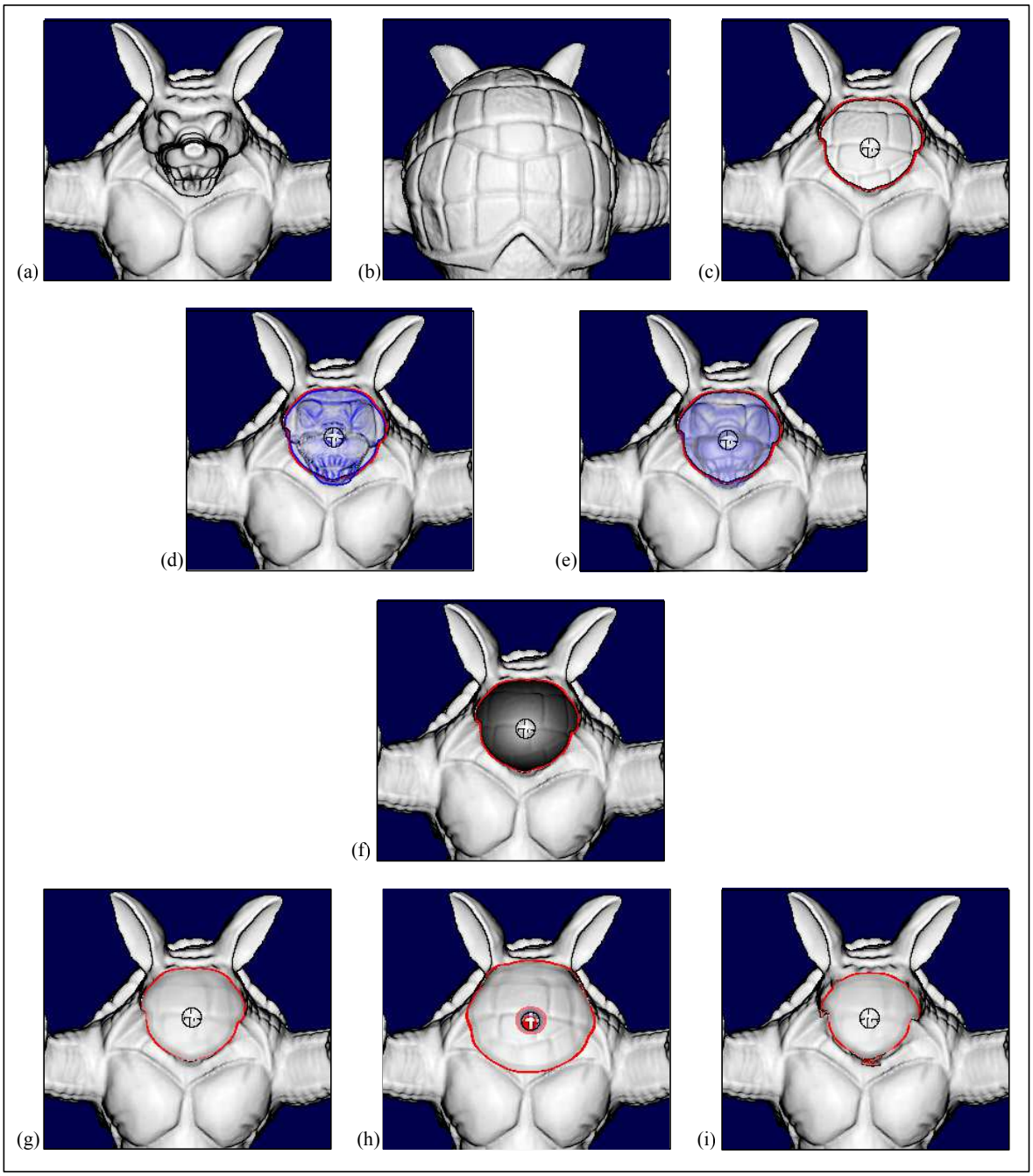


FIG. 5.10: Exemples de vues fantômes du contexte occultant et d'indices facilitant la perception de la distance au contexte arrière sur le modèle Armadillo. (a) Vue de face. (b) Vue de dos. (c) Vue de face, avec creusement par un focus sphérique sans ajout d'indices visuels. (d) Vue fantôme de la surface avant, transparence modulée par la courbure. (e) Vue fantôme de la surface avant, transparence modulée par l'orientation (silhouettes). (f) Assombrissement de la surface arrière en fonction de la distance à la surface d'extrusion. (g) Ajout de flou sur la surface arrière, dont l'extension varie avec la distance à la surface d'extrusion. (h) Idem, focus plus proche de la surface arrière. (i) Idem, focus plus proche de l'œil.

3.2.1 Contexte occultant

Etant donnée dans le shader la carte de profondeur de l'image avant, Z_{CO}, il est facile de calculer des informations géométriques supplémentaires telles que les normales et courbures en considérant cette carte comme un champ de hauteur. La normale $\vec{N}(P)$ en un pixel P peut être obtenue comme la somme des gradients horizontaux et verticaux de Z_{CO} en P. Une estimation de la courbure c_P peut être calculée en utilisant la méthode proposée par [KSW06], qui s'appuie sur l'opérateur en parapluie de Kobbelt et al. [KCVS98]. Si l'on considère les normales des quatre pixels voisins de P ($\vec{N}_{1,2,3,4}(P)$, une approximation de la courbure en P est obtenue comme la somme des distances entre ces normales et celle en P, à savoir :

$$c_P = \sum_{i\in[1:4]} \left\| \vec{N}(P) - \vec{N}_i(P) \right\| \tag{5.7}$$

Comme proposé par Kruger et al. [KSW06], il est alors possible d'afficher une vue fantôme de la surface occultante. Pour cela, la surface avant est rendue par transparence, avec une modulation de l'opacité α_P en fonction de l'orientation de la surface en ce point, pour faire ressortir les silhouettes (figure 5.10 (e)), ou de la courbure estimée (figure 5.10 (d)). Pour la modulation par silhouettes, il suffit par exemple de choisir $\alpha_P = 1 - \vec{N}(P).\vec{V}$, où $\vec{V}$ est la direction de vue ($\vec{V} = (0,0,1)$), et pour les courbures $\alpha_P \propto c_P$.

3.2.2 Contexte arrière

Pour un pixel P appartenant à la projection de la surface d'extrusion, on dispose de la profondeur associée à cette surface, z_P^e, et de la profondeur de la surface dessinée pour l'arrière plan, z_P^{CA}.

Une première information facilitant la lecture de l'image est le marquage de la zone découpée. En théorie, celle-ci correspond aux pixels vérifiant $z_P^e = z_P^{CA}$. En raison de la discrétisation de l'affichage, il faut plutôt considérer l'ensemble des pixels P tels que $(z_P^{CA} - z_P^e) < \varepsilon$. Dans la figure 5.10, la frontière de la découpe réalisée est marquée en rouge (ici $\varepsilon = 0.02$).

Lorsque les données affichées ne sont pas denses dans le volume, il peut être difficile de percevoir la distance entre le volume d'intérêt et les éléments de contexte situés en arrière plan. Ainsi, dans l'image 5.10(c), la carapace du modèle semble collée au focus sphérique et beaucoup plus proche de l'œil qu'elle ne l'est en réalité. Deux techniques ont ici été implémentées pour limiter ces artefacts visuels. En premier lieu, on peut obscurcir les zones de l'arrière plan les plus éloignées de la surface de creusement en fonction de la distance $z_P^{CA} - z_P^e$, en s'inspirant de l'a priori perceptif « dark is deep » (figure 5.10 (f)). Il est également possible de simuler un effet de profondeur de champ en ajoutant du flou à la surface arrière. La quantité de flou, c'est-à-dire la taille du voisinage sur lequel il est réalisé, peut également être modulé par la distance entre la surface arrière et la surface d'extrusion, de telles sorte qu'une surface lointaine apparaîtra très floutée (figures 5.10 (g)-(h)-(i)). Si l'ajout d'une profondeur de champ approchée apporte un peu d'information en statique, l'effet est plus important en dynamique, quand la modification continue de netteté de l'image facilite la perception de l'éloignement ou du rapprochement de la zone de focus par rapport au contexte environnant.

3.3 Quelques exemples

Nous montrons maintenant l'application de l'architecture de rendu implémentée avec l'extrusion conique sur quelques exemples de visualisation scientifique, en choisissant des régions de focalisation variées.

L'exemple de la figure 5.11 correspond au refroidissement d'une structure de complexe industriel par un circuit fluide. Pour vérifier l'efficacité du système de refroidissement, la température calculée est affichée au moyen de surfaces d'isovaleur. Dans cet exemple, les isosurfaces définissent la région

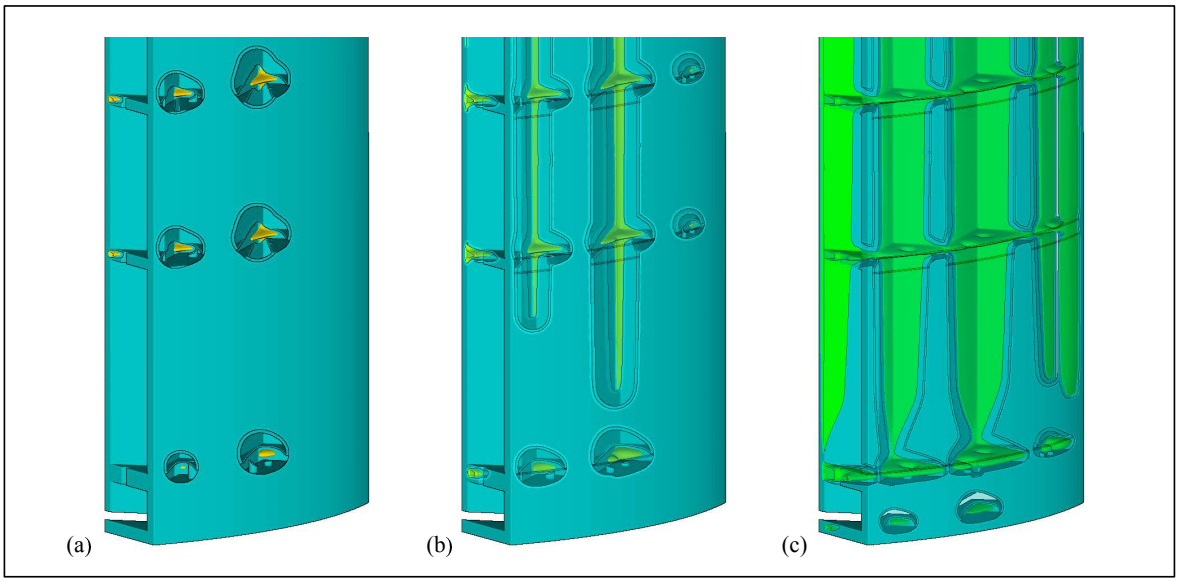


FIG. 5.11: Exemple d'extrusion conique par isosurface de température non connexe. (a) Les surfaces occultantes sont complètement retirées. (b) Modification de l'isovaleur. Les surfaces occultantes sont représentées en transparence simple, avec des lignes caractéristiques (plis) en noir. Les découpes sont marquées par des traits plus clairs. (c) Idem que (b), pour une isovaleur de température inférieure.

de focus, la peau du domaine de calcul masquant ces isosurfaces étant creusée par l'algorithme, ce qui permet de clarifier les positions et extensions spatiales des isosurfaces par rapport à la géométrie environnante. Il est important de noter qu'ici la région de focus est composée de plusieurs composantes connexes situées à des profondeurs distinctes. La peau du maillage retirée par extrusion est affichée par transparence ou au moyen de lignes caractéristiques. Par ailleurs, le calcul des isosurfaces sur ce modèle comprenant 3 millions de tétraèdres est réalisé directement sur le GPU, en utilisant les nouvelles fonctionnalités des cartes graphiques et d'OpenGL (*geometry shaders*), si bien que l'utilisateur peut interactivement modifier la valeur d'intérêt dans les données et observer directement le résultat.

La figure 5.12 représente la circulation d'un fluide entre deux cylindres. Une grille est implantée perpendiculairement à la direction principale, afin de modifier la structure de l'écoulement. Ici, la zone de focus est un cube d'intérêt positionné autour de la grille, de façon à appréhender l'apport de cette structure. A l'intérieur de cube sont affichées certaines lignes de courant caractérisant le mouvement du fluide, ainsi que des lignes caractéristiques du maillage de la grille (pliures). L'extrusion autour du cube aide à positionner mentalement ces éléments au sein de la structure complète, et un rendu transparent de la surface occultante clarifie l'intersection du cube avec le cylindre extérieur.

Pour finir, la figure 5.13 est construite à partir de données issues de la simulation du chauffage d'une pièce par un radiateur. Des isosurfaces sont extraites du champ de température, et des lignes de courant tubulées illustrent le mouvement de l'air. Cet exemple montre que l'extrusion conique peut reposer sur des structures spatiales relativement complexes sans recourir au calcul de représentations intermédiaires comme une enveloppe convexe.

Les exemples présentés s'appuient uniquement sur des représentations solides, mais il est possible d'intégrer également un rendu volumique dans les différentes régions, dans la mesure où la carte Z_{VI} définissant le volume d'intérêt et la surface d'extrusion fournit les limites en profondeur de ces régions. Dans ce cas, il est envisageable de modifier la fonction de transfert du DVR de manière à faire plus ou moins ressortir certaines informations, et à limiter l'occultation du focus par le volume

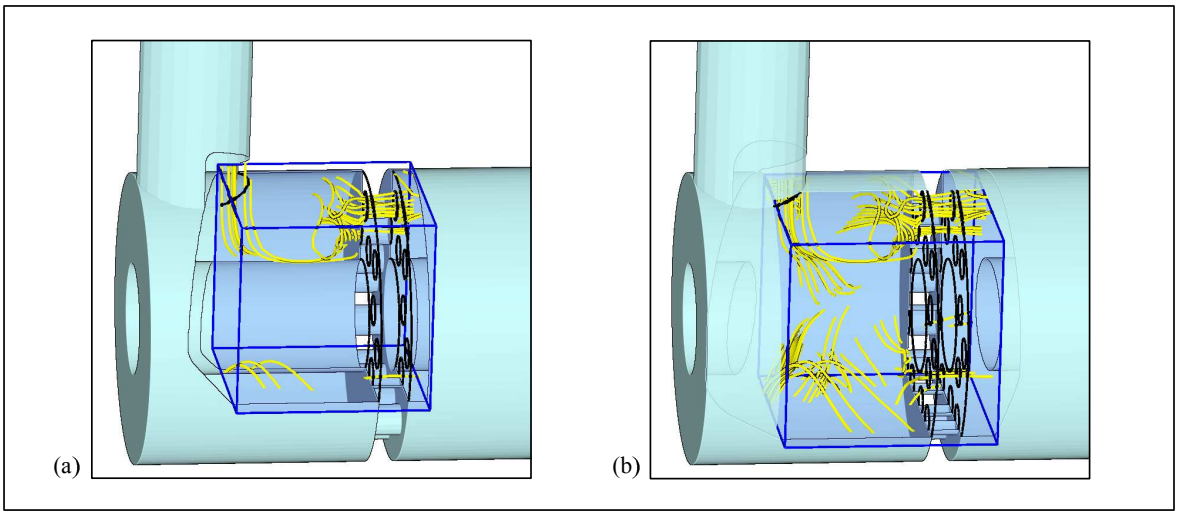


FIG. 5.12: Exemple d'extrusion conique par cube d'intérêt, pour une simulation de fluide circulant entre les parois de deux tubes cylindriques. Le cube intersecte le volume, certaines lignes de courant du fluide sont affichées à l'intérieur du volume d'intérêt, ainsi que les lignes caractéristiques de la peau du maillage. L'extrusion permet de visualiser la géométrie de la structure solide environnante, avec affichage par transparence de la surface coupée en (b).

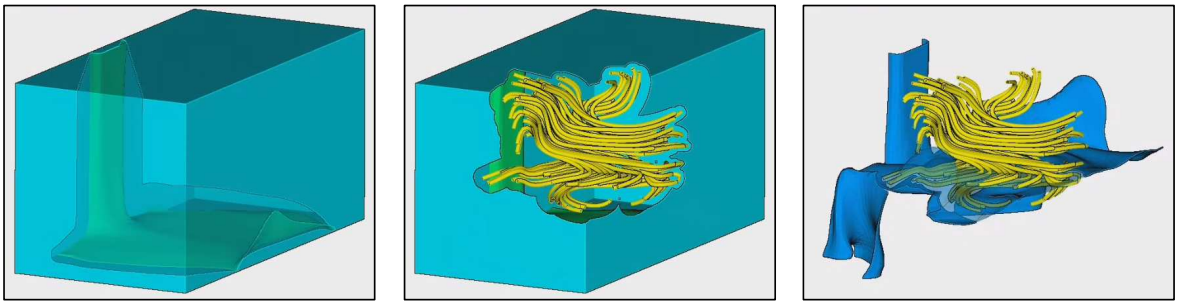


FIG. 5.13: Exemple d'extrusion conique par isosurface et lignes de courant, dans un cas de simulation de chauffage d'une pièce par un radiateur. (a) Le creusement est réalisé autour d'une isosurface de température. (b) Des lignes de courant décrivant le mouvement de l'air dans la pièce servent de région de focus, la peau du maillage et une isosurface sont affichées dans le contexte arrière. (c) Le focus est constitué des mêmes lignes de courant, une isosurface est affichée dans le contexte arrière et par transparence en avant plan.

situé en avant plan.

3.4 Performances et limites

Le principal goulet d'étranglement concerne le calcul de la surface d'extrusion, qui est d'autant plus coûteux que l'angle de découpe α^e est important. Pour rappel, des ordres de grandeur des temps de calcul requis ont été explicités dans la section 2.3.

Il n'est pas cependant nécessaire de recalculer la surface d'extrusion à chaque image, mais seulement lorsque la région de focus, le point de vue ou l'angle de découpe α^e sont modifiés. En dehors de ces cas, il est possible de changer les données et structures affichées dans les différentes régions de

la partition sans autre coût que celui du rendu de ces éléments. De plus, pour une image donnée, les paramètres utilisés dans l'étape de composition peuvent toujours être ajustés en temps réel, puisqu'ils ne nécessitent que l'affichage simultané des trois images correspondantes et quelques opérations par pixel. Il s'agit notamment du mode et de l'amplitude de transparence de la surface occultante et du type d'indices facilitant la perception de la distance à la surface arrière

Dans l'implémentation actuelle de l'architecture décrite dans la figure 5.9, les objets doivent être affichés autant de fois que le nombre de régions dans lesquelles on souhaite les voir apparaître. Cela peut limiter fortement l'interactivité pour des maillages de grandes tailles rendus par exemple dans les zones de contexte arrière et occultant.

4 Discussion et perspectives

4.1 Ecorché par découpe pseudo-conique

La technique proposée dans ce chapitre s'inspire des méthodes d'écorché pratiquées en illustration technique. L'accent a été mis sur la réalisation interactive d'une découpe pseudo-conique autour des régions focalisées, cette découpe étant définie en fonction du point de vue et indépendamment de la nature des structures devant être rendues, que ce soit celles délimitant la région d'intérêt ou celles appartenant au contexte environnant.

Cette méthode d'écorché permet dans certains cas de clarifier les relations spatiales entre la zone d'intérêt et la géométrie dans laquelle elle est intégrée. Ainsi, l'exemple illustré par la figure 5.11 a été présenté aux ingénieurs responsables de l'étude concernée. Ceux-ci se sont montrés enthousiastes quant à la possibilité offerte de discerner clairement la position des isosurfaces dans la structure industrielle servant de base aux simulations réalisées, tout en pouvant modifier interactivement l'isovaleur de température. Dans un tel cas, ils sont plus habitués à utiliser certaines lignes caractéristiques du maillage de peau (plis ou crêtes), mais une telle information filaire est plus délicate à appréhender. La personne responsable du calcul a adhéré au fait que le rendu soit complètement interactif, tout en soulignant qu'un simple rendu en transparence du déflecteur avant pouvait suffire dans le cas où un seul occulteur se trouve devant l'isosurface. Un autre ingénieur a par ailleurs proposé une application originale de l'algorithme, qui consisterait à utiliser le découpage conique pour s'assurer visuellement que trois isosurfaces représentant des champs différents (température, irradiation, contraintes mécaniques) ne présentent aucune intersection. L'interactivité permet alors de parcourir pour ces différentes variables des plages de valeurs correspondant à des régions critiques de la physique simulée.

Cependant, l'approche proposée ne peut être efficace en toutes circonstances, comme le montre la figure 5.14. Lorsque de très nombreuses surfaces sont interposées entre l'œil et la région d'intérêt, la compréhension de la découpe devient particulièrement complexe, et cette méthode générique n'est pas nécessairement la plus adaptée pour clarifier les relations spatiales entre les différents éléments visualisés. Une méthode plus proche des conventions utilisées en illustration technique, comme celle implémentée par Li et al. [LRA+07], serait plus efficace, mais elle ne peut être appliquée que dans une phase postexploratoire, car elle requiert une segmentation efficace des éléments affichés et certains précalculs que l'on souhaite ici éviter. De manière générale, notre approche peut être utilisée pour offrir interactivement un aperçu de la scène en déplaçant librement le point de vue, et d'autres types de rendu peuvent être employés pour continuer l'exploration visuelle ou pour produire des images convaincantes en vue de prendre les décisions induites par les résultats de la simulation réalisée.

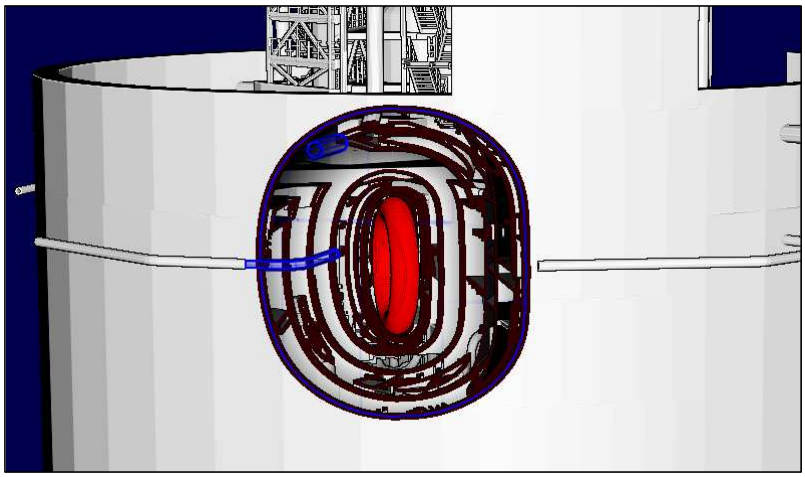


FIG. 5.14: Limites de l'algorithme lors du creusement de multiples surfaces. Dans ce cas, le focus torique est situé à l'intérieur d'un bâtiment comprenant de nombreux murs. Un creusement conique ne tenant pas compte des spécificités des surfaces rencontrées ne permet pas ici de comprendre clairement la structure spatiale de l'ensemble.

4.2 Perception de la profondeur

Dans les rendus proposés, un problème délicat concerne la perception de la distance entre les éléments appartenant à la région de focus et les éléments de contexte situés en avant ou en arrière plan. Nous avons ébauché certaines méthodes destinées à faciliter cette perception, mais le problème reste entier. De plus, une approche comme la simulation approchée de profondeur de champ brouille les éléments arrière, ce qui peut nuire à la reconnaissance des structures clés de la visualisation.

Il s'agit ici d'un problème délicat qui survient principalement dans des représentations volumiques éparses. En effet, si l'on choisit de rendre de manière quasi opaque les données situées en arrière plan de la zone d'extrusion, par exemple au moyen d'un rendu volumique ou en affichant les valeurs scalaires d'un champ sur cette surface de coupe, la lecture de l'image est plus aisée, mais on perd dans le même temps la notion des surfaces caractéristiques plus éloignées. Une solution que nous n'avons pas illustrée ici consisterait à utiliser des halos autour de la région de focus, en utilisant par exemple l'algorithme d'*unsharp masking* de [LCD06] en couplant les images de profondeur des régions de focus et de contexte arrière. Plus simplement, il est aussi possible de réaliser cet effet au moyen de l'EyeDome Lighting, qui dans la version actuelle n'est utilisé que dans la résolution la plus fine et indépendamment pour chacune des trois images de profondeur calculées.

4.3 Surfaces multiples

Une limite importante de la méthode proposée concerne le nombre de surfaces affichées simultanément. Dans la version actuelle, il n'est en effet possible de rendre qu'une surface dans le contexte occultant, et l'affichage d'un élément dans deux à trois régions de la partition spatiale nécessite autant de projections de cet élément. Il s'agit ici essentiellement de limites algorithmiques et matérielles, pour lesquelles certaines solutions peuvent être envisagées.

Concernant le rendu du contexte occultant, nous nous sommes pour le moment limités à afficher en transparence la surface la plus proche du point de vue, ce qui permet d'établir une continuité avec le contexte arrière sur les bords de la surface d'extrusion. Cependant, il est tout à fait possible

de choisir d'afficher de préférence les éléments les plus proches de la région focalisée. Pour cela, il suffit simplement en OpenGL de modifier la nature du test de profondeur réalisé, en choisissant l'option *GL_GREATER* plutôt que *GL_LESS* et en initialisant l'image de profondeur à 0 au lieu de 1. Il est également envisageable d'afficher ces deux surfaces conjointement pour un coût algorithmique négligeable, en ayant recours à la technique de double depth peeling déjà implémentée pour le calcul des limites avant et arrière de la région de focus.

En ce qui concerne le rendu simultané d'éléments situés à des profondeurs différentes, le problème principal vient de la nature de l'architecture des cartes graphiques, qui n'autorise à l'heure actuelle l'utilisation que d'un unique tampon de profondeur. Pour résoudre ce problème, Bavoil et al. [BCL+07] suggèrent d'utiliser un algorithme de *k-buffer* qui généralise le principe du *z-buffer* actuel à plusieurs couches. Leur implémentation permet de réaliser en une seule passe l'acquisition de plusieurs surfaces superposées, et de réaliser très rapidement des effets de transparence et de flou de profondeur. Dans l'état actuel, elle est cependant soumise à des aléas de lecture et d'écriture sur la carte pouvant générer des artefacts visuels, et les auteurs suggèrent donc qu'une telle architecture puisse être intégrée directement dans le schéma général de rendu des cartes graphiques. Dans notre cas, une telle approche permettrait d'une part de recourir à une unique passe pour rendre les structures appartenant aux trois régions de la partition, d'autre part d'afficher dans le contexte occultant plusieurs surfaces, par exemple au moyen de rendus par silhouettes. Dans ce dernier cas, il faudrait néanmoins prendre garde à gérer intelligemment les occultations induites par ces éléments supplémentaires, et s'assurer en outre que ces nouvelles informations n'altèrent pas la compréhension générale de l'image.

Conclusion

Contributions

Le travail présenté dans ce mémoire a été en premier lieu guidé par les entretiens réalisés en début de thèse avec des ingénieurs de la R&D d'EDF. Ces discussions avaient pour but d'essayer d'appréhender concrètement les diverses difficultés relatives à la visualisation de gros volumes de données issus de simulations numériques. Suite à ces entretiens, nous nous sommes particulièrement concentrés sur la perception de la profondeur, des formes et du volume dans les scènes de visualisation complexes. Nous avons considéré uniquement le cas de visualisations réalisées sur un écran ordinaire, et non pas doté de système de vision stéréoscopique. Nous n'avons pas non plus abordé les difficultés relatives à la visualisation de plusieurs pas de temps.

Dans un premier temps, nous avons évalué la capacité du système visuel humain à percevoir des relations de profondeur dans un rendu volumique direct (DVR). Ce mode de représentation consiste à accumuler par transparence les différents éléments présents dans le volume, en tenant compte de l'ordre de leur distribution en profondeur. Les images produites permettent ainsi de présenter l'ensemble des informations du volume. Cependant, de telles représentations ne sont pas naturelles pour le SVH qui est généralement confronté à des surfaces opaques. Les expériences de psychophysique que nous avons réalisées ont effectivement montré que même dans des cas simples les rendus volumiques induisent une forte ambiguïté de perception de la profondeur. L'ajout d'informations dynamiques à travers la rotation de la scène et le recours à une projection perspective amplifiée peuvent aider à lever en partie ces ambiguïtés.

Nous avons ensuite proposé deux nouveaux algorithmes de rendu pour lesquels les calculs sont réalisés entièrement sur la carte graphique, en utilisant essentiellement la carte de profondeur obtenue après projection des objets à visualiser. Cette approche permet de s'abstraire des caractéristiques géométriques de ces objets, une visualisation pouvant inclure en pratique des primitives très diverses. Elle contribue de plus fortement à l'interactivité de la visualisation si l'on considère les nouvelles architectures graphiques dans lesquelles le nombre de processeurs de flux est en constante augmentation.

L'EyeDome Lighting (EDL) est une nouvelle technique interactive d'ombrage non photoréaliste

utilisant uniquement les informations de profondeur au voisinage d'un pixel. Une approche multirésolution couplée à un filtrage bilinéaire pondéré par les différences de profondeur limite les calculs à réaliser, et permet de produire des pseudo-halos facilitant la séparation des objets en profondeur. Si l'EDL peut être vu comme un dôme de lumière situé au-dessus de chaque pixel, nous avons également proposé une variation de l'algorithme permettant d'orienter la direction apparente de l'éclairage, ce qui peut aider dans certains cas à clarifier interactivement certaines régions plus sombres de la scène. Une autre extension de l'EDL consiste à moduler le contraste de l'ombrage en profondeur et dans l'image, afin de focaliser l'attention de l'observateur sur certaines régions plus importantes.

Nous avons ensuite présenté une technique particulière de rendu en écorché, pour laquelle un creusement conique est calculé en fonction du point de vue autour d'un ou de plusieurs objets focalisés. Cet algorithme de focus+contexte permet de gérer les occultations éventuelles de la région sur laquelle l'analyse se concentre, tout en essayant de clarifier les relations spatiales des éléments de contexte environnants. Nous avons vu que cette technique peut effectivement faciliter la perception de certaines scènes de visualisation, mais qu'il existe des cas dans lesquels elle n'est pas efficace en l'état, notamment lorsque la scène comprend de nombreuses surfaces superposées.

Les algorithmes d'ombrage et d'extrusion conique ont été implémentés sur des prototypes indépendants, et les premiers résultats ont été montrés à un échantillon très restreint d'utilisateurs potentiels à EDF. Les premiers retours étaient plutôt positifs, et ont même fait jaillir de nouvelles pistes d'application pour un algorithme de type vue en écorché et extrusion conique (visualisation du recouvrement d'isosurfaces de grandeurs physiques différentes). Cela illustre l'intérêt, en visualisation scientifique, de confronter de nouvelles idées aux pratiques des ingénieurs, cet échange pouvant être particulièrement riche : comprendre les besoins spécifiques des utilisateurs peut aider à développer de nouvelles solutions plus adaptées aux caractéristiques de leur métier, et les discussions suivant la présentation de prototype peuvent faire émerger de nouvelles idées fécondes.

Pour pouvoir toucher un public plus large et générer plus de retours d'expérience, il reste désormais à intégrer l'EDL et les vues en écorché aux outils de visualisation utilisés à EDF. Cela ne peut cependant être réalisé que sur des logiciels développés en interne ou non propriétaires. Il est ainsi prévu d'intégrer le nouvel ombrage présenté dans cette thèse dans Paraview, le logiciel libre de visualisation parallèle basé sur VTK, ce qui permettra une utilisation concrète de cette technique dans des cas de visualisation scientifique à EDF. On peut ainsi espérer obtenir des retours d'expérience à plus long terme, qui permettront d'évaluer l'intérêt effectif de cette technique après mûrissement par les ingénieurs concernés.

Enfin, notons que ces algorithmes exploitant les nouvelles possibilités des cartes graphiques ont été montrés à différents acteurs du marché logiciel. Ceux-ci ont semblé apprécier les rendus produits, et ils ont souligné qu'ils étaient effectivement en retard quant à l'intégration dans leurs outils de solutions permettant de réaliser de tels rendus.

Perspectives

La visualisation scientifique est une discipline relativement récente [1] et qui connaît une évolution importante du fait de l'intérêt croissant des ingénieurs pour les solutions proposées. En effet, les utilisateurs potentiels ont souvent mis un temps important avant de se tourner vers l'utilisation de la 3D pour analyser les résultats de leurs calculs, soit par réticence à quitter des techniques de visualisations

[1] La première définition en a été donnée par McCormick en 1987 : « l'utilisation de l'informatique graphique pour créer des images visuelles qui aident à comprendre la représentation complexe et souvent massive de concepts ou résultats scientifiques » (source : http ://en.wikipedia.org/wiki/Scientific_visualization).

par graphes auxquelles ils étaient habitués, soit par manque de temps pour se former aux nouveaux outils, soit parce que les logiciels disponibles ne permettaient pas de répondre suffisamment interactivement et simplement à leurs attentes. Cependant, le développement de simulations de plus en plus poussées générant de très grands volumes de données fait émerger des besoins très forts en visualisation pour les ingénieurs concernés, et apparaître de nouvelles problématiques. Dans une entité comme EDF R&D, ce mouvement de la visualisation scientifique est actuellement très marqué. Il pousse à s'interroger sur les conditions requises pour que l'étape de visualisation soit la plus efficace possible, et à proposer de nouveaux modèles de représentation adaptés aux différents cas rencontrés.

Considérations générales

En premier lieu, il faut toujours garder à l'esprit que les images produites ont un but précis, qui est de faire comprendre les résultats issus de simulations numériques. Comme le décrit Durand [Dur02], il en résulte que la production d'images est un processus d'optimisation. En fonction du but fixé dans l'étude, et implicitement en fonction du domaine physique auquel cette dernière est rattachée, la visualisation doit permettre de produire les images les plus efficaces possibles pour répondre aux questions que l'on se pose. Par conséquent, il ne saurait y avoir une visualisation scientifique, mais plutôt des visualisations scientifiques partageant certains outils, certaines représentations. Les spécificités propres aux différents métiers constituent un élément important que l'on ne saurait ignorer. Ces spécificités peuvent résulter de la nature des champs physiques observés, mais aussi des phénomènes principaux caractérisant le domaine de recherche (turbulences, oscillations particulières...), qui ont pu forger au fil du temps des modes de représentations archétypiques. C'est en se tournant vers la littérature consacrée et vers les ingénieurs concernés, voire en impliquant directement ces derniers, que l'on sera en mesure de proposer des solutions adaptées aux problèmes étudiés. Une partie essentielle du travail dans ce domaine de recherche consiste donc à établir des passerelles entre la communauté graphique et celle des utilisateurs directs. Il est enrichissant de pouvoir observer en pratique les utilisateurs finaux des outils développés, de comprendre précisément leurs attentes, d'interagir pour essayer de leur faire ressentir ce qui est éventuellement réalisable afin que de nouvelles idées puissent émerger de ces échanges. Comme l'ont suggéré certains auteurs, il peut aussi être intéressant de faire participer des artistes au processus de création de nouveaux modes de visualisation. En effet, ceux-ci, de par leur pratique, ont en général développé un sens aigu des styles et conventions les plus efficaces pour faire passer visuellement certaines informations. Cependant, une telle collaboration représente un exercice délicat, car elle nécessite que tous parlent le même langage, ou du moins que chacun soit capable de comprendre ce que l'autre exprime.

Pour continuer, revenons sur le sens, sur la fonction de l'image scientifique. Initialement, le dessin scientifique, par exemple tel qu'il était pratiqué par Léonard de Vinci, avait pour but de permettre l'élaboration d'un raisonnement, il était le moyen de parvenir à une compréhension des phénomènes naturels observés. La représentation constituait un moyen d'accéder à la découverte. Aujourd'hui, l'utilisation de photographies ou de films à grande vitesse aide à comprendre certains phénomènes complexes non accessibles par un autre moyen de mesure, comme la formation des bulles de champagne dans un verre ou la méthode de propulsion de la méduse dans l'eau. De même, la visualisation scientifique vise à produire des images dont l'objectif principal est d'aider à appréhender les phénomènes complexes qui ont été simulés. Les images générées ne constituent pas une fin en soi, mais plutôt le moyen par lequel l'ingénieur peut analyser ses résultats et aboutir à des conclusions éventuellement étayées par des chiffres précis. Par conséquent, ces images doivent pouvoir être produites le plus interactivement possible, pour ne pas bloquer le raisonnement de celui qui visualise. Il ne s'agit pas de noyer l'utilisateur sous un flot ou une sélection d'images précalculées qu'il devrait

par la suite observer passivement. Au contraire, il est nécessaire de lui laisser toute liberté de créer sa propre visualisation, de laisser libre cours à ses idées, ses impressions, il faut permettre que les objets visualisés suivent l'ordre logique de ses déductions. Ces nécessités d'interactivité dans l'exploration et de possibilité de découverte par l'exploration permettent de définir plusieurs critères essentiels que doit respecter une visualisation scientifique, et ce d'autant plus que les tailles des simulations réalisées explosent.

D'abord, les algorithmes utilisés doivent impérativement être rapides et nécessiter le moins de prétraitements possibles. En effet, face à la croissance des volumes de données à analyser, il paraît de plus en plus difficile d'envisager une préparation exhaustive de l'ensemble des champs et structures directement issus des calculs ou qui pourraient être produits au cours de la visualisation. Une telle approche n'est pas à bannir complètement, le précalcul de différents niveaux de résolution d'un maillage constituant par exemple un apport important pour la navigation interactive au moment du rendu. Elle est cependant à utiliser avec parcimonie, dans des cas bien particuliers et limités. De plus, pour aller vers plus d'interactivité, il paraît indispensable que les algorithmes utilisés en visualisation puissent tirer partie des avancées technologiques, les nouvelles architectures graphiques offrant de plus en plus de possibilités relatives aux calculs propres au rendu. Et pour une utilisation concrète de ces nouvelles techniques, il faut ici souhaiter que les logiciels de visualisation disponibles s'adaptent à ces évolutions, ou permettent de le faire en proposant des architectures plus ouvertes sur lesquelles il soit possible d'intégrer efficacement de nouveaux algorithmes proposés dans les laboratoires de recherche.

Aussi, les visualisations proposées doivent lever les obstacles perceptifs pouvant gêner la compréhension de la scène affichée. En particulier, la perception des formes, de la profondeur et des relations spatiales des différentes structures en présence doit être la moins ambiguë possible, sans quoi l'observateur sera considérablement gêné dans son raisonnement. L'utilisation d'algorithmes d'éclairages perceptivement adaptés et manipulables interactivement plutôt que des modèles locaux de type Phong nous paraît en ce sens très importante. De même, la levée des ambiguïtés de profondeur générées par le mouvement dans des affichages peu denses, par exemple dans des rendus filaires ou de nuages de points, constitue un point important nécessitant sans doute une réflexion attentive : si l'ingénieur voit tourner la scène à l'envers lorsqu'il bouge la souris, il risque de se trouver déstabilisé, et perdre le fil de ses pensées, si ce n'est sa compréhension de l'objet observé.

Enfin, les images produites doivent éviter de procurer une fausse impression d'exactitude ou de vérité. Elles doivent certes être exactes par rapport à ce qu'elles représentent, ne pas introduire de distorsions faussant le raisonnement. Mais il ne faut pas qu'elles induisent une trop grande confiance dans ce qui est visualisé, il faut qu'elles permettent une certaine distanciation vis-à-vis du support de raisonnement, c'est-à-dire une possibilité permanente de prendre du recul vis-à-vis de ce qui est observé. En ce sens, le (photo)réalisme du rendu nous semble fortement pénalisant, car donner le sentiment à un ingénieur qu'il regarde un objet réel, une forme de vidéo d'expérience réalisée sur maquette, c'est risquer d'introduire en lui une trop grande confiance dans ce qui est observé. Comme le faisait remarquer de manière informelle un ingénieur rencontré, il est généralement hasardeux de présenter au final un résultat précis à la virgule près, en raison des imprécisions et erreurs relatives à la discrétisation du problème et aux équations utilisées lors des simulations. Souvent, une erreur de 10 à 20 % sur certaines grandeurs, voire un résultat qualitatif, sont tout à fait tolérables. Cela ne signifie pas que les représentations utilisées en visualisation puissent éliminer la précision présente dans les données, mais plutôt qu'elles doivent éviter de faire croire, même inconsciemment, que cette précision est réelle. De même que la tentation du réalisme de l'image semble à proscrire, l'esthétique ne devrait a priori pas être une qualité première recherchée en visualisation scientifique, du moins dans la phase d'exploration de données. Pour la présentation des résultats finaux, des qualités artistiques de l'image peuvent néanmoins aider à focaliser l'attention des destinataires sur le message principal, et le faire

passer plus efficacement.

Dans notre travail, nous nous sommes notamment appuyé sur la littérature relative à la perception de la transparence et de la forme par l'ombrage et par le mouvement pour évaluer et proposer des visualisations adaptées au regard des capacités et limites du système visuel humain. De nombreux autres aspects de la perception et de la cognition méritent également d'être connus et développés un peu plus dans le but de créer des solutions de visualisation scientifique plus performantes.

D'un point de vue perceptif, les connaissances relatives aux indices visuels facilitant la segmentation des objets qui composent une scène complexe peuvent aider à clarifier les images produites et à mieux isoler les phénomènes représentés. Par ailleurs, nous avons jusqu'à présent concentré notre attention sur le sens de la vision, mais il est envisageable d'intégrer d'autres modalités aux outils de visualisation, comme l'audition ou la proprioception (interfaces haptiques ou pseudo-haptiques). Il est alors nécessaire de comprendre précisément quel type d'information peut être véhiculé par ces canaux sensoriels complémentaires de manière à ajouter du sens à la visualisation, afin de déterminer les modalités permettant une intégration multisensorielle constructive.

Il reste également à investiguer plus en détails certains aspects cognitifs importants tels que l'identification d'objets ou de structures, leur manipulation, et la navigation et l'orientation spatiale dans des scènes 3D. Les résultats expérimentaux accumulés doivent permettre de concevoir des interfaces de visualisation capables de fluidifier l'exploration des volumes de données importants issus de la simulation. Concernant la navigation, il s'agit d'un problème crucial dès lors que l'on visualise un objet relativement étendu ou compliqué sur un écran de taille réduite. On peut essayer de déterminer les amers visuels minimaux qui sont nécessaires à un repérage optimal de la position actuelle dans la scène et limitent dans le même temps les occultations de données visualisées. Il sera également utile d'identifier les modes de navigation les plus adaptés pour se déplacer virtuellement sans perdre la compréhension globale de la scène, et comprendre comment ceux-ci peuvent interférer avec les rotations mentales réalisées lors de l'exploration visuelle d'objets tridimensionnels. En ce sens, un dialogue avec les ingénieurs concernés peut aussi s'avérer utile, car leur connaissance souvent très pointue des structures géométriques sur lesquelles s'appuie la visualisation influence la manière dont ils parcourent une scène de visualisation.

Au final, l'objectif consiste à produire des systèmes de visualisation plus efficaces et plus facilement utilisables pour les différents utilisateurs de visualisations scientifiques. Cela passe nécessairement par une application et une intégration en un cadre cohérent de l'ensemble des connaissances relatives à la perception et à la cognition chez l'homme.

Nouvelles représentations en visualisation scientifique

Pour terminer cet exposé, nous présentons trois problèmes ouverts pour lesquels de nouveaux modes de représentation doivent encore être élaborés, et pour lesquels une collaboration entre les chercheurs en informatique graphique et en perception visuelle est vraisemblablement nécessaire.

Le premier problème, déjà mentionné dans ce manuscrit, concerne la représentation de données volumiques. Quel type de rendu peut-on appliquer afin de mieux percevoir les variations de structures occupant l'espace, alors que notre système visuel a évolué dans un environnement constitué de surfaces opaques, et semble mieux à même d'appréhender de tels objets ? Les représentations surfaciques constituent incontestablement un frein à la visualisation en révélant une information par trop limitée, et en générant de nombreuses occultations. D'un autre côté, les rendus volumiques, dans leur forme actuelle, semblent trop difficiles à exploiter correctement, en tout cas pour des données autres qu'anatomiques. Le réglage de la fonction de transfert est un exercice délicat, et les images produites peuvent être particulièrement ambiguës, sauf si la fonction de transfert génère un rendu

quasi surfacique. Trouver des représentations adaptées à la perception de structures volumiques demeure donc un défi à résoudre. Sans doute d'ailleurs n'existe-t-il pas de solution générale, mais des réponses peuvent être envisagées pour répondre à des besoins bien précis. Pour commencer, il est envisageable d'investiguer des représentations intermédiaires, destinées à faire percevoir les données comprises dans des tranches fines du volume. On peut dans ce cas chercher à déterminer quels modes de représentations sont les plus adaptés au regard des capacités et limites de notre système visuel. On peut imaginer d'utiliser un rendu dynamique, qui fasse apparaître successivement certaines parties des informations, le mouvement aidant à extraire des formes et variations spatiales. Dans ce cas, on doit alors se poser aussi la question du rôle et des limites de la mémoire à court terme, pour savoir quelles informations nous sommes capables de retenir sur une courte période de temps afin de les intégrer dans un percept, ou un concept, cohérents.

Certaines simulations permettent de calculer en plus des variables principales les incertitudes relatives aux valeurs trouvées. Cela permet par exemple de caractériser la sensibilité du résultat final aux variations de certains paramètres, et donc de fournir un indice de confiance important lorsqu'il s'agit de formuler des conclusions dans une étude. Cependant, visualiser ces incertitudes en même temps que la variable correspondante est un exercice délicat, des représentations classiques risquant de se masquer mutuellement. Il faut donc pour les incertitudes trouver un mode de représentation clair qui n'interfère pas avec celui choisi pour la variable correspondante, une information visuelle qui soit d'une certaine façon orthogonale à celle à laquelle elle doit être associée. Par exemple, l'utilisation de flou dont l'extension varie avec l'ampleur de l'incertitude (dans un cas unidimensionnel) pourrait à première vue correspondre ; cependant, dans des régions où la variable étudiée est constante, toute information éventuelle de flou serait irrémédiablement non perceptible. Le problème est complexe, et nécessite de faire appel aux connaissances acquises sur la perception visuelle chez l'homme ; inversement, des expériences de psychophysique doivent permettre d'évaluer quantitativement les solutions proposées.

Enfin, le développement des couplages de codes et l'augmentation de la puissance de calcul permettent de plus en plus de réaliser des simulations comportant plusieurs physiques. Les domaines spatiaux correspondants peuvent être juxtaposés, par exemple dans le cas d'un couplage de mécanique des fluides et de mécanique des matériaux, ou bien coïncider. L'étude réalisée peut nécessiter de comprendre les relations entre les différents phénomènes simulés, ce qui exige de représenter conjointement plusieurs champs et de pouvoir établir des relations entre eux au moment de la visualisation. Cela augmente la quantité de données à afficher en même temps, et risque de poser des problèmes supplémentaires d'occultation. Là encore, de nouveaux modes de représentation adaptés à ce type particulier d'analyse doivent être trouvés. En raison de la spécificité importante de ce type d'approche, une collaboration étroite entre physiciens et spécialistes de l'informatique graphique et de la visualisation scientifique paraît indispensable.

Dès lors que l'explosion de la puissance de calcul permet de réaliser des simulations numériques de plus en plus fines et complexes, la visualisation scientifique doit sans cesse s'adapter et proposer de nouvelles méthodes adaptées à l'analyse et la compréhension des volumes de données générées. Des modes de représentation et d'interaction encore inexplorés doivent être imaginés. Pour inventer les visualisations de demain, il semble de plus en plus indispensable de faire travailler ensemble les spécialistes de domaines de recherche très différents. Plus que jamais, la visualisation scientifique se trouve au carrefour de multiples disciplines, de l'informatique graphique aux neurosciences de la perception visuelle, en passant par les sciences physiques concernées, voire par l'art visuel.

ANNEXE A

Calcul de la translation compensatoire dans l'expérience 3 sur le DVR

Dans cette expérience, la perspective linéaire utilisée introduit une différence de vitesse en valeur absolue entre les projections des points situés à l'arrière et à l'avant du volume. Il en résulte que la vitesse moyenne horizontale à l'écran est directement corrélée avec le sens de rotation, même si cet effet est partiellement atténué par la présence d'un masque qui occulte plus de points avants qu'arrières. Dès lors, les sujets pourraient répondre seulement sur la base de cette information 2D, et les résultats obtenus ne reflèteraient pas une réelle perception tridimensionnelle.

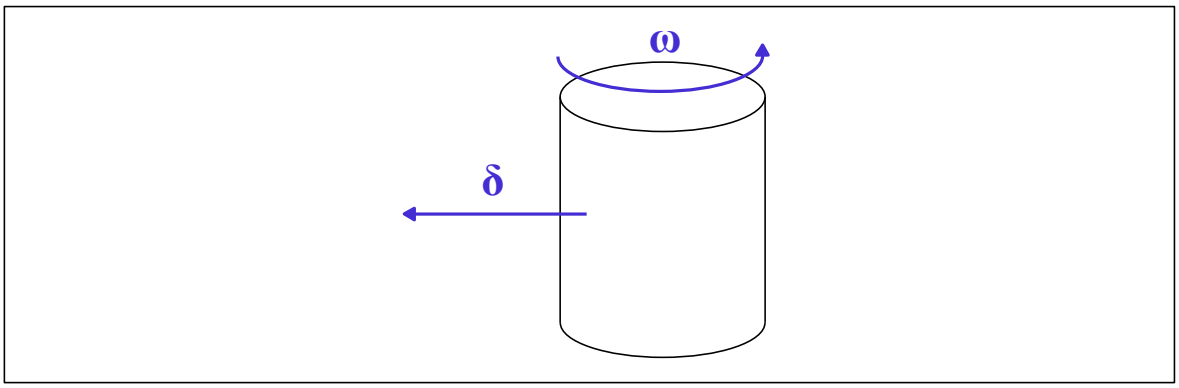


FIG. A.1: Dans l'expérience 3 sur la perception de la profondeur dans les rendus volumique, une session de contrôle est réalisée pour s'assurer que les participants répondent sur la base d'un percept 3D et non 2D. Une translation δ permet de compenser les différences de vitesses résultant de la rotation ω du cylindre.

Afin d'écarter cette hypothèse, nous avons implémenté des essais de contrôle dans le cas du rendu

par nuage de points volumique (CLOUD-VOL) en perspective amplifiée (P2). Dans ces essais, une translation horizontale du cylindre de vitesse δ s'oppose au biais mentionné précédemment, ce dernier résultant de la rotation de vitesse angulaire ω (voir figure A.1). Deux vitesses angulaires sont définies : dans la condition T1, δ_{T1} est choisie de manière à annuler la vitesse moyenne du nuage à l'écran : dans la condition T2, δ_{T2} permet d'égaliser en valeur absolue les vitesses maximales des projections de points arrières et avants.

Les notations utilisées ci-après sont illustrées sur la figure A.2. L'origine O du repère est fixée au centre du cylindre. Un point $M(x,y,z)$ du volume, de coordonnées cylindriques $M(\rho cos(\theta), y, \rho sin(\theta)$, est projeté sur le plan orthogonal à la direction de direction de vue centrale et passant par O en un point $Mp(X,Y,0)$. Le centre de projection C se trouve à une distance d du centre du cylindre, ce dernier étant de rayon R.

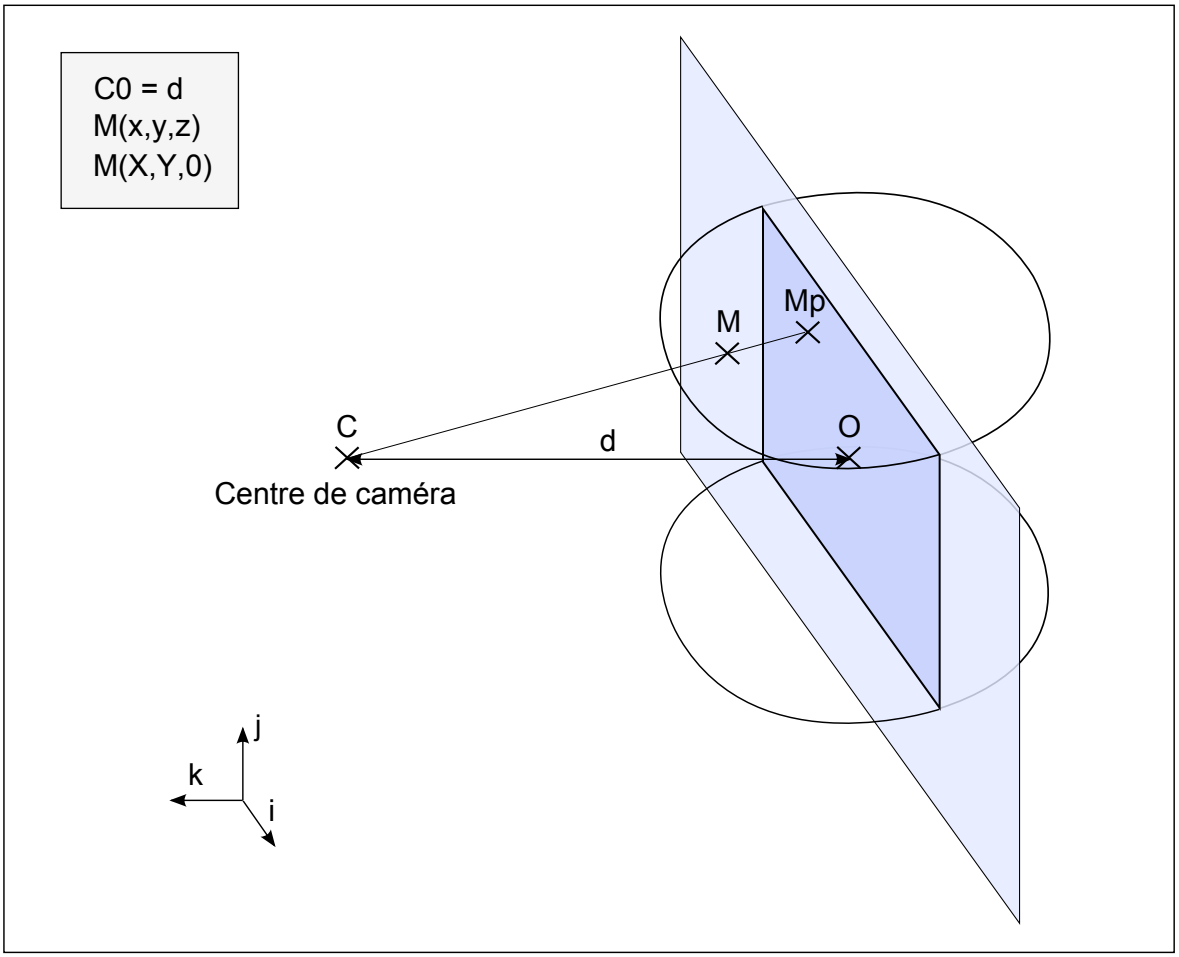


FIG. A.2: Notations utilisées pour calculer la translation compensant le différentiel de vitesse entre l'avant et l'arrière du volume dans le cas de projections perspectives.

Comme la translation horizontale et la rotation suivant l'axe du cylindre vertical n'induisent au-

cune vitesse verticale ($\dot{y} = 0$), on obtient les relations suivantes :

$$Mp \left\{ \begin{array}{lcl} X & = & x\frac{d}{d-z} \\ Y & = & y\frac{d}{d-z} \\ Z & = & 0 \end{array} \right. \quad \text{and} \quad \dot{M}p \left\{ \begin{array}{lcl} \dot{X} & = & \dot{x}\frac{d}{d-z} + x\dot{z}\frac{d}{(d-z)^2} \\ \dot{Y} & = & x\dot{z}\frac{d}{(d-z)^2} \\ \dot{Z} & = & 0 \end{array} \right. \tag{A.1}$$

Le mouvement moyen du nuage de points projeté est obtenu par une sommation pondérée de la rotation unitaire moyenne $\overline{\dot{X}_R}$ et de la translation unitaire moyenne $\overline{\dot{X}_T}$

$$\overline{\dot{X}} = \omega\,\overline{\dot{X}_R} \quad + \quad \delta\,\overline{\dot{X}_T}\,, \tag{A.2}$$

qui sont calculées en intégrant sur l'ensemble du volume du cylindre les vitesse projetées décrites plus haut :

$$\overline{\dot{X}_R} = \int_{y=-H/2}^{H/2}\int_{\theta=0}^{2\pi}\int_{\rho=0}^{R} -\rho\dot{\theta}\sin(\theta)\frac{d}{d-\rho\sin(\theta)} + \rho^2\dot{\theta}\cos(\theta)^2\frac{d}{(d-\rho\sin(\theta))^2}\,\mathrm{d}\rho\,\mathrm{d}\theta\,\mathrm{d}y \tag{A.3}$$

$$\overline{\dot{X}_T} = \int_{y=-H/2}^{H/2}\int_{\theta=0}^{2\pi}\int_{\rho=0}^{R} \dot{x}\frac{d}{d-\rho\sin(\theta)}\,\mathrm{d}\rho\,\mathrm{d}\theta\,\mathrm{d}y \tag{A.4}$$

Le calcul exact de ces intégrales n'est en pratique pas réalisable, d'autant plus qu'il nous faut nous limiter aux seuls points non cachés par le masque rectangulaire défini dans l'image. Par conséquent, la valeur de δ_{T1} est obtenue par une simulation de Monte Carlo réalisée avec 10 millions de points aléatoirement positionnés dans le cylindre et respectant la contrainte de visibilité induite par le masque. En revanche, la valeur de δ_{T2} peut être déduite directement de l'équation A.1.

TABLE DES FIGURES

1.1 Boucle simulation-visualisation 19
1.2 Exemples de maillages volumiques complexes 20
1.3 Exemples de représentations tridimensionnelles utilisées en visualisation scientifique 21
1.4 Visualisation par développée de surface 23
1.5 Primitives géométriques 3D utilisées pour discrétiser le domaine d'étude 24
1.6 Exemple d'évolution des tailles de simulation à EDF 25
1.7 Exemple d'utilisation du mur d'images installé sur le site EDF R&D de Clamart . . . 28
1.8 Exemple de difficulté liée à un problème d'échelles spatiales en visualisation 32
1.9 Importance du choix de l'échelle de couleurs 32
1.10 Difficulté de perception de l'espace dans une scène de visualisation complexe 34

2.1 Anatomie de l'œil 37
2.2 Répartition des capteurs rétiniens et acuité visuelle 38
2.3 Fonctions de sensibilité au contraste 38
2.4 Exemple d'exploration d'un tableau par saccades oculaires 39
2.5 Exemples d'affichages dépendant du regard (GCD) 41
2.6 Dépendance du trajet visuel à la tâche du sujet 42
2.7 Indices picturaux de perception de la forme et de la profondeur 47
2.8 Indices cinématiques de perception de la profondeur 50
2.9 Illusion de mouvement dans le dispositif de Ternus 54
2.10 Triangle de la visualisation scientifique 60

3.1 Exemples naturels d'objets transparents ou translucides 64
3.2 Principe du rendu volumique par lancer de rayon 66
3.3 Exemple de rendu volumique direct ordonné et de MIP 68
3.4 Modèle de transparence de Metelli 70
3.5 Transparence phénoménologique : importance des jonctions de contours 71
3.6 Inversions de contrastes au niveau des jonctions et ordre en profondeur 72
3.7 Stimulus utilisé pour étudier la perception statique de la profondeur en DVR 73
3.8 Perception de la profondeur en DVR, condition statique : résultats moyens 76
3.9 Perception de la profondeur en DVR, condition statique : réponses moyennes 77

3.10 Perception de la profondeur en DVR, condition statique : résultats moyens par couleur 78
3.11 Principaux profils de fixations oculaires 78
3.12 Evolution moyenne des performances au cours d'une session expérimentale 79
3.13 Evolutions individuelles des performances au cours d'une session expérimentale . . . 80
3.14 Effet d'apprentissage observé entre deux sessions 81
3.15 Proposition de modification de l'expérience d'évaluation du DVR en statique 82
3.16 Stimulus utilisé pour étudier la perception dynamique de la profondeur en DVR . . . 84
3.17 Choix des fonctions de transfert 85
3.18 Les six rendus sélectionnés pour le cylindre avec bruit de Perlin 86
3.19 Perception de la profondeur en DVR, condition dynamique : résultats moyens 87
3.20 Importance de la déformation perspective 90
3.21 Apport de l'information de perspective : configuration expérimentale 91
3.22 Influence de la projection pour la perception de la profondeur, classement des différents rendus et distributions par projection 93
3.23 Influence de la projection pour la perception de la profondeur, par condition de rendu 94
3.24 Influence du biais latéral sur les taux de réponses correctes 97
3.25 Informations contextuelles pour lever les ambiguïtés de profondeur 100

4.1 Différents effets de l'illumination 104
4.2 Eclairage de type Phong et Ambient Occlusion 109
4.3 Screen Space Ambient Occlusion : influence du voisinage et filtrage 111
4.4 Screen Space Ambient Occlusion et unsharp masking 111
4.5 EyeDome Lighting : définition du voisinage et calcul de l'ombrage 116
4.6 Résultat de la fonction d'ombrage de l'EDL et influence du calcul de l'élévation . . . 118
4.7 Architecture de rendu différé intégrant l'EDL 118
4.8 Approche multirésolution pour le calcul de l'EDL 119
4.9 Filtre gaussien contre filtre bilatéral croisé 121
4.10 Caractérisation 1D de l'ombrage produit par l'EDL 122
4.11 Comparaison de l'EDL avec les ombrages de type Phong et SSAO 123
4.12 Application de l'EDL à une visualisation par point sprites 124
4.13 Application de l'EDL à une visualisation de lignes de courant 125
4.14 Application de l'EDL à la visualisation d'un nuage de points 126
4.15 Hypothèse d'éclairage venant d'en haut 128
4.16 EDL directionnel (TorchDome Lighting) 130
4.17 Préservation des silhouettes et halos avec le TDL 131
4.18 EDL et focus en profondeur et en espace image 133
4.19 Modification du contraste avec le niveau de zoom 135
4.20 Comparaison de l'EDL avec les ombrages de type Phong et Ambient Occlusion en espace objet 136
4.21 Comparaison de l'EDL avec d'autres techniques d'ombrage 137

5.1 Exemples d'algorithmes produisant des vues en écorché 143
5.2 Partitionnement de l'espace en fonction du point de vue 144
5.3 Choix de la surface utilisée pour le calcul d'extrusion conique 145
5.4 Calcul de la surface d'extrusion conique en deux passes 146
5.5 Principe de l'algorithme d'extrusion conique en deux passes 147
5.6 Représentation de la surface d'extrusion conique 148
5.7 Choix de la norme pour l'extrusion conique 149

5.8 Réalisation du calcul d'extrusion à différentes résolutions 150
5.9 Architecture complète de rendu avec partitionnement et extrusion conique 152
5.10 Vues fantômes du contexte occultant et perception de la distance au contexte arrière . 153
5.11 Exemple d'extrusion conique par isosurface non connexe 155
5.12 Exemple d'extrusion conique par cube d'intérêt avec rendus variés 156
5.13 Exemple d'extrusion conique par isosurface et lignes de courant 156
5.14 Limites de l'algorithme lors du creusement de multiples surfaces 158

A.1 Translation du cylindre compensant les différences de vitesses. 167
A.2 Notations utilisées pour calculer la translation compensatoire. 168

LISTE DES TABLEAUX

3.1 Intensité moyenne et contraste RMS dans les différents rendus volumiques. 88

4.1 EDL : performances. 127

5.1 Performances de l'algorithme d'extrusion conique sur GPU. 150

TABLE DES MATIÈRES

Remerciements **5**

Résumé **7**

Abstract **9**

Liste des abréviations **11**

Introduction **13**
Contexte et motivations 15
Organisation du document 16

1 La visualisation scientifique à EDF **17**
1 Simulation et visualisation 18
1.1 Création du maillage 18
1.2 Le temps de la simulation 19
1.3 Visualisation des résultats 20
1.4 La prise de décision 22
1.5 Conclusion : place de la visualisation 22
2 Domaines et échelles 22
2.1 Domaines physiques 22
2.2 Echelles spatiales 23
2.3 Primitives géométriques 23
2.4 Grands volumes de données 24
3 Moyens matériels et logiciels 25
3.1 Codes de calcul 25
3.2 Calculateurs 26
3.3 Postes de visualisation 27
3.4 Logiciels de visualisation 27
4 Pratiques et difficultés 28

4.1 Contexte des entrevues . . . 29
4.2 Perception dans le volume . . . 29
4.3 Perception de la profondeur et de l'orientation . . . 30
4.4 Complexité de l'affichage . . . 30
4.5 Interactivité, latence . . . 31
4.6 Problèmes d'échelle . . . 31
5 Conclusion et objectifs . . . 33

2 Perception visuelle et informatique graphique 35
1 Rendre plus vite . . . 36
1.1 Limites bas-niveau . . . 36
1.1.1 Quelques éléments de vision . . . 36
1.1.2 CSF et masquage en informatique graphique . . . 38
1.1.3 Affichages fonction de l'excentricité . . . 40
1.2 Le problème de l'attention . . . 41
1.2.1 Introduction à l'attention visuelle . . . 41
1.2.2 Exploration visuelle . . . 42
1.2.3 Cécité attentionnelle . . . 43
1.2.4 Saillance de l'information visuelle . . . 44
1.2.5 Limites des modèles de saillance : le regard de l'expert . . . 45
2 Rendre mieux . . . 46
2.1 La perception de l'espace . . . 46
2.1.1 Forme, profondeur et volume . . . 46
2.1.2 Indices picturaux . . . 46
2.1.3 Indices oculomoteurs et binoculaires . . . 48
2.1.4 Indices cinématiques . . . 49
2.1.5 Indices cognitifs . . . 49
2.1.6 Combinaison d'indices multiples . . . 50
2.1.7 Perception du volume . . . 51
2.2 L'interactivité en visualisation scientifique . . . 52
2.3 Interactivité et temps réel . . . 52
2.4 Notions sur la perception du mouvement . . . 53
2.5 Interactivité requise pour les algorithmes de visualisation scientifique . . . 55
3 Place de la perception en visualisation scientifique . . . 55
3.1 Première limite : la complexité du stimulus . . . 55
3.2 Deuxième limite : l'expertise du sujet . . . 56
3.3 Troisième limite : la nature de la tâche expérimentale . . . 57
3.4 Conclusion . . . 58
4 Perspectives, orientation du travail . . . 59

3 Evaluation des rendus volumiques 63
1 Les techniques de rendu volumique direct . . . 65
1.1 Principe . . . 65
1.2 Choix de la fonction de transfert . . . 66
1.3 Rendus volumiques non ordonnés . . . 67
1.4 Améliorations du rendu . . . 67
1.5 Evaluation . . . 68
2 Perception de la transparence . . . 69

2.1 Conditions d'occurence . . . 69
2.2 Importance des jonctions . . . 70
2.3 Transparence et ordre en profondeur . . . 71
3 Expérience 1 : perception statique en DVR . . . 72
3.1 Dispositif expérimental . . . 73
3.1.1 Description du stimulus . . . 73
3.1.2 Déroulement de l'expérience . . . 74
3.1.3 Participants . . . 75
3.2 Résultats . . . 75
3.2.1 Performances des sujets . . . 75
3.2.2 Analyse des fixations . . . 76
3.2.3 Rôle éventuel de l'apprentissage . . . 77
3.3 Discussion . . . 79
3.3.1 Résultat principal . . . 79
3.3.2 Du plus gros au plus proche . . . 80
3.3.3 Temps de présentation et de réponse . . . 82
4 Expérience 2 : DVR et effet cinétique de profondeur . . . 83
4.1 Protocole expérimental . . . 83
4.1.1 Description du stimulus . . . 83
4.1.2 Déroulement de l'expérience . . . 85
4.2 Résultats . . . 86
4.3 Discussion . . . 87
4.3.1 Importante de la fonction de transfert . . . 87
4.3.2 Influence du contraste . . . 88
5 Expérience 3 : apport de la perspective linéaire . . . 89
5.1 Protocole expérimental . . . 89
5.1.1 Modification du stimulus de l'expérience 2 . . . 89
5.1.2 Contrôle de la présence d'artefacts 2D dans les réponses . . . 90
5.1.3 Déroulement de l'expérience . . . 92
5.1.4 Participants . . . 92
5.2 Résultats . . . 93
5.2.1 Session principale . . . 93
5.2.2 Sessions de contrôle . . . 95
5.3 Discussion . . . 95
5.3.1 Résultat principal . . . 95
5.3.2 Une perception 2D ou 3D ? . . . 96
5.3.3 Biais latéral . . . 96
6 Conclusion et perspectives . . . 97
6.1 Discussion des résultats . . . 97
6.1.1 Rendu volumique et perception de l'espace . . . 97
6.1.2 Contrastes et couleurs . . . 98
6.2 Améliorer l'exploration de données volumiques . . . 98
6.2.1 Faciliter la perception de profondeur par le mouvement . . . 98
6.2.2 Trouver de nouveaux rendus perceptivement adaptés . . . 99

4 EyeDome Lighting **101**
1 Perception de l'éclairage et modèles infographiques . . . 102
1.1 Eclairage et perception de la forme . . . 102
1.1.1 Biais de convexité . . . 104
1.1.2 Conditions d'éclairage . . . 105
1.1.3 Perception et importance de la direction d'éclairage . . . 106
1.1.4 Importance des reflets . . . 106
1.1.5 Ombres portées . . . 107
1.2 Eclairage en informatique graphique . . . 108
1.2.1 Méthodes d'éclairage global . . . 108
1.2.2 Méthodes d'éclairage local . . . 109
1.2.3 Eclairage ambiant . . . 109
1.2.4 Autres algorithmes . . . 111
1.2.5 Modèles d'ombrage non photoréalistes . . . 112
1.3 Eclairage en visualisation scientifique . . . 114
1.4 Conclusion : caractéristiques importantes pour un éclairage . . . 114
2 Algorithme de l'EyeDome Lighting . . . 115
2.1 Aperçu . . . 115
2.2 Algorithme de base . . . 116
2.2.1 Notations . . . 116
2.2.2 Calcul de la valeur d'ombrage . . . 116
2.2.3 Implémentation . . . 117
2.3 Approche multirésolution . . . 117
2.3.1 Principe . . . 117
2.3.2 Lissage par filtre bilatéral croisé . . . 119
2.4 Résultats . . . 121
2.4.1 Caractéristiques de l'EyeDome Lighting . . . 121
2.4.2 Comparaison de l'EDL avec des ombrages de type Phong et ambient occlusion . . . 122
2.4.3 Exemples particuliers . . . 124
2.4.4 Performances : avantages et inconvénients . . . 125
3 Extension de l'EyeDome Lighting . . . 127
3.1 EDL et lumière directionnelle . . . 127
3.1.1 Justifications perceptives : hypothèse de l'éclairage par le haut et intérêt d'un éclairage dynamique pour lever les ambiguïtés . . . 128
3.1.2 Lumière directionnelle : le TorchDome Lighting (TDL) . . . 129
3.1.3 Résultats . . . 130
3.1.4 Remarques et critiques . . . 131
3.2 Focus stylisé . . . 131
4 Conclusion et perspectives . . . 132
4.1 Limites de l'EDL . . . 134
4.1.1 Comparaisons avec d'autres méthodes d'ombrage . . . 135
4.2 Application à la visualisation chez EDF . . . 136
4.3 Retour à la Perception . . . 136

5 Focus+Contexte interactif avec extrusion conique **139**
1 Techniques variées de Focus+Contexte . . . 140
1.1 Principes généraux . . . 140
1.2 Multirésolutions . . . 140
1.3 Focus centré sur les données . . . 141
1.4 Déformations et vues éclatées . . . 141
1.5 Découpes et écorchés . . . 142
1.6 Objectifs . . . 142
2 Algorithme d'extrusion conique sur le GPU . . . 144
2.1 Partition de l'espace . . . 144
2.1.1 Définitions et notations . . . 144
2.1.2 Choix de la surface arrière du focus . . . 144
2.2 Creusement de l'image de profondeur . . . 145
2.2.1 Formulation . . . 145
2.2.2 Optimisation en deux passes . . . 146
2.2.3 Choix de la norme . . . 148
2.3 Performances . . . 149
3 Rendu complet en écorché . . . 151
3.1 Architecture générale . . . 151
3.2 Gestion des occultations et distance à l'arrière plan . . . 151
3.2.1 Contexte occultant . . . 154
3.2.2 Contexte arrière . . . 154
3.3 Quelques exemples . . . 154
3.4 Performances et limites . . . 156
4 Discussion et perspectives . . . 157
4.1 Ecorché par découpe pseudo-conique . . . 157
4.2 Perception de la profondeur . . . 158
4.3 Surfaces multiples . . . 158

Conclusion et perspectives **161**
Contributions . . . 161
Perspectives . . . 162

A Calcul de la translation compensatoire **167**

Table des figures **171**

Table des matières **175**

Bibliographie **181**

BIBLIOGRAPHIE

[AA90] E. H. Adelson and P. Anandan. Ordinal characteristics of transparency. In *AAAI-90 Workshop on Qualitative Vision*, 1990. 71, 72

[AAB+84] E.H. Adelson, C.H. Anderson, J.R. Bergen, P.J. Burt, and J.M. Ogden. Pyramid methods in image processing. *RCA Engineer*, 29(6), 1984. 119

[AC99] R. R. Althoff and N. J. Cohen. Eye-movement-based memory effect : a reprocessing effect in face perception. *J Exp Psychol Learn Mem Cogn*, 25(4) :997–1010, Jul 1999. 45

[AGE04] W J Adams, E W Graf, and M O Ernst. Experience can change the 'light-from-above' prior. *Nat Neurosci*, 7(10) :1057–1058, Oct 2004. 50

[ALK+03] D. Akers, F. Losasso, J. Klingner, M. Agrawala, J. Rick, and P. Hanrahan. Conveying shape and features with image-based relighting. In *VIS '03 : Proceedings of the 14th IEEE Visualization 2003 (VIS'03)*, page 46, Washington, DC, USA, 2003. IEEE Computer Society. 113

[AM08] R. Allen and P. McGeorge. Enumeration : Shape information and expertise. *Acta Psychologica*, 129(1) :26–31, 2008. 56

[And97] B. L. Anderson. A theory of illusory lightness and transparency in monocular and binocular images : the role of contour junctions. *Perception*, 26(4) :419–453, 1997. 69, 70

[Ans74] S. M. Anstis. Letter : A chart demonstrating variations in acuity with retinal position. *Vision Res*, 14(7) :589–592, Jul 1974. 38

[Ara03] Daniel Arasse. *Léonard de Vinci*. Hazan, 2003. 13, 47

[ARHM00] M. Agrawala, R. Ramamoorthi, A. Heirich, and L. Moll. Efficient image-based methods for rendering soft shadows. In *SIGGRAPH '00 : Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, pages 375–384, New York, NY, USA, 2000. ACM Press/Addison-Wesley Publishing Co. 112

[BB85] C. L. Baker and O. J. Braddick. Temporal properties of the short-range process in apparent motion. *Perception*, 14(2) :181–192, 1985. 54

[BB07] D. C. Banks and K. Beason. Fast global illumination for visualizing isosurfaces with a 3d illumination grid. *Computing in Science and Engineering*, 9(1) :48–54, 2007. 114

[BBS98] I. Bülthoff, H. Bülthoff, and P. Sinha. Top-down influences on stereoscopic depth-perception. *Nat Neurosci*, 1(3) :254–257, Jul 1998. 50

[BCL+07] L. Bavoil, S. P. Callahan, A. Lefohn, J. L. D. Comba, and C. T. Silva. Multi-fragment effects on the gpu using the k-buffer. In *I3D '07 : Proceedings of the 2007 symposium on Interactive 3D graphics and games*, pages 97–104, New York, NY, USA, 2007. ACM. 159

[BCS01] R. Borgo, P. Cignoni, and R. Scopigno. An easy-to-use visualization system for huge cultural heritage meshes. In *VAST '01 : Proceedings of the 2001 conference on Virtual reality, archeology, and cultural heritage*, pages 121–130, New York, NY, USA, 2001. ACM. 110

[BG06] S. Bruckner and E. Gröller. Exploded views for volume data. In *IEEE Transactions on Visualization and Computer Graphics*, pages 1077–1084, 2006. 141

[BG07] S. Bruckner and E. Gröller. Enhancing depth-perception with flexible volumetric halos. *IEEE Transactions on Visualization and Computer Graphics*, 13(6) :1344–1351, 2007. 67, 99, 114

[BH07] A. Bair and D. House. Grid with a view : Optimal texturing for perception of layered surface shape. *IEEE Transactions on Visualization and Computer Graphics*, 13(6) :1656–1663, 2007. 52

[BHW+07] M. Burns, M. Haidacher, W. Wein, I. Viola, and E. Gröller. Feature emphasis and contextual cutaways for multimodal medical visualization. In *Eurographics/IEEE-VGTC Symposium on Visualization*, pages 275–282, 2007. 142

[BKK07] B. Battu, A. M. Kappers, and J. J. Koenderink. Ambiguity in pictorial depth. *Perception*, 36(9) :1290–1304, 2007. 48

[BKR+05] Michael Burns, Janek Klawe, Szymon Rusinkiewicz, Adam Finkelstein, and Doug DeCarlo. Line drawings from volume data. *ACM Trans. Graph.*, 24(3) :512–518, 2005. 58, 100

[BKY97] P. N. Belhumeur, D. J. Kriegman, and A. L. Yuille. The bas-relief ambiguity. In *CVPR '97 : Proceedings of the 1997 Conference on Computer Vision and Pattern Recognition (CVPR '97)*, page 1060, Washington, DC, USA, 1997. IEEE Computer Society. 101, 103, 106

[Bli77] J. F. Blinn. Models of light reflection for computer synthesized pictures. *ACM SIGGRAPH Computer Graphics*, 11(2) :192–198, 1977. 109

[BM98] M. R. Bolin and G. W. Meyer. A perceptually based adaptive sampling algorithm. In *SIGGRAPH '98 : Proceedings of the 25th annual conference on Computer graphics and interactive techniques*, pages 299–309, New York, NY, USA, 1998. ACM. 38

[BM08] L. Bavoil and K. Myers. Order independent transparency with dual depth peeling. in NVIDIA OpenGL SDK Guide, 2008. 146

[Bor42] E. G. Boring. *Sensation and Perception in the History of Experimental Psychology*. New York : Appleton-Century, 1942. 46

[BPP05] J. Burge, M. A. Peterson, and S. E. Palmer. Ordinal configural cues combine with metric disparity in depth perception. *J Vis*, 5(6) :534–542, 6 2005. 50

[Bra74] O. Braddick. A short-range process in apparent motion. *Vision Res*, 14(7) :519–527, Jul 1974. 53

[Bry05] S. Bryson. *The Visualization Handbook*, chapter Direct manipulation in virtual reality, pages 413–430. Elsevier Academic Pr., 2005. 53

[BSP+94] E. A. Bier, M. C. Stone, K. Pier, K. Fishkin, T. Baudel, M. Conway, W. Buxton, and T. DeRose. Toolglass and magic lenses : the see-through interface. In *CHI '94 : Conference companion on Human factors in computing systems*, pages 445–446. ACM, 1994. 140

[BTM06] P. Barla, J. Thollot, and L. Markosian. X-toon : An extended toon shader. In *International Symposium on Non-Photorealistic Animation and Rendering (NPAR)*. ACM, 2006. 112

[Bun05] M. Bunnell. *GPU GEMS 2*, chapter Dynamic Ambient Occlusion and Indirect Lighting. Addison-Wesley Professional, 2005. 110

[Cas08] R. Casati. The copycat solution to the shadow correspondence problem. *Perception*, 37(4) :495–503, 2008. 107

[Cav04] P. Cavanagh. *Cognitive Neuroscience of Attention*, chapter Attention Routines and the Architecture of Selection, pages 13–28. Guilford Press, 2004. 41

[Cav05] P. Cavanagh. The artist as neuroscientist. *Nature*, 434(7031) :301–307, Mar 2005. 58, 106, 107, 112

[Cav06] P. Cavanagh. Les licences de fra carnevale. *Pour la Science*, 16 :637–643, 2006. 106

[CCL02] K. Cater, A. Chalmers, and P. Ledda. Selective quality rendering by exploiting human inattentional blindness : Looking but not seeing. In *Symposium on Virtual Reality Software and Technology 2002*, pages 17–24. ACM, November 2002. 43

[CDF+06] F. Cole, D. DeCarlo, A. Finkelstein, K. Kin, K. Morley, and A. Santella. Directing gaze in 3D models with stylized focus. *Eurographics Symposium on Rendering*, pages 377–387, June 2006. 45, 131

[CF07] F. Caniard and R. W. Fleming. Distortion in 3d shape estimation with changes in illumination. In *APGV '07 : Proceedings of the 4th symposium on Applied perception in graphics and visualization*, pages 99–105, New York, NY, USA, 2007. ACM. 103, 105, 128

[CH06] C. Coffin and T. Höllerer. Interactive perspective cut-away views for general 3d scenes. In *VR '06 : Proceedings of the IEEE conference on Virtual Reality*, page 118, 2006. 142

[CL93] B. Cabral and L. C. Leedom. Imaging vector fields using line integral convolution. In *SIGGRAPH '93 : Proceedings of the 20th annual conference on Computer graphics and interactive techniques*, pages 263–270, New York, NY, USA, 1993. ACM. 52

[Cla76] J. H. Clark. Hierarchical geometric models for visible surface algorithms. *Communications of the ACM*, 19(10) :547–554, 1976. 39

[CMS94] P. Cignoni, C. Montani, and R. Scopigno. Magicsphere : an insight tool for 3d data visualization. *Computer Graphics Forum*, 13(3) :317–328, 1994. 140, 141

[CSC07] C. Correa, D. Silver, and M. Chen. Illustrative deformation for data exploration. *IEEE Transactions on Visualization and Computer Graphics*, 13(6) :1320–1327, 2007. 141

[CST05] P. Cignoni, R. Scopigno, and M. Tarini. A simple normal enhancement technique for interactive non-photorealistic renderings. *Computer & Graphics*, 29(1) :125–133, feb 2005. 113

[CT82] R. L. Cook and K. E. Torrance. A reflectance model for computer graphics. *ACM Transactions on Graphics*, 1(1) :7–24, 1982. 109

[CV95] J. E. Cutting and P. M. Vishton. *Perception of Space and Motion*, chapter Perceiving layout and knowing distances : The integration, relative potency, and contextual use of different information about depth, pages 69–117. Academic Pr., 1995. 46

[Dal93] S. Daly. The visible differences predictor : an algorithm for the assessment of image fidelity. In *Digital images and human vision*, pages 179–206, Cambridge, MA, USA, 1993. MIT Press. 38

[DBD+07] G. Drettakis, N. Bonneel, C. Dachsbacher, S. Lefebvre, M. Schwarz, and I. Viaud-Delmon. An interactive perceptual rendering pipeline using contrast and spatial masking. In *Eurographics Symposium on Rendering*. Eurographics, June 2007. 40

[DBS08] R. Dimitrov, L. Bavoil, and M. Sainz. Horizon-split ambient occlusion. In *I3D '08 : Proceedings of the 2008 symposium on Interactive 3D graphics and games*, pages 1–1, New York, NY, USA, 2008. ACM. 111

[DCH88] R. A. Drebin, L. Carpenter, and P. Hanrahan. Volume rendering. In *SIGGRAPH '88 : Proceedings of the 15th annual conference on Computer graphics and interactive techniques*, pages 65–74, New York, NY, USA, 1988. ACM. 52, 64, 65

[DCM04] A. T. Duchowski, N. Cournia, and H. Murphy. Gaze-contingent displays : a review. *Cyberpsychol Behav*, 7(6) :621–634, Dec 2004. 40

[DD90] R. L. DeValois and K. K. DeValois. *Spatial Vision*. Oxford University Press, 1990. 38

[DD02] F. Durand and J. Dorsey. Fast bilateral filtering for the display of high-dynamic-range images. In *SIGGRAPH '02 : Proceedings of the 29th annual conference on Computer graphics and interactive techniques*, pages 257–266, New York, NY, USA, 2002. ACM. 120

[DDGM+04] F. Duguet, G. Drettakis, D. Girardeau-Montaut, J. Martinez, and F. Schmitt. A point-based approach for capture, display and illustration of very complex archeological artefacts. In K. Cain, Y. Chrysanthou, F. Niccolucci, and N. Silberman (Editors), editors, *Proceedings of the International Symposium on Virtual Reality, Archeology and Cultural Heritage*. Eurographics, December 2004. 110

[Dec96] P. Decaudin. Cartoon looking rendering of 3D scenes. Research Report 2919, INRIA, June 1996. 112, 113

[DGH03] H. Doleisch, M. Gasser, and H. Hauser. Interactive feature specification for focus+context visualization of complex simulation data. In *VISSYM '03 : Proceedings of the symposium on Data visualisation 2003*, pages 239–248, Aire-la-Ville, Switzerland, Switzerland, 2003. Eurographics Association. 141

[DGM04] F. Duguet and D. Girardeau-Montaut. Rendu en portion de ciel visible de gros nuages de points 3d. In *Actes des journees de l'AFIG*, 2004. 110

[DS02] D. DeCarlo and A. Santella. Stylization and abstraction of photographs. In *SIGGRAPH '02 : Proceedings of the 29th annual conference on Computer graphics and interactive techniques*, pages 769–776, New York, NY, USA, 2002. ACM. 44, 131

[DSDD07] C. Dachsbacher, M. Stamminger, G. Drettakis, and F. Durand. Implicit visibility and antiradiance for interactive global illumination. *ACM Transactions on Graphics*, 26(3) :61, 2007. 109

[Dur02] F. Durand. An invitation to discuss computer depiction. In *NPAR '02 : Proceedings of the 2nd international symposium on Non-photorealistic animation and rendering*, pages 111–124, New York, NY, USA, 2002. ACM. 163

[DUS91] M. Dick, S. Ullman, and D. Sagi. Short- and long-range processes in structure-from-motion. *Vision Res*, 31(11) :2025–2028, 1991. 54

[DWE03] J. Diepstraten, D. Weiskopf, and T. Ertl. Interactive cutaway illustrations. In *Proccee-dings of EUROGRAPHICS '03*, pages 523–532, 2003. 142

[DWS+88] M. Deering, S. Winner, B. Schediwy, C. Duffy, and N. Hunt. The triangle processor and normal vector shader : a vlsi system for high performance graphics. In *SIGGRAPH '88 : Proceedings of the 15th annual conference on Computer graphics and interactive techniques*, pages 21–30, New York, NY, USA, 1988. ACM. 108, 110, 115

[ED04] E. Eisemann and F. Durand. Flash photography enhancement via intrinsic relighting. *ACM Transactions on Graphics*, 23(3) :673–678, 2004. 120

[EH99] R. A. Eagle and M. A. Hogervorst. The role of perspective information in the recovery of 3d structure-from-motion. *Vision Res*, 39(9) :1713–1722, May 1999. 89

[EKE01] K. Engel, M. Kraus, and T. Ertl. High-quality pre-integrated volume rendering using hardware-accelerated pixel shading. In *HWWS '01 : Proceedings of the ACM SIGGRAPH/EUROGRAPHICS workshop on Graphics hardware*, pages 9–16, New York, NY, USA, 2001. ACM Press. 66

[EKK93] R. G. Erens, A. M. Kappers, and J. J. Koenderink. Perception of local shape from shading. *Percept Psychophys*, 54(2) :145–156, Aug 1993. 104, 107

[ER90] J. T. Enns and R. A. Rensink. Influence of scene-based properties on visual search. *Science*, 247(4943) :721–723, Feb 1990. 106

[FA03] R. W. Fleming and B. L. Anderson. *The Visual Neurosciences*, chapter The perceptual organization of depth, pages 1284–1299. Cambridge, MA : MIT Press, 2003. 46, 51, 69, 72

[FAR07] R. Fattal, M. Agrawala, and S. Rusinkiewicz. Multiscale shape and detail enhancement from multi-light image collections. *ACM Transactions on Graphics*, 26(3) :51, 2007. 113

[FD06] E. D. Freeman and J. Driver. Subjective appearance of ambiguous structure-from-motion can be driven by objective switches of a separate less ambiguous context. *Vision Res*, 46(23) :4007–4023, Nov 2006. 99

[Fer01] J. A. Ferwerda. Elements of early vision for computer graphics. *IEEE Computer Graphics and Applications*, 21(5) :22–33, September/October 2001. pdf paper read interesting paper that gives a good quick understanding about human vision. 37

[FG98] Anton Fuhrmann and Eduard Gröller. Real-time techniques for 3d flow visualization. In *VIS '98 : Proceedings of the conference on Visualization '98*, pages 305–312, Los Alamitos, CA, USA, 1998. IEEE Computer Society Press. 140, 141

[FH04] F. Fang and S. He. Stabilized structure from motion without disparity induces disparity adaptation. *Curr Biol*, 14(3) :247–251, Feb 2004. 99

[FS93] T. A. Funkhouser and C. H. Séquin. Adaptive display algorithm for interactive frame rates during visualization of complex virtual environments. In *SIGGRAPH '93 : Proceedings of the 20th annual conference on Computer graphics and interactive techniques*, pages 247–254, New York, NY, USA, 1993. ACM. 39, 149

[FSM06] J. M. Fulvio, M. Singh, and L. T. Maloney. Combining achromatic and chromatic cues to transparency. *J Vis*, 6 :760–776, 2006. 69

[FTA04] R. W. Fleming, A. Torralba, and E. H. Adelson. Specular reflections and the perception of shape. *J Vis*, 4(9) :798–820, Sep 2004. 107

[FWSP04] J. A. Ferwerda, S. H. Westin, R. C. Smith, and R. Pawlicki. Effects of rendering on shape perception in automobile design. In *APGV '04 : Proceedings of the 1st Symposium on Applied perception in graphics and visualization*, pages 107–114, New York, NY, USA, 2004. ACM. 114

[Gas04] M. Gasser. Fast focus+context visualization of large scientific data. In *CESCG 2004 : 8th Central European Seminar on Computer Graphics*, 2004. 141

[GB06] T. Grossman and R. Balakrishnan. An evaluation of depth perception on volumetric displays. In *AVI '06 : Proceedings of the working conference on Advanced visual interfaces*, pages 193–200, New York, NY, USA, 2006. ACM. 46

[GdMM07] P. Gerardin, M. de Montalembert, and P. Mamassian. Shape from shading : new perspectives from the polo mint stimulus. *J Vis*, 7(11) :1–11, 2007. 129

[Get96] I. Getz. *L'expertise cognitive aux échecs*. Presses Universitaires de France, 1996. 45, 56

[GGLE04] K. Gaither, B. Geisler, D. H. Laidlaw, and D. Ebert. Panel : In the eye of the beholder : The role of perception in scientific visualization. In *VIS'04 Proceedings of IEEE Visualization Conference*, pages 567–568, October 10 - 15 2004. 55

[GGSC98] A. Gooch, B. Gooch, P. Shirley, and E. Cohen. A non-photorealistic lighting model for automatic technical illustration. In *SIGGRAPH '98 : Proceedings of the 25th annual conference on Computer graphics and interactive techniques*, pages 447–452, New York, NY, USA, 1998. ACM. 112

[GH90] M. A. Georgeson and M. G. Harris. The temporal range of motion sensing and motion perception. *Vision Res*, 30(4) :615–619, 1990. 54

[GH97] M. Garland and P. S. Heckbert. Surface simplification using quadric error metrics. In *SIGGRAPH '97 : Proceedings of the 24th annual conference on Computer graphics and interactive techniques*, pages 209–216, New York, NY, USA, 1997. ACM Press/Addison-Wesley Publishing Co. 44

[GLC+01] F. Gobet, P. C. Lane, S. Croker, P. C. Cheng, G. Jones, I. Oliver, and J. M. Pine. Chunking mechanisms in human learning. *Trends Cogn Sci*, 5(6) :236–243, Jun 2001. 57

[GMS+07] J. Giesen, K. Mueller, E. Schuberth, L. Wang, and P. Zolliker. Conjoint analysis to measure the perceived quality in volume rendering. *IEEE Transactions on Visualization and Computer Graphics*, 13(6) :1664–1671, 2007. 68

[Gom02] E. H. Gombrich. *L'art et l'illusion (Psychologie de la représentation picturale)*. Phaidon Press Limited, 2002. 112

[Gou71] H. Gouraud. Continuous shading of curved surfaces. *IEEE Transactions on Computers*, 20(6) :623–629, 1971. 102, 109

[GP06] C. P. Gribble and S. G. Parker. Enhancing interactive particle visualization with advanced shading models. In *APGV '06 : Proceedings of the 3rd symposium on Applied perception in graphics and visualization*, pages 111–118, New York, NY, USA, 2006. ACM. 114, 124

[Gre61] B. F. Green. Figure coherence in the kinetic depth effect. *J Exp Psychol*, 62 :272–282, Sep 1961. 49, 83, 89

[GSTdW90] W. Gerbino, C. I. Stultiens, J. M. Troost, and C. M. de Weert. Transparent layer constancy. *J Exp Psychol Hum Percept Perform*, 16(1) :3–20, Feb 1990. 69

[GTGB84] C. M. Goral, K. E. Torrance, D. P. Greenberg, and B. Battaile. Modeling the interaction of light between diffuse surfaces. In *SIGGRAPH '84 : Proceedings of the 11th annual conference on Computer graphics and interactive techniques*, pages 213–222, New York, NY, USA, 1984. ACM. 108

[HB94] H. Hill and V. Bruce. A comparison between the hollow-face and 'hollow-potato' illusions. *Perception*, 23(11) :1335–1337, 1994. 104

[HBW06] D. H. House, A. S. Bair, and C. Ware. An approach to the perceptual optimization of complex visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 12(4) :509–521, 2006. 52

[HHO05] S. Howlett, J. Hamill, and C. O'Sullivan. Predicting and evaluating saliency for simplified polygonal models. *ACM Transactions on Applied Perception*, 2(3) :286–308, 2005. 44

[Hib00] B. Hibbard. Confessions of a visualization skeptic. *Computer Graphics*, 34(3) :11–13, August 2000. 68

[HM03] M. Halle and J. Meng. Lightkit : A lighting system for effective visualization. In *VIS '03 : Proceedings of the 14th IEEE Visualization 2003 (VIS'03)*, page 48, Washington, DC, USA, 2003. IEEE Computer Society. 129

[Hop96] H. Hoppe. Progressive meshes. In *SIGGRAPH '96 : Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, pages 99–108, New York, NY, USA, 1996. ACM. 39

[HPAD06] K. Hegeman, S. Premože, M. Ashikhmin, and G. Drettakis. Approximate ambient occlusion for trees. In *I3D '06 : Proceedings of the 2006 symposium on Interactive 3D graphics and games*, pages 87–92, New York, NY, USA, 2006. ACM. 110

[HR05] S. Haller and E. W. Radue. What is different about a radiologist's brain ? *Radiology*, 236(3) :983–989, 2005. 56

[HWH99] J. M. Henderson, P. A. Jr. Weeks, and A. Hollingworth. The effects of semantic consistency on eye movements during complex scene viewing. *J Exp Psychol Hum Percept Perform*, 25(1) :210–228, 1999. 44, 45

[HWLB04] J. M. Hillis, S. J. Watt, M. S. Landy, and M. S. Banks. Slant from texture and disparity cues : optimal cue combination. *J Vis*, 4(12) :967–992, Dec 2004. 50

[HWSB99] G. S. Hubona, P. N. Wheeler, G. W. Shirah, and M. Brandt. The relative contributions of stereo, lighting, and background scenes in promoting 3d depth visualization. *ACM Transactions on Computer-Human Interaction*, 6(3) :214–242, 1999. 51, 107

[IFP97] V. Interrante, H. Fuchs, and S. M. Pizer. Conveying the 3D shape of smoothly curving transparent surfaces via texture. *IEEE Transactions on Visualization and Computer Graphics*, 3(2) :98–117, 1997. 52, 114

[IG97] V. Interrante and C. Grosch. Strategies for effectively visualizing 3d flow with volume lic. In *VIS '97 : Proceedings of the 8th conference on Visualization (VIS'97)*, pages 421–ff., Los Alamitos, CA, USA, 1997. IEEE Computer Society Press. 52

[IKN98] L. Itti, C. Koch, and E. Niebur. A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(11) :1254–1259, 1998. 44

[IKSZ03] A. Iones, A. Krupkin, M. Sbert, and S. Zhukov. Fast, realistic lighting for video games. *IEEE computer graphics and applications*, 23(3) :54–64, 2003. 110

[Jam90] W. James. *The principles of psychology*. New York : Holt, 1890. 41

[Jul71] B. Julesz. *Foundations of Cyclopean Perception*. Chicago : The University of Chicago Press, 1971. 48

[KBSK92] D. Kersten, H. H. Bülthoff, B. L. Schwartz, and K. J. Kurtz. Interaction between transparency and structure from motion. *Neural Computation*, 4(4) :573–589, 1992. 71, 74, 83

[KCVS98] L. Kobbelt, S. Campagna, J. Vorsatz, and H.-P. Seidel. Interactive multi-resolution modeling on arbitrary meshes. In *SIGGRAPH '98 : Proceedings of the 25th annual conference on Computer graphics and interactive techniques*, pages 105–114, New York, NY, USA, 1998. ACM. 154

[KH84] J. T. Kajiya and B. P. Von Herzen. Ray tracing volume densities. In *SIGGRAPH '84 : Proceedings of the 11th annual conference on Computer graphics and interactive techniques*, pages 165–174, New York, NY, USA, 1984. ACM. 65

[KHI+03] R. Kosara, C. G. Healey, V. Interrante, D. H. Laidlaw, and C. Ware. User studies : Why, how, and when ? *IEEE Computer Graphics and Applications*, 23(4) :20–25, 2003. 55, 56, 58

[KK01] R. Kasrai and F. A. Kingdom. Precision, accuracy, and range of perceived achromatic transparency. *J Opt Soc Am A Opt Image Sci Vis*, 18(1) :1–11, Jan 2001. 69

[KK02] R. Kasrai and F. A. Kingdom. Achromatic transparency and the role of local contours. *Perception*, 31(7) :775–790, 2002. 70

[KKMB96] D. Kersten, D. C. Knill, P. Mamassian, and I. Bülthoff. Illusory motion from shadows. *Nature*, 379(6560) :31–31, Jan 1996. 48, 107

[KKVL05] D. F. Keefe, D. B. Karelitz, E. L. Vote, and D. H. Laidlaw. Artistic collaboration in designing vr visualizations. *IEEE Computer Graphics and Applications*, 25(2) :18–23, 2005. 58

[KL05] J. Kontkanen and S. Laine. Ambient occlusion fields. In *I3D '05 : Proceedings of the 2005 symposium on Interactive 3D graphics and games*, pages 41–48. ACM Press, 2005. 110

[KLO06] M. Kass, A. Lefohn, and J. D. Owens. Interactive depth of field using simulated diffusion. Technical report, Pixar Animation Studios, 2006. 48

[KM05] A. Kaufman and K. Mueller. *The Visualization Handbook*, chapter Overview of Volume Rendering, pages 127–174. Academic Press, December 2005. 64, 65

[KMH+02] R. Kosara, S. Miksch, H. Hauser, J. Schrammel, V. Giller, and M. Tscheligi. Useful properties of semantic depth of field for better f+c visualization. In *VISSYM '02 : Proceedings of the symposium on Data Visualisation 2002*, pages 205–210, Aire-la-Ville, Switzerland, Switzerland, 2002. Eurographics Association. 45, 131

[Koe84] J. J. Koenderink. What does the occluding contour tell us about solid shape ? *Perception*, 13(3) :321–330, 1984. 105

[Koe98] J. J. Koenderink. Pictorial relief. *Philosophical Transactions of the Royal Society A : Mathematical, Physical and Engineering Sciences*, 356(1740) :1071–1086, 1998. 51

[KPI+03] J. Kniss, S. Premoze, M. Ikits, A. Lefohn, C. Hansen, and E. Praun. Gaussian transfer functions for multi-field volume visualization. In *VIS '03 : Proceedings of the 14th IEEE Visualization 2003 (VIS'03)*, page 65, Washington, DC, USA, 2003. IEEE Computer Society. 67, 88

[KPvD+07] J. J. Koenderink, S. C. Pont, A. J. van Doorn, A. M. Kappers, and J. T. Todd. The visual light field. *Perception*, 36(11) :1595–1610, 2007. 106

[Kra05] M. Kraus. Scale-invariant volume rendering. In *VIS '05 : Proceedings of the 16th IEEE Visualization 2005 (VIS'05)*, 2005. 68

[KS03] D. C. Knill and J. A. Saunders. Do humans optimally integrate stereo and texture information for judgments of surface slant ? *Vision Res*, 43(24) :2539–2558, Nov 2003. 50

[KS07] M. Kraus and M. Strengert. Depth-of-field rendering by pyramidal image processing. In *Proceedings Eurographics 2007*, 2007. 48

[KSTE06] M. Kersten, J. Stewart, N. F. Troje, and R. E. Ellis. Enhancing depth perception in translucent volumes. *IEEE Transactions on Visualization and Computer Graphics*, 12(5) :1117–1124, 2006. 68, 83, 84, 95

[KSW06] J. Kruger, J. Schneider, and R. Westermann. Clearview : An interactive context preserving hotspot visualization technique. *IEEE Transactions on Visualization and Computer Graphics*, 12(5) :941–948, 2006. 142, 143, 154

[KV06] Y. Kim and A. Varshney. Saliency-guided enhancement for volume visualization. *IEEE Transactions on Visualization and Computer Graphics*, 12(5) :925–932, 2006. Youngmin Kim and Amitabh Varshney. 45, 68, 141

[KvD92] J. J. Koenderink and A. J. van Doorn. Surface shape and curvature scales. *Image and Vision Computing*, 10(8) :557–565, 1992. 113

[KvD03] J. J. Koenderink and A. J. van Doorn. *The Visual Neurosciences*, chapter Shape and Shading, pages 1090–1105. Bradford Books, 2003. 48, 56, 102, 103, 105

[KvDCL96] J. J. Koenderink, A. J. van Doorn, C. Christou, and I. S. Lappin. Perturbation study of shading in pictures. *Perception*, 25(9) :1009–1026, 1996. 103, 105, 106, 110

[KvDK92] J. J. Koenderink, A. J. van Doorn, and A. M. Kappers. Surface perception in pictures. *Percept Psychophys*, 52(5) :487–496, Nov 1992. 129

[KvDPR08] J. J. Koenderink, A. J. van Doorn, S. C. Pont, and W. Richards. Gestalt and phenomenal transparency. *J Opt Soc Am A Opt Image Sci Vis*, 25(1) :190–202, Jan 2008. 69, 71

[KWTM03] G. Kindlmann, R. Whitaker, T. Tasdizen, and T. Moller. Curvature-based transfer functions for direct volume rendering : Methods and applications. In *VIS '03 : Proceedings of the 14th IEEE Visualization 2003 (VIS'03)*, page 67, Washington, DC, USA, 2003. IEEE Computer Society. 67

[KWYM06] F. A. A. Kingdom, K. Wong, A. Yoonessi, and G. Malkoc. Colour contrast influences perceived shape in combined shading and texture patterns. *Spatial Vision*, 19 :147–159(13), April 2006. 137

[LAM01] J. Lext, U. Assarsson, and T. Möller. A benchmark for animated ray tracing. *IEEE Computer Graphics and Applications*, 21(2) :22–31, 2001. 135

[Lav07] G. Lavoué. A roughness measure for 3d mesh visual masking. In *APGV '07 : Proceedings of the 4th symposium on Applied perception in graphics and visualization*, pages 57–60, New York, NY, USA, 2007. ACM. 40

[LB00] M. S. Langer and H. H. Bülthoff. Depth discrimination from shading under diffuse lighting. *Perception*, 29(6) :649–660, 2000. 106, 110, 132

[LB01] M. S. Langer and H. H. Bülthoff. A prior for global convexity in local shape-from-shading. *Perception*, 30(4) :403–410, 2001. 105

[LCD06] T. Luft, C. Colditz, and O. Deussen. Image enhancement by unsharp masking the depth buffer. *ACM Transactions on Graphics*, 25(3) :1206–1213, jul 2006. 111, 113, 134, 135, 158

[Lev88] M. Levoy. Display of surfaces from volume data. *IEEE Computer Graphics and Applications*, 8(3) :29–37, 1988. 67

[LH01] D. P. Luebke and B. Hallen. Perceptually-driven simplification for interactive rendering. In *Rendering Techniques*, pages 223–234, 2001. 39, 40

[LHJ01] E. LaMar, B. Hamann, and K. I. Joy. A magnification lens for interactive volume visualization. In *PG '01 : Proceedings of the 9th Pacific Conference on Computer Graphics and Applications*, page 223, 2001. 141

[LHV04] C. H. Lee, X. Hao, and A. Varshney. Light collages : Lighting design for effective visualization. In *VIS '04 : Proceedings of the conference on Visualization '04*, pages 281–288, Washington, DC, USA, 2004. IEEE Computer Society. 113

[LM05] L. C. Loschky and G. W. McConkie. How late can you update ? detecting blur and transients in gaze-contingent multi-resolutional displays. In *Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting*, pages 1527–1530, Santa Monica CA, 2005. 41

[LME+02] A. Lu, C. J. Morris, D. S. Ebert, P. Rheingans, and C. Hansen. Non-photorealistic volume rendering using stippling techniques. In *VIS '02 : Proceedings of the 13th conference on Visualization 2002 (VIS '02)*, pages 211–218, Washington, DC, USA, 2002. IEEE Computer Society. 100

[LP04] M. Lipps and J. B. Pelz. Yarbus revisited : task-dependent oculomotor behavior. *J. Vis.*, 4(8) :115–115, 8 2004. 42

[LRA+07] W. Li, L. Ritter, M. Agrawala, B. Curless, and D. Salesin. Interactive cutaway illustrations of complex 3d models. In *SIGGRAPH '02 : Proceedings of the 34th annual conference on Computer graphics and interactive techniques*, page 31, New York, NY, USA, 2007. ACM. 142, 143, 157

[LT04] B. Liu and J. T. Todd. Perceptual biases in the interpretation of 3d shape from shading. *Vision Res*, 44(18) :2135–2145, 2004. 105, 107

[LVJ05] C. H. Lee, A. Varshney, and D. W. Jacobs. Mesh saliency. *ACM Transactions on Graphics*, 24(3) :659–666, 2005. 44

[LW90] M. Levoy and R. Whitaker. Gaze-directed volume rendering. In *SI3D '90 : Proceedings of the 1990 symposium on Interactive 3D graphics*, pages 217–223, New York, NY, USA, 1990. ACM. 40, 140

[Mam04] P. Mamassian. Impossible shadows and the shadow correspondence problem. *Perception*, 33(11) :1279–1290, 2004. 107

[Mas06] S. C. Masin. Test of models of achromatic transparency. *Perception*, 35(12) :1611–1624, 2006. 70

[Max88] N. Max. Horizon mapping : Shadows for bump-mapped surfaces. *The Visual Computer*, 4(2) :109–117, 1988. 110, 111

[Max95] N. Max. Optical models for direct volume rendering. *IEEE Transactions on Visualization and Computer Graphics*, 1(2) :99–108, 1995. 65

[MBW04] I. C. McManus, J. Buckman, and E. Woolley. Is light in pictures presumed to come from the left side ? *Perception*, 33(12) :1421–1436, 2004. 129

[MC96] G. W. McConkie and C. B. Currie. Visual stability across saccades while viewing complex pictures. *J Exp Psychol Hum Percept Perform*, 22(3) :563–581, Jun 1996. 43

[MD01] H. Murphy and A. T. Duchowski. Gaze-contingent level of detail rendering. In *EuroGraphics*, 2001. 41

[MD07] H. Murphy and A. T. Duchowski. Hybrid image-/model-based gaze-contingent rendering. In *APGV '07 : Proceedings of the 4th symposium on Applied perception in graphics and visualization*, pages 107–114, New York, NY, USA, 2007. ACM. 40

[MDM07] S. Marchesin, J.-M. Dischler, and C. Mongenet. Feature enhancement using locally adaptive volume rendering. In *IEEE/EG International Symposium on Volume Graphics*. IEEE/EG, september 2007. 68

[ME04] B. Mora and D. S. Ebert. Instant volumetric understanding with order-independent volume rendering. *Computer Graphics Forum*, 23(3) :489–498, 2004. 67

[Met74] F. Metelli. The perception of transparency. *Sci Am*, 230(4) :90–98, Apr 1974. 69

[MG01] P. Mamassian and R. Goutcher. Prior knowledge on the illumination position. *Cognition*, 81(1) :1–9, Aug 2001. 128, 129

[MHB+00] M. Meißner, J. Huang, D. Bartz, K. Mueller, and R. Crawfis. A practical evaluation of popular volume rendering algorithms. In *VVS '00 : Proceedings of the 2000 IEEE symposium on Volume visualization*, pages 81–90, New York, NY, USA, 2000. ACM. 68

[Mil94] G. Miller. Efficient algorithms for local and global accessibility shading. In *SIGGRAPH '94 : Proceedings of the 21st annual conference on Computer graphics and interactive techniques*, pages 319–326, New York, NY, USA, 1994. ACM. 110

[Mit07] M. Mittring. Finding next gen : Cryengine 2. In *SIGGRAPH '07 : ACM SIGGRAPH 2007 courses*, pages 97–121, New York, NY, USA, 2007. ACM. 110, 111, 123, 126, 134

[MK96] P. Mamassian and D. Kersten. Illumination, shading and the perception of local orientation. *Vision Res*, 36(15) :2351–2367, Aug 1996. 105

[MKK98] P. Mamassian, D. C. Knill, and D. Kersten. The perception of cast shadows. *Trends in Cognitive Sciences*, 2(8) :288–295, August 1998. 107

[ML98] P. Mamassian and M. S. Landy. Observer biases in the 3d interpretation of line drawings. *Vision Res*, 38(18) :2817–2832, Sep 1998. 105

[ML01] P. Mamassian and M. S. Landy. Interaction of visual prior constraints. *Vision Res*, 41(20) :2653–2668, Sep 2001. 105

[ML02] G. W. McConkie and L. C. Loschky. Perception onset time during fixations in free viewing. *Behav Res Methods Instrum Comput*, 34(4) :481–490, Nov 2002. 40

[MSE+06] D. Merhof, M. Sonntag, F. Enders, C. Nimsky, and G. Greiner. Hybrid visualization for white matter tracts using triangle strips and point sprites. *IEEE Transactions on Visualization and Computer Graphics*, 12(5) :1181–1188, 2006. 125

[MT86] E. Mingolla and J. T. Todd. Perception of solid shape from shading. *Biol Cybern*, 53(3) :137–151, 1986. 103, 107

[MTB03] M. J. McGuffin, L. Tancau, and R. Balakrishnan. Using deformations for browsing volumetric data. In *VIS '03 : Proceedings of the 14th IEEE Visualization 2003 (VIS'03)*, page 53, Washington, DC, USA, 2003. IEEE Computer Society. 141

[MTHG03] O. Mattausch, T. Theusl, H. Hauser, and E. Gröller. Strategies for interactive exploration of 3d flow using evenly-spaced illuminated streamlines. In *SCCG '03 : Proceedings of the 19th spring conference on Computer graphics*, pages 213–222, New York, NY, USA, 2003. ACM. 114, 140, 141

[MvL00] J. D. Mulder and R. van Liere. Fast perception-based depth of field rendering. In *VRST '00 : Proceedings of the ACM symposium on Virtual reality software and technology*, pages 129–133, New York, NY, USA, 2000. ACM. 40

[MZH+05] M. Meißner, K. Zuiderveld, G. Harris, J. R. Lesser, A. Persson, and M. Vannier. End users' perspectives on volume rendering in medical imaging : A job well done or not over yet ? *IEEE Visualization Conference*, page 119, 2005. 68

[NKK05] H. T. Nefs, J. J. Koenderink, and A. M. Kappers. The influence of illumination direction on the pictorial reliefs of lambertian surfaces. *Perception*, 34(3) :275–287, 2005. 106, 110

[NN85] T. Nishita and E. Nakamae. Continuous tone representation of three-dimensional objects taking account of shadows and interreflection. In *SIGGRAPH '85 : Proceedings of the 12th annual conference on Computer graphics and interactive techniques*, pages 23–30, New York, NY, USA, 1985. ACM. 108

[NTO04] J. F. Norman, J. T. Todd, and G. A. Orban. Perception of three-dimensional shape from specular highlights, deformations of shading, and other types of visual information. *Psychol Sci*, 15(8) :565–570, Aug 2004. 107

[OBA08] J. P. O'Shea, M. S. Banks, and M. Agrawala. The assumed light direction for perceiving shape from shading. In *APGV '08 : Proceedings of the 5th symposium on Applied perception in graphics and visualization*, pages 135–142, New York, NY, USA, 2008. ACM. 128, 129

[OCS05] Y. Ostrovsky, P. Cavanagh, and P. Sinha. Perceiving illumination inconsistencies in scenes. *Perception*, 34(11) :1301–1314, 2005. 106, 128

[OD01] C. O'Sullivan and J. Dingliana. Collisions and perception. *ACM Trans. Graph.*, 20(3) :151–168, 2001. 40

[ODH03] C. O'Sullivan, J. Dingliana, and S Howlett. *The Mind's Eyes : Cognitive and Applied Aspects of Eye Movement Research.*, chapter Eye-movements and Interactive Graphics, pages 555–571. Elsevier Science, Oxford, 2003. 43

[Oli91] J. Oliensis. Shape from shading as a partially well-constrained problem. *CVGIP : Image Understanding*, 54(2) :163–183, 1991. 103

[ORC99] J. K. O'Regan, R. A. Rensink, and J. J. Clark. Change-blindness as a result of 'mudsplashes'. *Nature*, 398(6722) :34, Mar 1999. 43

[Pal99] S. E. Palmer. *Vision Science : Photons to Phenomenology*. MIT Press, 1999. 46

[PD84] T. Porter and T. Duff. Compositing digital images. *Computer Graphics*, 18(3) :253–259, 1984. 69

[PD04] J. T. Petersik and J. L. Dannemiller. Factors influencing the ability to detect motion reversals in rotation simulations. *Spat Vis*, 17(3) :201–234, 2004. 89

[Pfi05] H. Pfister. *The Visualization Handbook*, chapter Hardware-Accelerated Volume Rendering, pages 229–258. Academic Press, December 2005. 65

[Pho75] B. T. Phong. Illumination for computer generated pictures. *Communications of the ACM*, 18(6) :311–317, 1975. 102, 109

[PI08] R. J. Peters and L. Itti. Applying computational tools to predict gaze direction in interactive visual environments. *ACM Transactions on Applied Perception*, 5(2) :1–19, 2008. 44

[PLS+00] H. Pfister, B. Lorensen, W. Schroeder, C. Bajaj, and G. Kindlmann. The transfer function bake-off (panel session). In *VIS '00 : Proceedings of the conference on Visualization '00*, pages 523–526, Los Alamitos, CA, USA, 2000. IEEE Computer Society Press. 67, 88

[PSA+04] G. Petschnigg, R. Szeliski, M. Agrawala, M. Cohen, H. Hoppe, and K. Toyama. Digital photography with flash and no-flash image pairs. *ACM Transactions on Graphics*, 23(3) :664–672, 2004. 120

[PvV05] T. Q. Pham and L. J. van Vliet. Separable bilateral filtering for fast video preprocessing. In *ICME*, pages 454–457, 2005. 120, 126

[PWG04] T. J. Palmeri, A. C. Wong, and I. Gauthier. Computational approaches to the development of perceptual expertise. *Trends Cogn Sci*, 8(8) :378–386, Aug 2004. 57

[PWNL04] D. Purves, S. M. Williams, S. Nundy, and R. B. Lotto. Perceiving the intensity of light. *Psychol Rev*, 111(1) :142–158, Jan 2004. 98, 136

[Ram88] V. S. Ramachandran. Perception of shape from shading. *Nature*, 331(6152) :163–166, Jan 1988. 103, 105, 106, 128

[RBD06] S. Rusinkiewicz, M. Burns, and D. DeCarlo. Exaggerated shading for depicting shape and detail. *ACM Transactions on Graphics*, 25(3) :1199–1205, 2006. 113, 135, 137

[RE01] P. Rheingans and D. Ebert. Volume illustration : Nonphotorealistic rendering of volume models. *IEEE Transactions on Visualization and Computer Graphics*, 7(3) :253–264, 2001. 58, 67, 99

[Red] Martin Reddy. Reducing lags in virtual reality systems using motion-sensitive level of detail. 149

[Red01] M. Reddy. Perceptually optimized 3d graphics. *IEEE Comput. Graph. Appl.*, 21(5) :68–75, 2001. 39, 40

[RHS96] D. L. Ringach, M. J. Hawken, and R. Shapley. Binocular eye movements caused by the perception of three-dimensional structure from motion. *Vision Res*, 36(10) :1479–1492, May 1996. 83, 96

[Rit86] D. Rittenhouse. Explanation of an optical deception. *Transactions of the American Philosophical Society*, 2 :37–42, 1786. 128

[RL95] P. Rheingans and C. Landreth. Perceptual principles for effective visualizations. In *Perceptual Issues in Visualization*, pages 59–74. Springer Verlag, 1995. 137

[RLMS03] E M Reingold, L C Loschky, G W McConkie, and D M Stampe. Gaze-contingent multiresolutional displays : an integrative review. *Hum Factors*, 45(2) :307–328, 2003. 40

[ROJ97] R. A. Rensink, J.K. O'Regan, and Clark J.J. To see or not to see : The need for attention to perceive changes in scenes. *Psychological Science*, 8 :368–373, 1997. 43

[RPG99] M. Ramasubramanian, S. N. Pattanaik, and D. P. Greenberg. A perceptually based physical error metric for realistic image synthesis. In Alyn Rockwood, editor, *SIGGRAPH '99 : Proceedings of the 26th annual conference on Computer graphics and interactive techniques*, Annual Conference Series, pages 73–82. ACM SIGGRAPH, Addison Wesley, Aug 1999. 38

[RR92] S. Rogers and B. J. Rogers. Visual and nonvisual information disambiguate surfaces specified by motion parallax. *Percept Psychophys*, 52(4) :446–452, Oct 1992. 89, 99

[RSH06] T. Ropinski, F. Steinicke, and K. H. Hinrichs. Visually supporting depth perception in angiography imaging. In *Proceedings of the 6th International Symposium on Smart Graphics (SG06)*, pages 93–104, Vancouver, 2006. Springer. 69

[RSK06] C. Rezk-Salama and A. Kolb. Opacity peeling for direct volume rendering. *Computer Graphics Forum*, 25(3) :597–606, 2006. 99

[RT90] F. D. Reichel and J. T. Todd. Perceived depth inversion of smoothly curved surfaces due to image orientation. *J Exp Psychol Hum Percept Perform*, 16(3) :653–664, Aug 1990. 105, 132

[SA02] M. Singh and B. L. Anderson. Toward a perceptual theory of transparency. *Psychol Rev*, 109(3) :492–519, Jul 2002. 69, 97

[SA07] P. Shanmugam and O. Arikan. Hardware accelerated ambient occlusion techniques on gpus. In *I3D '07 : Proceedings of the 2007 symposium on Interactive 3D graphics and games*, pages 73–80, New York, NY, USA, 2007. ACM. 110

[Sab88] P. Sabella. A rendering algorithm for visualizing 3d scalar fields. In *SIGGRAPH '88 : Proceedings of the 15th annual conference on Computer graphics and interactive techniques*, pages 51–58, New York, NY, USA, 1988. ACM Press. 52, 64, 65

[SAH04] J. Schumacher, R. S. Allison, and R. Herpers. Using saccadic suppression to hide graphic updates. In *10th Eurographics Symposium on Virtual Environments*, pages 17–24, June 2004. 43

[Sai94] T. Saito. Real-time previewing for volume visualization. In *VVS '94 : Proceedings of the 1994 symposium on Volume visualization*, pages 99–106, New York, NY, USA, 1994. ACM. 53, 67

[SB97] S. M. Smith and J. M. Brady. Susan—a new approach to low level image processing. *International Journal of Computer Vision*, 23(1) :45–78, 1997. 120

[SC99] D. J. Simons and C. F. Chabris. Gorillas in our midst : sustained inattentional blindness for dynamic events. *Perception*, 28(9) :1059–1074, 1999. 43

[SD94] G. Sperling and B. A. Dosher. *Early vision and beyond*, chapter Depth from motion, pages 133–142. Cambridge, MA : MIT Press, 1994. 55

[SDL+05] V. Sundstedt, K. Debattista, P. Longhurst, A. Chalmers, and T. Troscianko. Visual attention for efficient high-fidelity graphics. In *SCCG '05 : Proceedings of the 21st spring conference on Computer graphics*, pages 169–175, New York, NY, USA, 2005. ACM. 44

[SDM05] J. L. Snyder, K. Doerschner, and L. T. Maloney. Illumination estimation in three-dimensional scenes with and without specular cues. *J Vis*, 5(10) :863–877, 2005. 137

[SE03] N. A. Svakhine and D. S. Ebert. Interactive volume illustration and feature halos. In *PG '03 : Proceedings of the 11th Pacific Conference on Computer Graphics and Applications*, page 347, Washington, DC, USA, 2003. IEEE Computer Society. 58

[SFFP04] S. Savarese, L. Fei-Fei, and P. Perona. What do reflections tell us about the shape of a mirror ? In *APGV '04 : Proceedings of the 1st Symposium on Applied perception in graphics and visualization*, pages 115–118, New York, NY, USA, 2004. ACM. 107

[SGS05] C. Stoll, S. Gumhold, and H.-P. Seidel. Visualization with stylized line primitives. In *VIS '05 : Proceedings of the 16th IEEE Visualization 2005 (VIS'05)*, page 88, 2005. 114, 125

[SH98] M. Singh and D. D. Hoffman. Part boundaries alter the perception of transparency. *Psychol Sci*, 9 :370–378(9), 1998. 70

[Sin04] M. Singh. Lightness constancy through transparency : internal consistency in layered surface representations. *Vision Res*, 44(15) :1827–1842, 2004. 69

[SL89] G. Sperling and M. S. Landy. Kinetic depth effect and identification of shape. *J Exp Psychol Hum Percept Perform*, 15(4) :826–840, Nov 1989. 91

[SL01] R. Shacked and D. Lischinski. Automatic lighting design using a perceptual quality metric. *Computer Graphics Forum*, 20 :215–227(13), September 2001. 113

[SLM02] A. Stompel, E. B. Lum, and K.-L. Ma. Visualization of multidimensional, multivariate volume data using hardware-accelerated non-photorealistic rendering techniques. In *PG '02 : Proceedings of the 10th Pacific Conference on Computer Graphics and Applications*, page 394, Washington, DC, USA, 2002. IEEE Computer Society. 58

[SMGG01] P.-P. J. Sloan, W. Martin, A. Gooch, and B. Gooch. The lit sphere : a model for capturing npr shading from art. In *GRIN'01 : Graphics interface 2001*, pages 143–150, Toronto, Ont., Canada, Canada, 2001. Canadian Information Processing Society. 112

[SML06] W. Schroeder, K. Martin, and B. Lorensen. *The Visualization Toolkit An Object-Oriented Approach To 3D Graphics*. Kitware, Inc., 2006. 28

[SP96a] P. Sinha and T. Poggio. Role of learning in three-dimensional form perception. *Nature*, 384(6608) :460–463, Dec 1996. 50

[SP96b] J. Sun and P. Perona. Early computation of shape and reflectance in the visual system. *Nature*, 379(6561) :165–168, Jan 1996. 106

[SP98] J. Sun and P. Perona. Where is the sun ? *Nat Neurosci*, 1(3) :183–184, Jul 1998. 128, 129

[SP00] C. R. Sears and Z. W. Pylyshyn. Multiple object tracking and attentional processing. *Can J Exp Psychol*, 54(1) :1–14, Mar 2000. 41

[SR05] D. J. Simons and R. A. Rensink. Change blindness : past, present, and future. *Trends Cogn Sci*, 9(1) :16–20, Jan 2005. 43

[SS99] M. E. Sereno and M. I. Sereno. 2-d center-surround effects on 3-d structure-from-motion. *J Exp Psychol Hum Percept Perform*, 25(6) :1834–1854, Dec 1999. 96, 99, 100

[SSH01] N. E. Scott-Samuel and R. F. Hess. What does the ternus display tell us about motion processing in human vision ? *Perception*, 30(10) :1179–1188, 2001. 53

[SSMK05] M. Sattler, R. Sarlette, T. Mücken, and R. Klein. Exploitation of human shadow perception for fast shadow rendering. In *APGV '05 : Proceedings of the 2nd symposium on Applied perception in graphics and visualization*, pages 131–134, New York, NY, USA, 2005. ACM. 107

[SSN+98] Y. Sato, N. Shiraga, S. Nakajima, S. Tamura, and R. Kikinis. Local maximum intensity projection (lmip) : a new rendering method for vascular visualization. *J Comput Assist Tomogr*, 22(6) :912–917, Nov-Dec 1998. 67

[ST90] T. Saito and T. Takahashi. Comprehensible rendering of 3-d shapes. *ACM SIGGRAPH Computer Graphics*, 24(4) :197–206, 1990. 108, 112

[Ste98] A. James Stewart. Fast horizon computation at all points of a terrain with visibility and shading applications. *IEEE Transactions on Visualization and Computer Graphics*, 4(1) :82–93, 1998. 110, 111

[Ste03] A. J. Stewart. Vicinity shading for enhanced perception of volumetric data. In *VIS '03 : Proceedings of the 14th IEEE Visualization 2003 (VIS'03)*, page 47, Washington, DC, USA, 2003. IEEE Computer Society. 114

[SW05] A. Seitz and T. Watanabe. A unified model for perceptual learning. *Trends Cogn Sci*, 9(7) :329–334, Jul 2005. 57

[TCM06] M. Tarini, P. Cignoni, and C. Montani. Ambient occlusion and edge cueing for enhancing real time molecular visualization. *IEEE Transactions on Visualization and Computer Graphics*, 12(5) :1237–1244, 2006. 109, 110, 114

[TG80] A. M. Treisman and G. Gelade. A feature-integration theory of attention. *Cognit Psychol*, 12(1) :97–136, Jan 1980. 44

[TM83] J. T. Todd and E. Mingolla. Perception of surface curvature and direction of illumination from patterns of shading. *J Exp Psychol Hum Percept Perform*, 9(4) :583–595, Aug 1983. 107

[TM98] C. Tomasi and R. Manduchi. Bilateral filtering for gray and color images. In *Proceedings of the 1998 IEEE International Conference on Computer Vision*, pages 839–846, 1998. 120

[TOCH06] A. Torralba, A. Oliva, M. S. Castelhano, and J. M. Henderson. Contextual guidance of eye movements and attention in real-world scenes : the role of global features in object search. *Psychol Rev*, 113(4) :766–786, Oct 2006. 44

[Tod85] J. T. Todd. Perception of structure from motion : is projective correspondence of moving elements a necessary condition ? *J Exp Psychol Hum Percept Perform*, 11(6) :689–710, Dec 1985. 95

[Tod04] J. T. Todd. The visual perception of 3d shape. *Trends Cogn Sci*, 8(3) :115–121, Mar 2004. 46

[TR89] J. T. Todd and F. D. Reichel. Ordinal structure in the visual perception and cognition of smoothly curved surfaces. *Psychol Rev*, 96(4) :643–657, Oct 1989. 105, 132

[TR90] J. T. Todd and F. D. Reichel. Visual perception of smoothly curved surfaces from double-projected contour patterns. *J Exp Psychol Hum Percept Perform*, 16(3) :665–674, Aug 1990. 56

[UK88] C. Upson and M. Keeler. V-buffer : visible volume rendering. In *SIGGRAPH '88 : Proceedings of the 15th annual conference on Computer graphics and interactive techniques*, pages 59–64, New York, NY, USA, 1988. ACM. 64, 65

[Ull79] S. Ullman. *The interpretation of visual motion*. PhD thesis, Mass. Inst. of Technol., Dept. of Elect. Eng. Comput. Sci., 1979. 49

[VBGS08] Romain Vergne, Pascal Barla, Xavier Granier, and Christophe Schlick. Apparent relief : a shape descriptor for stylized shading. In *NPAR '08 : Proceedings of the 6th international symposium on Non-photorealistic animation and rendering*, pages 23–29, New York, NY, USA, 2008. ACM. 113

[VFSG06] I. Viola, M. Feixas, M. Sbert, and E. Gröller. Importance-driven focus of attention. *IEEE Transactions on Visualization and Computer Graphics*, 12(5) :933–940, October 2006. 141

[VG05] Ivan Viola and E. Gröller. Smart visibility in visualization. In L. Neumann, M. Sberta, B. Gooch, and W. Purgathofer, editors, *Proceedings of EG Workshop on Computational Aesthetics Computational Aesthetics in Graphics, Visualization and Imaging*, pages 209–216, 5 2005. 68, 142

[VGB05] D. Vishwanath, A. R. Girshick, and M. S. Banks. Why pictures look right when viewed from the wrong place. *Nat Neurosci*, 8(10) :1401–1410, Oct 2005. 51

[Vil06] C. Villarruel. *Computer graphics and human depth perception with gaze-contingent depth of field*. PhD thesis, Faculty of Mount Holyoke College, 2006. 40

[VM07] S. Vogt and S. Magnussen. Expertise in pictorial perception : eye-movement patterns and visual memory in artists and laymen. *Perception*, 36(1) :91–100, 2007. 45

[vNB67] F. L. van Nes and M. A. Bouman. Spatial modulation transfer in the human eye. *J opt Soc Am*, 57 :401–406, March 1967. 88

[vW06] J. J. van Wijk. Views on visualization. *IEEE Transactions on Visualization and Computer Graphics*, 12(4) :421–433, 2006. 58

[WB08] C. Weigle and D. C. Banks. A comparison of the perceptual benefits of linear perspective and physically-based illumination for display of dense 3d streamtubes. In *VIS '08 : Proceedings of the 19th IEEE Visualization 2008 (VIS'08)*, 2008. 56, 114

[WFG92] L. C. Wanger, J. A. Ferwerda, and D. P. Greenberg. Perceiving spatial relationships in computer-generated images. *IEEE Computer Graphics and Applications*, 12(3) :44–51, 54–58, 1992. 51, 107

[WH06] D. Weiskopf and H. Hauser. Cycle shading for the assessment and visualization of shape in one and two codimensions. In *GI '06 : Proceedings of Graphics Interface 2006*, pages 219–226, Toronto, Ont., Canada, Canada, 2006. Canadian Information Processing Society. 113

[Whi80] T. Whitted. An improved illumination model for shaded display. *Communications of the ACM*, 23(6) :343–349, 1980. 102, 108

[WKE99] R. Westermann, L. Kobbelt, and T. Ertl. Real-time exploration of regular volume data by adaptive reconstruction of iso-surfaces. *The Visual Computer*, 15(2) :100–111, 1999. 140, 141

[WLC+03] N. Williams, D. Luebke, J. D. Cohen, M. Kelley, and B. Schubert. Perceptually guided simplification of lit, textured meshes. In *SI3D '03 : Proceedings of the 2003 symposium on Interactive 3D graphics*, pages 113–121, New York, NY, USA, 2003. ACM Press. 39

[WLT08] Y.-S. Wang, T.-Y. Lee, and C.-L. Tai. Focus+context visualization with distortion minimization. In *VIS '08 : Proceedings of the 19th IEEE Visualization 2008 (VIS'08)*, volume 14, 2008. 141

[WNS01] T. Watanabe, J. E. Náñez, and Y. Sasaki. Perceptual learning without perception. *Nature*, 413(6858) :844–848, Oct 2001. 57

[WO53] H. Wallach and D. N. O'Connell. The kinetic depth effect. *J Exp Psychol*, 45(4) :205–217, Apr 1953. 49

[WOG06] H. Winnemöller, S. C. Olsen, and B. Gooch. Real-time video abstraction. *ACM Transactions on Graphics*, 25(3) :1221–1226, 2006. 120

[WPLD01] M. Wexler, F. Panerai, I. Lamouret, and J. Droulez. Self-motion and the perception of stationary objects. *Nature*, 409(6816) :85–88, Jan 2001. 49

[WQW+06] H.-C. Wong, H. Qu, U.-H. Wong, Z. Tang, and K. Mueller. A perceptual framework for comparisons of direct volume rendered images. In Long-Wen Chang and Wen-Nung Lie, editors, *PSIVT 2006 : Pacific-Rim Symposium on Image and Video Tecnology*, volume 4319 of *Lecture Notes in Computer Science*, pages 1314–1323. Springer, 2006. 68

[WWHW97] B. Watson, N. Walker, L. F. Hodges, and A. Worden. Managing level of detail through peripheral degradation : effects on search performance with a head-mounted display. *ACM Transactions on Computer-Human Interaction*, 4(4) :323–346, 1997. 40

[WZMK05] L. Wang, Y. Zhao, K. Mueller, and A. Kaufman. The magic volume lens : An interactive focus+context technique for volume rendering. *VIS '05 : Proceedings of the 16th IEEE Visualization 2005 (VIS'05)*, pages 367–374, 2005. 141

[Yar67] A. L. Yarbus. *Eye Movements and Vision*. New York : Plenum Press, 1967. 42

[YPG01] H. Yee, S. Pattanaik, and D. P. Greenberg. Spatiotemporal sensitivity and visual attention for efficient rendering of dynamic environments. *ACM Transactions on Graphics*, 20(1) :39–65, January 2001. 38, 149

[ZIK98] S. Zhukov, A. Inoes, and G. Kronin. An ambient light illumination model. In George Drettakis and Nelson Max, editors, *Rendering Techniques '98*, Eurographics, pages 45–56. Springer-Verlag Wien New York, 1998. 110

[ZSH96] M. Zöckler, D. Stalling, and H.-C. Hege. Interactive visualization of 3d-vector fields using illuminated stream lines. In *VIS '96 : Proceedings of the 7th IEEE Visualization 2006 (VIS'96)*, pages 107–113, Los Alamitos, CA, USA, 1996. IEEE Computer Society Press. 114