



HAL
open science

Génération d'images intermédiaires pour la création d'animations 2D stylisées à base de marques

Nicolas Barroso

► **To cite this version:**

Nicolas Barroso. Génération d'images intermédiaires pour la création d'animations 2D stylisées à base de marques. Sciences de l'information et de la communication. Université de Toulouse, 2024. Français. NNT : 2024TLSES083 . tel-04700257

HAL Id: tel-04700257

<https://theses.hal.science/tel-04700257v1>

Submitted on 17 Sep 2024

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.

Doctorat de l'Université de Toulouse

préparé à l'Université Toulouse III - Paul Sabatier

Génération d'images intermédiaires pour la création
d'animations 2D stylisées à base de marques

Thèse présentée et soutenue, le 12 juin 2024 par

Nicolas BARROSO

École doctorale

EDMITT - Ecole Doctorale Mathématiques, Informatique et Télécommunications de Toulouse

Spécialité

Informatique et Télécommunications

Unité de recherche

IRIT : Institut de Recherche en Informatique de Toulouse

Thèse dirigée par

Mathias PAULIN et David VANDERHAEGHE

Composition du jury

M. Damien ROHMER, Rapporteur, Institut Polytechnique de Paris

M. Daniel MENEVEAUX, Rapporteur, Université de Poitiers

Mme Johanna DELANOY, Examinatrice, INSA Lyon

Mme Florence SEDES, Examinatrice, Université Toulouse III - Paul Sabatier

M. Mathias PAULIN, Directeur de thèse, Université Toulouse III - Paul Sabatier

M. David VANDERHAEGHE, Co-directeur de thèse, Université Toulouse III - Paul Sabatier

Table des matières

1	Introduction	7
1.1	Rendu expressif	7
1.1.1	Simulation des techniques manuelles	8
1.1.2	Stylisation automatique	12
1.1.3	Stylisation et perception	16
1.2	Principes de base du rendu à base de marque	17
1.2.1	Les différents types de marques	17
1.2.2	Comparaison avec les approches textures	18
1.3	L'animation traditionnelle 2D	19
1.3.1	Application à l'animation	19
1.3.2	Les problématiques liées à la stylisation de l'animation	20
1.3.3	Automatisation des techniques traditionnelles	22
1.4	Contributions de la thèse	25
2	Fondements théoriques	27
2.1	État de l'art	27
2.1.1	Le mouvement	27
2.1.2	Stylisation basée synthèse de texture	31
2.1.3	Stylisation à base de marques	34
2.2	Positionnement scientifique	38
2.2.1	Les enjeux de la génération d'images intermédiaires	38
3	Génération d'images intermédiaires	42
3.1	Fondements pratiques	42
3.1.1	Représentation des coups de pinceau	42
3.1.2	Simulation de peinture	43
3.1.3	Composition d'une animation	44
3.1.4	Représentation du mouvement	44
3.2	Vue d'ensemble	45
3.2.1	Principe de la méthode	46
3.2.2	Flux de travail	46
3.3	Mouvement de l'animation	48
3.3.1	Différentes sources de mouvement	48
3.3.2	Extrapolation du mouvement	51
3.3.3	Poids de confiance	53
3.4	Génération de coups de pinceau	54
3.4.1	Propagation des coups de pinceau	54
3.4.2	Organisation des marques candidates	58

3.5	Sélection des marques candidates	60
3.5.1	Génération des cartes de quantité de matière	61
3.5.2	Heuristique de sélection des marques	62
3.5.3	Évaluation des coups de pinceau	64
3.5.4	Les marques d'étalements	66
4	Synthèse des résultats obtenus	69
4.1	Présentation des animations générées	69
4.1.1	Tableau des scènes	71
4.1.2	Impact du choix du simulateur	73
4.2	Implémentations	73
4.2.1	Performance	73
4.2.2	Mesure des temps de calcul	74
4.3	Comparaison avec les animations traditionnelles	74
4.4	Application aux cas problématiques d'animations	75
4.5	Limitations de la méthode	77
4.6	Perspectives des travaux menés	81
5	Conclusion	84
6	Annexes	86
6.1	Vue d'ensemble du système	86

Table des figures

1.1	Modèles de simulation physique de la dynamique d'un pinceau . . .	9
1.2	Modèles de simulation de médium	10
1.3	Plusieurs pigments synthétiques rendus avec le modèle de réflec- tance <i>Kubelka-Munk</i>	11
1.4	Simulations phénoménologiques dirigées par les données	12
1.5	Stylisation à base de marques d'une image	13
1.6	Stylisation à base de marques d'une scène 3D	14
1.7	Principe de l'analogie d'images	15
1.8	Stylisation par l'exemple d'une image avec l'utilisation d'un réseau de neurones	15
1.9	Stylisation d'une image améliorant la perception des structures . .	16
1.10	Stylisation d'une image améliorant la perception des détails	17
1.11	Le style pointilliste	18
1.12	Mouvements exagérés suggérant l'intensité du mouvement.	20
1.13	Principes du <i>Timing and spacing</i>	20
1.14	Diagramme de la cohérence temporelle	21
1.15	La technique du pose à pose	23
1.16	La technique <i>sous la caméra</i>	23
1.17	Animation traditionnelle combinant la peinture et le dessin au trait	24
2.1	Stylisations semi-automatiques à base de marque d'images	40
2.2	Étapes de l'interpolation d'images clés	41
3.1	Encodage du mouvement dans le champ vectoriel	45
3.2	Vue d'ensemble de la génération d'images intermédiaires	46
3.3	Flux de travail	47
3.4	Mouvement dessiné	49
3.5	Champ vectoriel obtenu avec le logiciel <i>Blender</i>	50
3.6	Champ vectoriel estimé depuis le flux vidéo de l'animation <i>Sintel</i> . .	51
3.7	Extrapolation par diffusion bi-harmonique	52
3.8	Singularités de mouvement après extrapolation	53
3.9	Interpolation linéaire et non linéaire des courbes	57
3.10	Correction des courbes advectées	58
3.11	Intérêt de l'ordonnancement par groupe de marque	60
3.12	Sélection de toutes les marques candidates	61
3.13	Génération de la carte de quantité de matière attendue	62
3.14	Intérêt de la découpe des marques candidates	65
3.15	Impact de la découpe des marques candidates	66
3.16	Impact de la découpe des marques d'étalement	68

4.4	Extraits d'animations traditionnelles	74
4.5	Limite d'un mouvement non bijectif	78
4.6	Impact de la qualité du mouvement	79
4.7	Limite de l'ordonnancement de notre système	81
4.8	Représentation temporelle du mouvement	83
6.1	Pipeline de la méthode	87

Remerciements

Je souhaiterais exprimer ma profonde gratitude envers mes directeurs de thèse, Mathias Paulin et David Vanderhaeghe, qui ont su susciter ma passion pour l'informatique graphique dès le Master et m'ont accordé leur confiance pour mener à bien ces travaux de recherche. Un merci particulièrement chaleureux à David pour sa patience, ses conseils et sa bienveillance. Je tiens également à remercier Loïc et Nicolas pour leur disponibilité et leurs explications au combien précieuses concernant les problèmes que j'ai pu rencontrer. Mes remerciements vont également à l'équipe *STORM* du laboratoire *IRIT* pour ses discussions stimulantes et ses soirées d'équipes autour d'un bon barbecue : Hugo, Chems, Amélie, les deux Pierre, les deux François, Gauthier, Willhem, Mégane, Jules, mais aussi Arthur et Vassili même si vous ne faisiez pas partie de la même équipe et tout ceux ayant fait un passage éclair dans la grotte des doctorants.

À mes amis, pour leur compréhension, leurs moments de détente bienvenus et leur soutien moral précieux qui ont égayé les phases intenses de rédaction et de recherche. Je dédie ce manuscrit à ma compagne Élodie, qui m'a toujours soutenu dans les moments difficiles comme dans les meilleurs, ce qui me permet aujourd'hui de terminer ce document.

Resumé

Dans le cadre de ma thèse, je m'intéresse à la problématique de la création d'animations 2D traditionnelles, où toutes les images sont réalisées à la main. Je cherche à explorer comment l'ordinateur peut aider les artistes à produire de manière plus efficace sans pour autant empiéter sur le spectre de la création artistique. Pour répondre à ce problème, mes travaux se placent dans le domaine des méthodes automatiques, où l'animateur travaille de manière itérative avec l'ordinateur. Je propose une méthode qui, à partir de deux images clés et d'une série de champs vectoriels 2D décrivant le mouvement dans l'espace image de l'animation, génère des images intermédiaires et reproduit le style donné en exemple. Ma méthode se situe à l'intersection de deux techniques manuelles d'animation : l'animation *pose à pose* et l'animation *sous la caméra*, et assure un contrôle fort en permettant d'éditer n'importe quelle image générée de la même manière que celles données en exemple. Mes travaux englobent plusieurs domaines, notamment l'analyse du mouvement, le contrôle de courbe 2D, le rendu à base de marques et la simulation de la peinture.

Abstract

As part of my thesis, I am interested in the issue of creating traditional 2D animations, where all the images are handcrafted. Specifically, I explore how computers can assist artists in producing animations efficiently without reducing the artistic creative process. To address this problem, my work falls within the scope of automatic methods, where the animator collaborates iteratively with the computer. I propose a method that takes two keyframe images and a series of 2D vector fields describing the motion in image space of the animation, and generates intermediate images while preserving the given style as an example. My method combines two manual animation techniques : pose-to-pose and frame-by-frame animation, providing strong control by allowing any generated image to be edited in the same way as the example images provided. My research covers several domains : motion analysis, 2D curve control, mark-based rendering, and paint simulation.

Chapitre 1

Introduction

Les travaux de ma thèse s’inscrivent dans le domaine de l’informatique graphique, qui regroupe : la géométrie, le rendu, l’animation et la simulation. Le rendu expressif est une composante du rendu et se différencie du rendu photoréaliste. Il regroupe les différentes méthodes qui capturent et reproduisent les styles artistiques dans les œuvres numériques. Plus précisément, mes travaux portent sur la génération d’animations 2D en style peinture, au sein du rendu expressif. L’objectif de cette thèse est de proposer des méthodes novatrices pour générer des animations 2D qui dépassent les limitations des approches conventionnelles, en offrant aux artistes et aux créateurs de nouveaux outils.

Ces travaux ont été rendus possible grâce au soutien financier du projet ANR ANR-19-CE38-0009-01 dont l’objectif principal est d’étudier les représentations intermédiaires entre la scène d’entrée et le processus de stylisation afin de guider le contrôle artistique des animations.

1.1 Rendu expressif

L’expressivité est la capacité d’une image à transmettre un message, une émotion ou un sentiment avec intention. Le style est un concept abstrait qui influence la qualité de cette transmission en caractérisant l’apparence [111]. Le style est défini par l’ensemble des moyens qui influencent l’apparence. Cette définition fait consensus dans la littérature scientifique du rendu expressif, et lie ce que l’on souhaite représenter [110] : la scène, et la façon dont on le représente : le rendu. On parle alors de fonction de stylisation définissant les caractéristiques de scène : l’éclairage, la vue, les objets représentés et les caractéristiques de rendu comme la couleur, les contours et l’épaisseur.

Le rendu expressif est alors défini comme l’ensemble des techniques qui visent à modifier l’apparence visuelle d’une image ou d’une animation de manière à faire transparaître un style artistique spécifique. L’expressivité s’exprime de façon statique par la stylisation d’une seule image ou de façon dynamique par l’impact du mouvement sur la stylisation d’une séquence d’images.

Intuitivement, la première idée pour parvenir à une stylisation par ordinateur est d'arriver à reproduire les outils utilisés par les artistes et l'interaction entre un médium et un substrat. Que ce soit de manière physiquement réaliste ou seulement plausible, les chercheurs ont trouvé des moyens de reproduire dans un premier temps les différents styles à travers la simulation de divers médiums et dans un second temps d'étendre les possibilités artistiques pour subjuguer l'information et la perception de l'image. Les prochaines sections (1.1.1, 1.1.2) se focalisent sur le style artistique peinture et proposent une revue non exhaustive des méthodes de chaque domaine.

1.1.1 Simulation des techniques manuelles

Simulation physique des médiums (peinture, aquarelle, crayon...)

L'approche basée physique consiste à établir un modèle utilisant des paramètres temporels mesurables physiquement, étudier leur évolution et produire une simulation mathématique ou informatique réaliste. Ces approches cherchent à trouver un lien exprimable entre la reproduction d'un phénomène réel et ses paramètres physiques.

• **Dynamique de pinceaux** Parmi les outils utilisés pour exprimer la créativité des artistes, les pinceaux jouent un rôle essentiel dans le rendu artistique final en raison de leur maniabilité et de leurs propriétés de déformation. Ils sont adaptés pour peindre des contours, ajouter des détails fins et réaliser des remplissages. Un premier modèle de simulation physique des coups de pinceaux est introduit par STRASSMANN [97] et modélise l'interaction de chaque poil avec le papier. Son modèle permet de contrôler la distribution des couleurs dans le pinceau, ce qui permet de créer des coups de pinceau avec des couleurs uniformes ou interpolées (Figure 1.1a). Il jette les bases d'une première approche sur l'évolution des quantités de peinture dans les poils du pinceau, sous forme d'automate cellulaire régit par des règles de transfert.

Les techniques permettent d'atteindre une simulation temps réels en calculant la déformation des poils du pinceau par groupes au lieu d'effectuer une simulation poil par poil. Au début des années 2000, les poils du pinceaux sont modélisés par un unique maillage qui une fois transformé donne la déformation de l'ensemble des poils [5, 20]. Ces technique utilisent des approximations physiques pour réduire la charge de calcul afin de proposer une déformation du pinceau interactive. L'utilisation de maillages est reprise par VAN LAERHOVEN et al. [103], qui utilisent plusieurs maillages indépendants pour représenter différentes déformations correspondant aux poils principaux du pinceau. Ces maillages sont ensuite interpolés au gré des déformations du pinceau [6]. Ces solutions temps réel ne permettent pas de simuler les déformations les plus complexes comme par exemple les irrégularités de distribution de poils. Les capacités de calculs augmentant fortement au début des années 2010, la simulation physique de milliers de poil devient alors possible à une fréquence de 30Hz grâce au modèle de DIVERDI et al. [29] (Figure 1.1b). Basé sur une résolution d'optimisation numérique, il consiste à transformer les

forces qui s'appliquent à la base du pinceau pour les répercuter jusqu'à la pointe de chaque poil en mettant à jour l'état de chaque maillage.

Certains styles requièrent l'utilisation de modélisations physiques de pinceaux virtuels spécifiques, tels que la calligraphie chinoise [62, 114]. Ces techniques mettent l'accent sur différentes caractéristiques comme par exemple les déformations de poils liées aux effets d'aplatissement ou d'écartement. La simulation de la plasticité des pinceaux mouillés joue un rôle important dans la calligraphie chinoise. Le modèle CHU et al. [19] apporte cette spécificité et représente le pinceau en une combinaison de poils longs et de poils courts (Figure 1.1c). Chaque poil possède plusieurs noeuds de calcul où est évalué la quantité d'encre absorbée le long de leur courbe. Par la suite la méthode est modifiée pour la rendre compatible avec le temps réel. Les poils sont alors regroupés, ce qui permet d'évaluer la déformation du pinceau et l'absorption de l'encre directement par groupe de poils [117].

Toujours dans le cadre de la peinture mais sans pinceaux, une technique s'intéresse à l'aérogaphie et simule la pulvérisation de peinture sur une surface. Le modèle proposé utilise le lancer de rayon pour simuler la pulvérisation de particules [59, 60]. La direction des particules pulvérisées est choisie selon une fonction de distribution et la simulation physique a pour paramètres : la viscosité de la peinture, la distance de pulvérisation, la pression de l'air, et le flux de peinture pulvérisée.

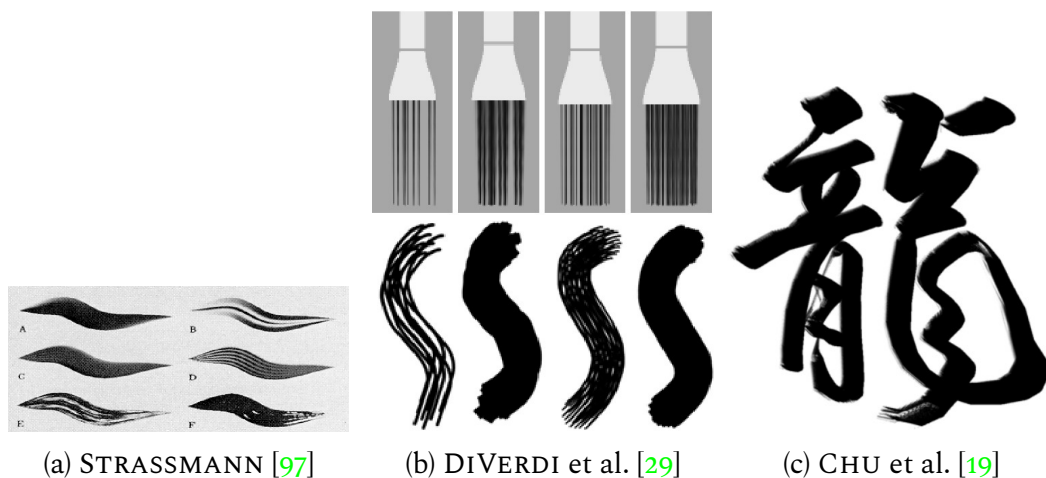


FIGURE 1.1 – Différents modèles de simulation physique de la dynamique d'un pinceau. Images tirées des articles correspondants.

• **Diffusion du médium** La simulation de la diffusion du médium correspond au transfert de matière contenue dans le pinceau au support, le substrat, et à son évolution au cours du temps. La plupart des modèles représentent le support (i.e. toile ou papier) comme une carte de cellule stockant le type, la quantité de peinture déposée et la quantité de peinture maximale stockée. Le modèle *Wet and Sticky* de COCKSHOT [23] est basé sur cette représentation et définit des règles de diffusion en fonction des caractéristiques de la peinture utilisée, sous la forme d'automate cellulaire, reprenant ainsi le concept de transfert de matière entre les poils du pinceau de STRASSMANN [97]. Il permet au surplus de peinture déposée de glisser

vers des cellules voisines non remplies, de simuler des éclaboussures de peinture lorsque les poils du pinceaux sont secs, rigides et soumis à une tension. Le modèle *IMPASTO* de BAXTER et al. [9] vient compléter la simulation physique, pour la diffusion de peinture acrylique ou à huile, en proposant une advection conservatrice des quantités de matières (Figure 1.2a). D'autres comme CHU et al. [21], proposent un modèle capable de reproduire les détails d'advection des quantités de matières tout en réduisant la charge de calcul pour la rendre compatible avec une utilisation en temps réel.

Spécifique à l'aquarelle, le modèle de SMALL [94] reprend la formulation du support en un amas cellulaire et propose une diffusion de pigment de couleur en fonction de paramètres physiques comme l'absorption des fibres du papier, la gravité ou la masse moléculaire du pigment utilisé. Le modèle est amélioré par CURTIS et al. [25] qui combine une structure du papier en couches translucides superposées avec une méthode de rendu qui tient compte du flux optique qui traverse ces couches. Cela permet de produire des résultats convaincants, mais aussi très coûteux ce qui conduit à l'émergence de techniques temps réel d'abord avec VAN LAERHOVEN et al. [102] (Figure 1.2b) puis CHU et al. [22]. XU et al. [116] proposent un modèle de diffusion intégrant l'évaporation de la solution utilisée pour l'aquarelle (Figure 1.2c). Le modèle s'inspire des principes de diffusion dans l'industrie chimique et textile. La simulation de diffusion de médium ne se limite pas seulement à la peinture, elle s'étend également au graphite [96], et au pastel [86].



FIGURE 1.2 – Différents modèles de simulation de médium. *Images tirées des articles correspondants.*

• **Réfectance** La simulation de la réfectance s'intéresse à la proportion de lumière réfléchi à la surface d'un matériau. Elle permet de déterminer la couleur perçue par l'œil après que la lumière ait été absorbée et diffusée par ce matériau. HAASE et al. [43] introduisent la théorie de Kubelka-Munk pour le mélange réaliste de pigment dans le domaine du rendu. Le modèle de Kubelka-Munk est une approximation du comportement de la lumière entre différentes couches de pigments avec des concentrations variables. Ils montrent les différentes variations du modèle, ses limites et ses hypothèses sur la représentation d'un pigment ou d'un mélange à différentes concentrations. Le modèle sera ensuite largement utilisé dans le rendu expressif pour décrire la réfectance issue de l'utilisation de pigments synthétiques [25, 101] (Figure 1.3). Le modèle est simplifié par la suite pour

atteindre des performances temps réels [85]. La méthode BAXTER et al. [9] opte pour un échantillonnage plus précis des longueurs d'ondes composant le spectre de la lumière. De son côté, XU et al. [116] utilisent un modèle basé sur les réseaux de neurones pour reproduire le mélange apparent de pigments en entraînant le réseau sur une série de mesures physiques de réflectance.

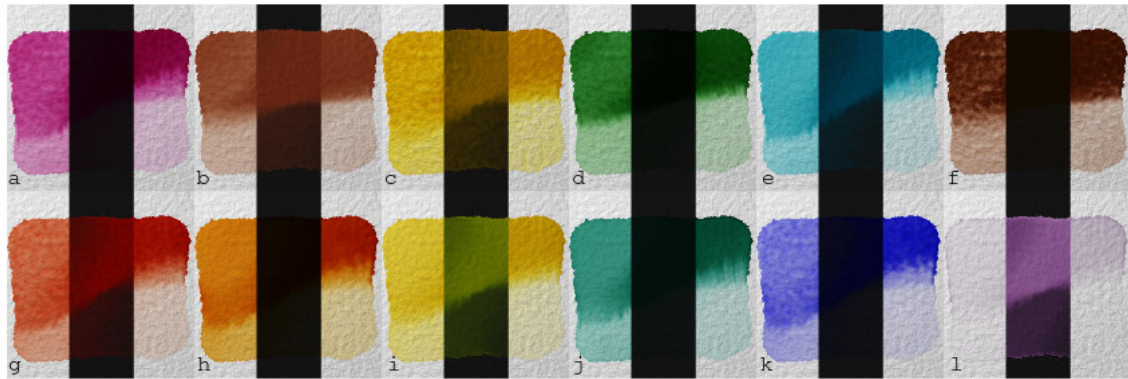


FIGURE 1.3 – Plusieurs pigments synthétiques rendus avec le modèle de réflectance *Kubelka-Munk*. CURTIS et al. [25].

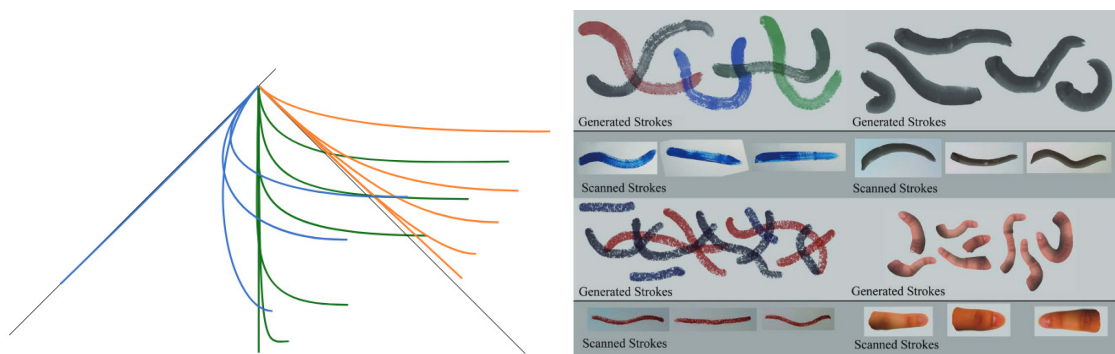
Simulation phénoménologique

La simulation phénoménologique consiste à reproduire un phénomène de manière empirique en utilisant une succession d'opérations numériques. Elle repose principalement sur l'observation et ne permet pas une compréhension approfondie du phénomène.

• **Approches empiriques** La première approche pour simuler la trace d'un coup de pinceau consiste à synthétiser sa texture sous la forme d'une image. Certains travaux initiaux se sont attachés à synthétiser des textures ayant l'apparence d'un coup de pinceau. Par exemple, LEWIS [65] propose d'utiliser le domaine fréquentiel pour générer les détails d'une trace de pinceau. Chaque niveau de détail correspond à un niveau de fréquence. Il crée un spectre fréquentiel, une phase et applique une transformation de Fourier inverse pour synthétiser la texture. Le domaine fréquentiel est largement utilisé dans la synthèse de texture pour produire des détails à différentes échelles. Orienté synthèse de texture par l'exemple, la méthode de RITTER et al. [82] synthétise un coup de pinceau à partir d'une image exemple représentant une empreinte existante d'un coup de pinceau. Cette méthode se base sur des correspondances optimales entre différentes régions d'image de texture et du support. L'impact du médium avec le substrat peut être modélisé par un mélange entre une texture représentant l'empreinte du pinceau et une autre représentant la rugosité du support [97]. Une autre approche consiste à ajouter de manière stochastique un grain à la couleur de l'image en simulant la distribution des fibres dans le cas de la toile ou du papier [57].

• **Approches dirigées par les données** D'autres approches se basent sur des ensembles d'exemples d'un phénomène pour le reproduire. Par exemple, BAXTER et al. [8] proposent de simuler la déformation globale des poils d'un pinceau à

partir de déformations données en exemple (Figure 1.4a). Chaque exemple correspond à une déformation des poils du pinceau en fonction de la pression exercée par le pinceau et de son angle d'attaque. Chaque déformation est enregistrée sous forme de segment de Bézier cubique dans une table de donnée. Les exemples sont ensuite interpolés afin de produire la déformation d'un cas quelconque. Ce type d'approche dirigée par les données est également utilisée pour la génération d'une texture de la trace d'un coup de pinceau, notamment avec ANDO et al. [1], qui proposent d'utiliser une bibliothèque d'images (Figure 1.4b). Ils représentent un coup de pinceau par un assemblage de quadrilatères sur lesquels sont plaquées une ou plusieurs textures représentant une partie de la trace d'un vrai coup de pinceau. Ensuite, ils reconstruisent la trace complète par optimisation numérique dans le but d'obtenir la séquence optimale d'images collant aux paramètres du pinceau virtuel.



(a) Ensemble de données concernant la déformation des poils d'un pinceau. Cet ensemble est ensuite interpolé pour calculer une déformation quelconque [8].

(b) Les coups de pinceau synthétiques sont créés à partir de morceaux de coups de pinceau réels [1].

FIGURE 1.4 – Simulations phénoménologiques dirigées par les données. Images tirées des articles correspondants.

La reproduction du comportement de la peinture, du pinceau au support, peut se faire à l'aide d'un modèle physique ou d'un modèle phénoménologique. Les avancées de la recherche dans ces domaines ont permis aux méthodes d'atteindre des performances en temps réel, ce qui permet d'assister un artiste dans son processus de création artistique. Ces méthodes sont conçues pour offrir aux artistes des outils et des techniques permettant d'obtenir des résultats proches de ce qu'ils auraient pu réaliser manuellement. Parallèlement aux méthodes qui assistent l'artiste dans le processus de création, il existe un autre genre de méthode où la stylisation se fait de manière automatique, sans l'intervention directe d'un artiste.

1.1.2 Stylisation automatique

La stylisation automatique repose sur un ensemble de données pour créer la stylisation d'une image. Ces données varient en fonction de la méthode utilisée. L'utilisateur n'intervient dans la stylisation automatique que lors du paramétrage du système. Ce paramétrage peut s'effectuer à différents niveaux [105], de l'utilisation de variables à l'utilisation d'exemples de stylisation.

Approches à base de marques

Les méthodes à base de marque représentent directement les traces issues de la gestuelle de l'artiste par des marques. Ces marques sont représentées à travers un ensemble de paramètres, pouvant être une texture sous forme d'image ou un ensemble de coefficients utilisés pour rendre la marque. Les marques permettent de représenter les images d'une manière différente en enregistrant les données liées aux gestes artistiques, et offrent un contrôle sur la stylisation grâce des outils manipulant leur forme, leur densité, leur orientation, leur couleur et d'autres caractéristiques artistiques. L'objectif des méthodes à base de marque est de générer un ensemble de marques qui seront ensuite rendues via des modèles de simulation physique ou phénoménologique afin de créer l'image finale, composée de pixels et affichable. La disposition des marques peut s'effectuer en 2D sur un support [47] ou en 3D sur une surface géométrique [75, 89] (Figure 1.6). La disposition et les paramètres des marques sont généralement guidés par une heuristique qui peut être stochastique [44] ou peut tenir compte des attributs et caractéristiques de l'image à styliser [69, 91]. Les marques peuvent être structurées en niveaux de détails, modifiant alors l'ordonnancement général qui détermine l'ordre de rendu des marques [46] (Figure 1.5).



FIGURE 1.5 – Stylisation à base de marques d'une image. HAYS et al. [46].



FIGURE I.6 – Stylisation à base de marques d’une scène 3D. *SCHMID et al. [89]*.

Approches synthèse de texture

L’approche de synthèse de texture regroupe les méthodes qui synthétisent directement des pixels. Les approches d’apprentissage automatique à base de réseaux de neurones dont la sortie est constituée de pixels entrent dans le cadre des approches de synthèse de texture. La génération d’images par ce genre d’approches peut être réalisée en construisant des filtres de stylisation, qui, une fois appliqués à l’image, la transforment dans un style artistique [14]. Ces filtres de stylisation sont élaborés à partir de données extraites de l’image, comme un découpage en régions ou une analyse des contours et des gradients [113]. Les approches de synthèse de texture pour la stylisation automatique contrôlent la stylisation à travers les paramètres algorithmiques qui modifient le filtre appliqué. Cependant, cette approche n’est ni pratique ni intuitive pour un artiste car il ne peut prédire la stylisation aisément. Pour remédier à ce problème, certaines méthodes utilisent un guide donné par l’artiste pour orienter la stylisation d’une image. On parle alors de stylisation contrôlée par l’exemple.

Stylisation par l’exemple

Le contrôle de la stylisation par l’exemple est indépendant du type de méthode utilisée et s’adresse aussi bien aux méthodes de synthèse de texture que celles à base de marques. Les méthodes de stylisation par l’exemple définissent la fonction de style de manière implicite à partir d’un ou plusieurs ensembles de données encodant les informations de style comme des marques ou des images. En synthèse

de texture, la stylisation peut s'effectuer à partir d'un couple d'image $\langle A, A' \rangle$, comprenant une image base A et son correspondant stylisée A' , qui permet d'appliquer la fonction à une autre image de base B afin d'obtenir sa stylisation B' . Ce principe est appelé *Analogie d'images* est introduit par HERTZMANN et al. [48] et est illustré sur la figure 1.7. La fonction de style permettant de passer d'une image à l'autre peut être représentée par un filtre de stylisation calculé par régression [48], puis appliquée à l'image à styliser B pour produire B' . La fonction de style peut aussi être représentée par une fonction de mise en correspondance de voisinage de pixels à travers des mesures de similarité :

Soit entre A et B , la région identifiée dans A est alors copiée depuis A' dans B pour produire B' [35].

Soit entre A' et B lorsqu'une seule image stylisée est utilisée [2].

D'autres méthodes utilisent les réseaux de neurones à convolution et entraînent leur réseau sur de large base d'images à extraire la fonction de style [37, 67] (Figure 1.8).



FIGURE 1.7 – Stylisation d'un couple d'images $B \rightarrow B'$ à partir d'un couple d'images stylisées $A \rightarrow A'$ donné en exemple. HERTZMANN et al. [48].

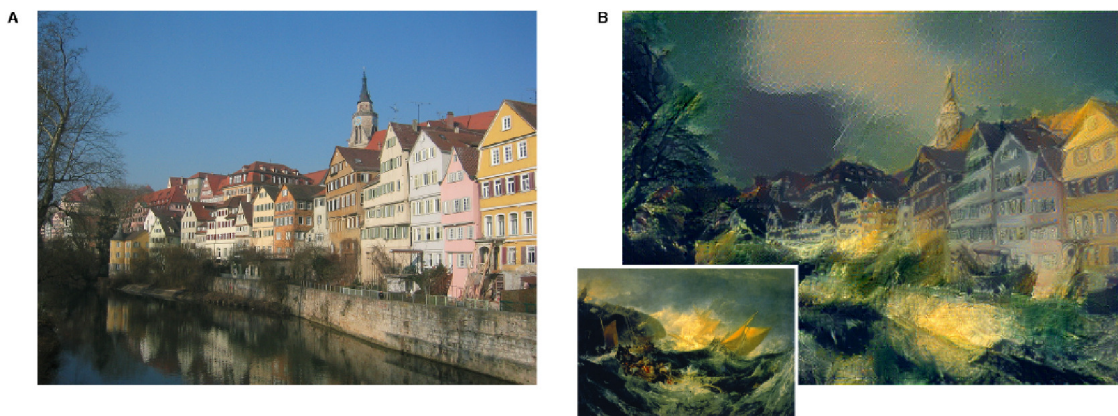


FIGURE 1.8 – Stylisation par l'exemple d'une image avec l'utilisation d'un réseau de neurones. GATYS et al. [37].

Les méthodes à base de marques proposent également un contrôle de la stylisation par l'exemple. Les attributs et caractéristiques des marques peuvent être calculés directement depuis les exemples par correspondance si l'image à styliser et l'exemple possèdent les mêmes structures [121] ou par apprentissage [108]. LU et al. [72] proposent une méthode d'aide au dessin capable de styliser des morceaux

de marque en temps réels en faisant correspondre par optimisation numérique la marque incomplète avec des marques stylisées présentes dans un ensemble de données en exemple.

1.1.3 Stylisation et perception

Certaines méthodes s'intéressent à la perception de l'information contenu par l'image et étudient sa lisibilité. En fonction du type de message que l'image transmet, ces méthodes mettent en avant les détails de l'image qui ont un impact sur sa transmission. Par exemple, GOOCH et al. [40] proposent une méthode qui améliore la visibilité des formes et des structures d'objets 3D grâce à un modèle d'éclairage basé sur les nuances de couleurs. Cette méthode s'inspire des codes artistiques de création d'illustrations, en assombrissant les arêtes et en utilisant une faible dynamique de couleur chaude afin de conserver un contraste élevé permanent entre les arêtes et l'intérieur des formes, quel que soit l'éclairage (Figure 1.9). Dans le même esprit, DECARLO et al. [27] proposent une méthode qui met en évidence les lignes importantes d'une image 2D ou de modèles 3D. Ils utilisent le terme "contour suggestif" pour décrire des lignes qui sont complémentaires aux contours de silhouette et qui permettent de mieux distinguer la forme des objets. Ces lignes correspondent à des contours qui apparaissent comme des arêtes vives depuis un point de vue proche du point de vue courant. La façon dont une scène ou un objet est exposé à la lumière a un impact majeur sur la perception de sa surface et de ses détails. MESTRES et al. [76] montrent qu'en maximisant le contraste local à la surface de l'objet, il est possible d'améliorer la perception des détails de cette surface. Pour cela, ils ajustent la direction de la lumière reçue en tout point de la surface, à différentes échelles et de manière itérative pour chaque niveau de détail, de sorte à préserver le contraste. Il en résulte une direction optimale de la lumière utilisée par le modèle d'éclairage, sublimant ainsi les détails (Figure 1.10).

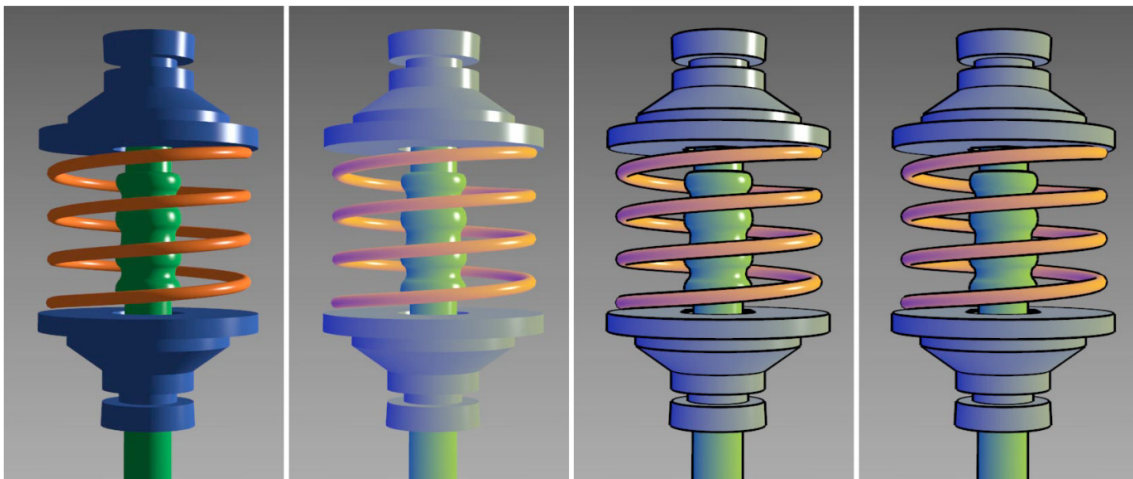


FIGURE 1.9 – Différentes stylisations d'un objet 3D améliorant la visibilité des formes et des structures. Les stylisations sont comparées avec l'objet à gauche, rendu avec le modèle de shading Phong [81]. GOOCH et al. [40].

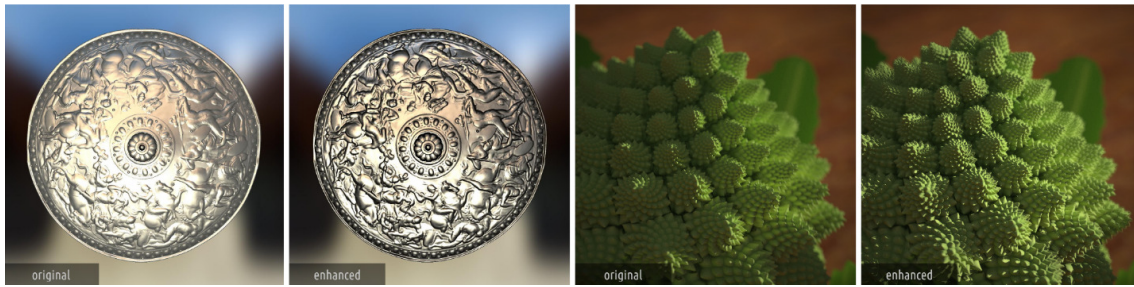


FIGURE 1.10 – Stylisation de deux objets 3D améliorant la perception des détails. MESTRES *et al.* [76].

1.2 Principes de base du rendu à base de marque

1.2.1 Les différents types de marques

Deux types de paramètres sont distingués lorsqu'il s'agit de marques. Tout d'abord, il y a les paramètres qui découlent du geste, qui sont obtenus pendant le processus de dessin. Ils dépendent du degré de liberté de l'outil utilisé et du système d'acquisition. Bien que pouvant être acquis, ces paramètres peuvent être tout aussi être générés procéduralement. Ces paramètres associés à une marque comprennent les positions 2D, la pression et l'orientation de l'outil. Ils serviront ensuite à dériver le deuxième type de paramètres, à savoir les paramètres de rendu. Ces paramètres de rendu permettront à l'ordinateur de rendre la marque et de produire l'image que nous percevons. On calculera ainsi des caractéristiques telles que l'épaisseur, la largeur et la couleur, qui viendront définir la texture et l'apparence visuelle de la marque.

Il existe différents types de marques en fonction de l'application, mais on peut les regrouper en deux catégories principales. Dans sa forme la plus basique, la marque s'apparente à un point. Elle est alors définie par une position (x, y) dans un espace 2D ou 3D le long d'une surface. Les marques prenant la forme de points sont utilisées dans plusieurs styles artistiques, comme le pointillisme (Figure 1.11), où la marque est caractérisée par sa position, sa couleur, son diamètre et sa pression.



FIGURE I.II – Le style pointilliste. *Sous les nuages*, 2022, Nicolas Cattin.

Les marques en forme de points peuvent également servir comme points d’ancrage pour placer un trait simple ou une texture à motifs.

Le plus souvent, les marques représentent des courbes. Chaque marque est alors définie par une fonction : $f : \mathbb{R} \rightarrow \mathbb{R}^n$, avec n étant le nombre de propriétés (i.e. positions, orientation, pression etc...) définies le long de la courbe et x le paramètre de courbe. Ces marques sont couramment utilisées dans le dessin au trait et la peinture, couvrant ainsi un large éventail de styles artistiques, allant du croquis aux dessins d’assemblages, en passant par les styles de peintures acryliques et aquarelles.

1.2.2 Comparaison avec les approches textures

Les approches textures et marques opèrent à des niveaux différents. Les approches textures opèrent au niveau des pixels alors que les approches à base de marques opèrent à partir d’une représentation haut niveau de l’image. Cette représentation abstraite implique généralement une liste de marques regroupant les caractéristiques de l’outil, ce qui permet de modifier les images générées en éditant chacune des marques. Il est bien plus pratique pour un artiste d’ajouter ou de retirer une marque plutôt que d’éditer un ensemble de pixels. Les méthodes textures ne facilitent pas une édition locale de l’image car elles utilisent des paramètres de génération communs pour toute l’image. Il est alors nécessaire de relancer la génération, ou de passer par un outil d’édition externe à la méthode tel que Gimp [99].

En ce qui concerne le rendu et la reproduction d’un style, les méthodes à base de marques restent limitées. Elles parviennent très bien à reproduire un style donné en exemple, à condition que celui-ci soit également réalisé à base de marques et que sa structure soit accessible. Les méthodes à base de synthèse de textures sont nettement plus performantes et reproduisent plus fidèlement un style donné en

exemple. Néanmoins, lorsque le style à reproduire présente des traces d'outils visibles, comme la trace des poils d'un pinceau sur le support, les méthodes de synthèse de texture peuvent créer des discontinuités visibles dans la stylisation. Pour la reproduction d'un style, la performance des méthodes à base de marques est nuancée.

En matière de cohérence temporelle, chacune des deux approches est confrontée aux mêmes problèmes. Les méthodes basées sur les marques peuvent suivre l'évolution des marques au cours du temps, tout comme les méthodes de synthèse de texture peuvent suivre l'évolution des pixels.

Les méthodes d'apprentissage automatique basées sur les réseaux de neurones accentuent ces différences. La très grande majorité de ces méthodes fonctionnent dans l'espace image et possèdent les mêmes problèmes d'édition et de contrôle. À cela s'ajoute l'effet boîte noire et le temps d'entraînement du réseau.

1.3 L'animation traditionnelle 2D

Cette section met en évidence les problématiques de génération d'animation 2D qui se décomposent en deux parties distinctes : la création du mouvement et la stylisation. Les contraintes et les données disponibles varient suivant le type d'animation stylisée : une animation 2D, 3D ou une vidéo.

1.3.1 Application à l'animation

La stylisation de chaque image d'une séquence d'animation de manière indépendante conduit à une stylisation discontinue dans le temps. C'est pourquoi les méthodes de stylisation d'animations utilisent une information liant le contenu d'une image à une autre. Cette information correspond au mouvement de l'animation représenté dans l'espace image par un champ vectoriel indiquant le déplacement d'image en image, en avant ou en arrière d'un ou plusieurs pixels. Il est obtenu par l'estimation du flux optique depuis une vidéo, par le rendu d'une scène 3D, par le dessin ou de manière procédurale. On peut également le représenter sous forme de transformations opérant sur des images ou des éléments de l'image permettant de passer d'un état à un autre. Ces transformations sont obtenues par le suivi temporel de points d'intérêt dans les images de l'animation et plus généralement par les techniques du domaine de la vision par ordinateur.

Les animateurs utilisent des techniques d'animation pour altérer la perception du mouvement d'une forme, comme par exemple l'exagération de mouvement. Prenons l'exemple de la figure 1.12. La silhouette du personnage est déformée en fonction de l'action décrite par la scène ce qui donne une impression d'accélération et d'intensité. Le message transmis par l'animation est différent de celui qu'il aurait été si la silhouette du personnage était dessinée avec des proportions normales et correspond à une impression de vitesse.

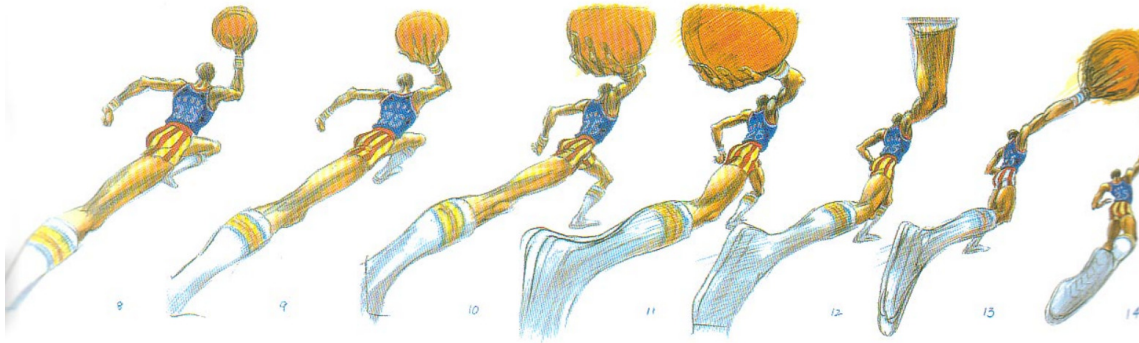


FIGURE 1.12 – Mouvements exagérés suggérant l'intensité du mouvement. *The Animator's Survival Kit* - Richard Williams [112]

Les animateurs réalisent ce type d'effet en manipulant ce que l'on appelle le *timing and spacing*. Le terme *timing* correspond au fait de déterminer les moments importants de l'action, comme les collisions d'un objet avec une surface, et le terme *spacing* correspond au positionnement des images dans la progression du mouvement. Les animateurs utilisent les concepts de *timing and spacing* dans la phase de conception de l'animation puis le *spacing* est affiné tout au long du processus. Lors de la réalisation des animations, les artistes utilisent le *spacing* pour modifier la répartition des images dans la séquence afin de créer un effet d'accélération ou de décélération du mouvement (Figure 1.13).

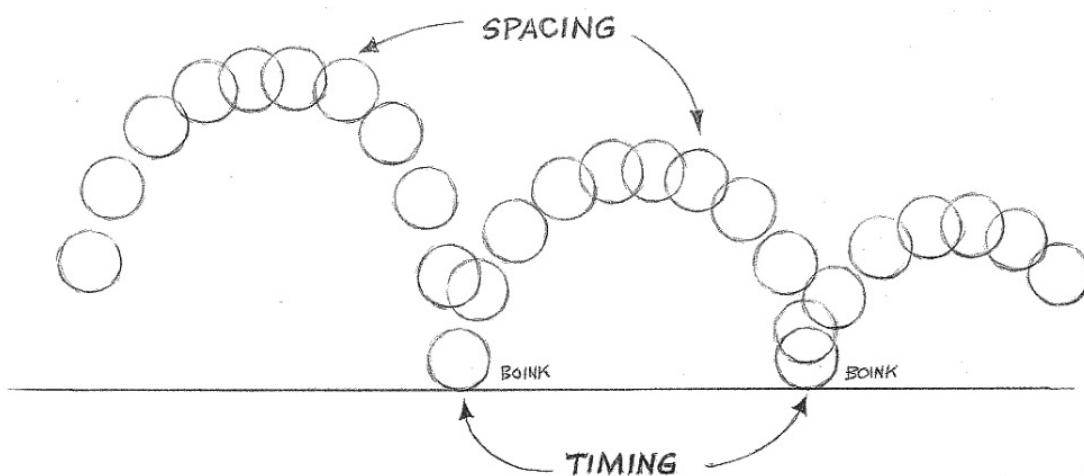


FIGURE 1.13 – Principe du *Timing and spacing*, *The Animator's Survival Kit* - Richard Williams [112]

1.3.2 Les problématiques liées à la stylisation de l'animation

La manière dont les méthodes de stylisation abordent la génération du mouvement contribue à définir le flux de travail de l'animateur et façonne son environnement de production. Le traitement du mouvement dans les techniques de stylisation soulève plusieurs problématiques :

- **Cohérence temporelle** La cohérence temporelle se définit par un compromis entre la cohérence du mouvement, la continuité temporelle et l'apparence 2D de

la stylisation. La cohérence du mouvement consiste à se demander si la stylisation suit correctement le mouvement de la scène. Lorsque le niveau de cohérence du mouvement est nul, la stylisation reste inchangée d'image en image, créant ainsi un effet de calque où des morceaux de la stylisation restent fixes et d'autres mobiles. Cet effet visuel peut être comparé à l'observation d'une scène à travers une vitre texturée. La continuité temporelle concerne la question de savoir si les changements de stylisation d'une image à l'autre se font de manière cohérente. Lorsqu'elle est faible, voir absente, des changements brusques de stylisation apparaissent, créant un effet visuel de scintillement. Une fois stylisées, les animations 2D ne présentent pas toutes le même niveau de cohérence temporelle.

Les animations avec une stylisation présentant un haut niveau de cohérence de mouvement et de continuité temporelle ont tendance à s'éloigner de l'apparence 2D. Ce constat s'explique par le fait que les animations 2D traditionnelles ne présentent pas, à l'origine, de tels niveaux de cohérence. Lorsque des images dessinées à la main mise bout à bout forme une animation, des scintillements sont perceptibles, dus aux changements discontinus de la stylisation d'une image à l'autre. BÉNARD et al. [12] illustrent ce phénomène en représentant par un diagramme les objectifs des méthodes de stylisation d'animations : la recherche de l'apparence d'une stylisation 2D, la recherche d'une forte cohérence de mouvement et la recherche d'une forte continuité temporelle (Figure 1.14). Ils mettent en évidence le compromis fait par les méthodes de stylisations, qui ne peuvent pas atteindre parfaitement ces trois objectifs.

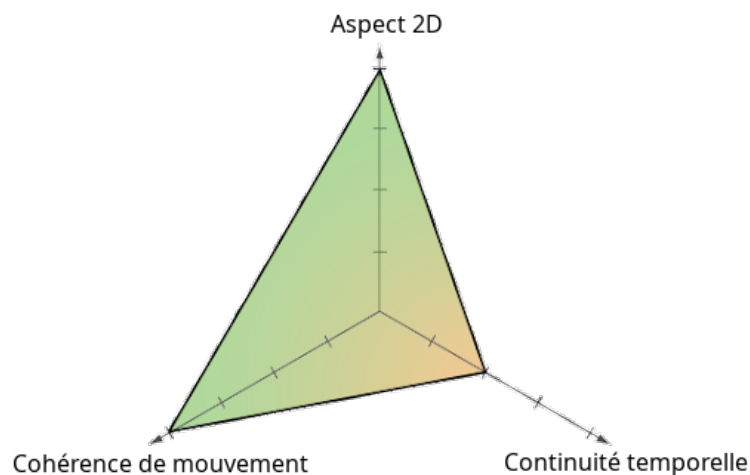


FIGURE I.14 – Diagramme montrant le compromis de la cohérence temporelle pour une méthode de stylisation. BÉNARD et al. [12].

• **Les occlusions** L'occlusion (ou occultation) désigne le fait qu'une partie de la scène est cachée par une autre placée devant par rapport au point de vue. Lors d'un changement de point de vue ou dans le cas d'une scène animée, les changements d'occlusions peuvent provoquer des artefacts lors de la stylisation. En effet, ces changements de visibilité posent problème lorsque la stylisation introduit des changements en espace image plus visibles que les changements de visibilité. Par exemple en rendu à base de marques, les marques ajoutées pour représenter une partie de la scène qui devient visible au cours de l'animation peuvent provoquer un

changement brusque plus notable que celui auquel on pourrait s'attendre. Des méthodes de stylisation abordent ce problème dans l'espace 3D se servant ainsi des relations de profondeur entre éléments de la scène pour anticiper et lisser temporellement les changements d'occlusion. Cependant les stylisations produites ne conservent pas leur apparence 2D au cours du temps. D'autres optent pour une stratégie hybride définissant la stylisation d'objets 3D depuis l'espace image mais cela nécessite de trouver un compromis entre une stylisation qui suit le mouvement 3D et son apparence 2D [15]. Lorsque les stylisations opèrent dans l'espace image, les méthodes de stylisation s'orientent vers des approches interactives où l'animateur apporte les informations qui permettent de traiter les occlusions [54]. D'autres méthodes ne traitent pas directement le problème d'occlusion et offrent la possibilité aux animateurs de le résoudre par l'utilisation de calques superposables, comme ce qui est fait dans les techniques traditionnelles.

1.3.3 Automatisation des techniques traditionnelles

L'animation traditionnelle regroupe toutes les techniques d'animation utilisées sans l'intervention de l'ordinateur dans ses procédés. Elle est définie par l'apparence et la qualité de ses animations, qui présentent des défauts visuels spécifiques aux problèmes de cohérence temporelle. Ces défauts sont directement liés au fait que toutes les images des animations traditionnelles sont créées à la main.

Les méthodes de travail

Deux techniques principales sont utilisées pour la création d'animations traditionnelles :

La première est le *pose à pose*, qui consiste à définir une série de poses d'un personnage à des points clés de la séquence, puis à les raffiner jusqu'à obtenir l'ensemble des poses intermédiaires. Quatre niveaux de poses sont proposés par WILLIAMS [112]. Le premier, les poses clés, sert à définir ce qui doit être animé. Le second niveau, les poses extrêmes, représente les positions les plus importantes ou les faits marquants d'un mouvement. Le troisième niveau correspond aux poses de rupture ou de passage. Ces trois premiers niveaux de poses servent à décrire correctement le *timing and spacing* de l'animation sans ambiguïté. Ensuite, l'artiste termine en créant les poses de transition entre chaque pose d'un niveau supérieur, jusqu'à atteindre le nombre d'images par seconde souhaité. Mises bout à bout, ces poses donnent alors l'illusion du mouvement (Figure 1.15). Selon le style artistique utilisé et la manière de créer l'animation, le terme *poses clés* fait référence aux *images clés*. C'est la technique la plus répandue car elle permet de prévisualiser et de structurer la création du mouvement par affinages successifs.

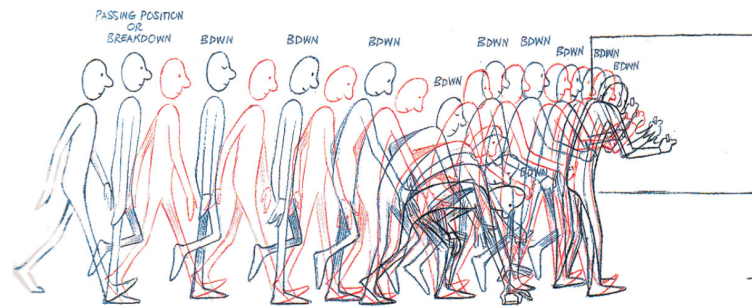


FIGURE I.15 – Animation d'un personnage par la technique du pose à pose. Trois niveaux de poses sont représentés ici : Les poses clés en noir, les poses extrêmes en rouge et les poses de passage en bleu. *The Animator's Survival Kit* - Richard Williams [112].

La deuxième technique est celle où l'animation se fait d'image en image, ou *straight ahead*, elle est connue sous le nom : *sous la caméra*. L'artiste crée une première image et la capture sur pellicule. L'image suivante est obtenue par de légères modifications de la première, et ainsi de suite pour toute la séquence. Chaque image de l'animation n'existe que lorsqu'elle est capturée, puis est détruite pour produire l'image suivante. En conséquence, l'artiste doit anticiper toute la séquence, il n'y a pas d'option pour revenir en arrière. Cette technique est appréciée des artistes spécialisés dans la peinture sur verre car elle répond aux contraintes matérielles du médium (Figure 1.16). Bien qu'elle tende à rendre la création du mouvement moins structurée, elle reste une technique intuitive et laisse place à la créativité par sa spontanéité [112].



FIGURE I.16 – La technique *sous la caméra*, principalement utilisée par les animateurs de peinture sur verre - *Traces de Hugo Frassetto*

Les animateurs peuvent combiner ces deux techniques pour profiter de leurs avantages respectifs en utilisant les trois premiers niveaux de pose comme guide pour dessiner *sous la caméra* image par image l'intégralité de la séquence avec la possibilité de ne pas respecter les poses précédemment établies. Ils profitent ainsi de la planification du mouvement grâce à la réalisation des poses clés, des poses extrêmes et des poses de passage, tout en bénéficiant du flux créatif en réalisant une image après l'autre.

La nouvelle méthode que je propose, et qui sera développée au cours de ce manuscrit de thèse, reprend les concepts des techniques *pose à pose* et *sous la caméra* et proposent une nouvelle approche hybride utilisant l'ordinateur.

Les différents styles

Les artistes créent les animations traditionnelles avec un ou plusieurs médiums. De manière non exhaustive, on retrouve le dessin animé avec des médiums comme l'encre, le crayon, la peinture acrylique, à l'eau ou à l'huile, ou encore le pastel. Certains artistes combinent différents médiums pour définir leur style, comme par exemple *Isao Takahata* qui réalise le film d'animation *Le Conte de la princesse Kaguya* en utilisant deux calques (Figure 1.17). Le premier lui permet d'utiliser le crayon pour réaliser les esquisses de chaque image de l'animation, et le second correspond aux calques de colorisation peints à la main.



FIGURE I.17 – Animation traditionnelle combinant la peinture et le dessin au trait. *Le conte de la princesse Kaguya* de *Isao Takahata*

Les enjeux de l'automatisation

Mes travaux se concentrent sur la génération d'animations 2D visant à reproduire, par le biais de l'ordinateur, l'esthétique des animations traditionnelles, tout en offrant un niveau de contrôle au moins équivalent à celui que proposent les techniques traditionnelles.

L'introduction de l'ordinateur dans les techniques traditionnelles de l'animation peut être bénéfique à plusieurs égards. La création d'animations traditionnelles comprend deux types de tâches distinctes : d'une part, celles qui permettent l'expression artistique, et d'autre part, celles plus fastidieuses qui consistent à répéter et reproduire des formes et motifs existants. L'un des enjeux majeurs de l'automatisation que j'identifie dans ce contexte est de rendre plus efficaces ces tâches fastidieuses, ce qui permettrait ainsi de gagner un temps précieux dans le processus créatif. En déchargeant les artistes de ces tâches répétitives, ils se concentreront davantage sur leur expression artistique et pourrons donner libre cours à leur

créativité. Les méthodes de stylisation d’animation peuvent contribuer également à enrichir les techniques de travail traditionnelles, en offrant de nouvelles fonctionnalités d’édition leur permettant d’expérimenter plus librement. Néanmoins elles peuvent s’éloigner de la manière dont les artistes travaillent aujourd’hui ce qui peut remettre en cause l’acceptation de ces méthodes dans un contexte de production professionnelle. Cette inquiétude est confirmée par l’animateur 2D professionnel du secteur, *David Martin*, que nous avons rencontré le 19 janvier 2022. La méthode de création artistique doit rester similaire à ce qui se fait manuellement, tant du point de vue de l’utilisation que du rendu. L’enjeu qui en découle est l’acceptabilité des méthodes proposées. Un autre enjeu essentiel de ces méthodes est leur capacité à reproduire fidèlement l’apparence et le rendu des animations traditionnelles, en y incluant les défauts et qualités qui leur donnent ce charme si particulier. La génération de ce type d’animation fait face aux nombreux styles et médiums composant le spectre artistique.

Les critères qui différencient les méthodes de génération d’animations 2D entre elles sont : le gain de temps, les possibilités techniques, l’acceptabilité, la qualité de reproduction de l’aspect traditionnel et la variété des médiums pris en charge. Ces critères permettent d’évaluer la méthode que je propose et de la comparer par rapport aux autres méthodes.

1.4 Contributions de la thèse

De nos jours, il existe beaucoup d’outils qui permettent de faciliter la création d’animations 2D. La plupart sont destinés à la production de dessins animés et utilisent l’approche *pose à pose*. Cependant, l’animation traditionnelle ne se limite pas au style cartoon. Elle englobe une grande variété de styles tels que la peinture, le pastel, et bien d’autres. Ces styles comportent davantage de textures, de détails, et chaque coup de crayon ou de pinceau est visible par rapport aux animations cartoon. Ils nécessitent donc des outils spécifiques qui sont parfois inexistantes. Je réponds aux enjeux évoqués à travers mes travaux afin d’aider les artistes à produire des animations 2D reproduisant l’aspect des techniques traditionnelles de manière plus efficace. Ces techniques d’animation impliquent que l’artiste dessine toutes les images individuellement, ce qui représente un temps de travail considérable.

Mes travaux se sont naturellement orientés vers la création d’animations traditionnelles de type peinture. Il existe très peu de méthodes pour produire ce genre d’animation par ordinateur qui respectent les techniques de création manuelles et qui reproduisent leur apparence. Je propose une nouvelle méthode de génération par ordinateur, permettant à l’animateur d’automatiser les tâches les plus fastidieuses dans la réalisation d’animations sans restreindre sa liberté et ses possibilités artistiques. Cette méthode allie les techniques du *pose à pose* et *sous la caméra*, et propose un rendu et une apparence comparable aux animations où toutes les images sont faites à la main. Elle permet à l’artiste de dessiner des images clés et de propager le contenu dans la séquence à l’aide d’un mouvement pour créer de

nouvelles images. Le contenu est propagé entre deux images clés et conserve les éléments du style de ces deux images.

La méthode proposée permet de lever les verrous techniques de la génération d'images intermédiaires à base de marques. Mon intérêt se porte sur les problématiques induites par l'utilisation du mouvement pour la propagation des marques des images clés, notamment comment associer un mouvement aux marques pour pouvoir les déplacer, et comment contrôler la forme des marques une fois altérées par le mouvement. Je mets en lumière les contraintes structurelles qu'implique l'utilisation d'un modèle physique de simulation de peinture sur l'ordonnement des marques provenant de deux images clés. De plus, je m'intéresse également à la reproduction des niveaux de cohérence temporelle des techniques traditionnelles à travers l'élaboration d'un algorithme de sélection de marques. Je mets les travaux en perspective en portant un intérêt à l'impact du temps de calcul sur le flux de travail de la méthode, aux différentes manières d'obtenir le mouvement et à la dépendance aux données que cela implique, ainsi qu'à la qualité des résultats que nous comparons avec ceux des techniques manuelles et traditionnelles.

Sur un plan technique, je développe quatre contributions :

La première contribution est la création d'un champ vectoriel global qui définit le mouvement pour tout pixel de l'image à partir d'un champ vectoriel incomplet. Elle permet de ne plus limiter l'advection d'un objet à la forme représentée par les contours du mouvement utilisé. L'intérêt de cette procédure est illustrée par le déplacement des coups de pinceau sur le support, en fonction du mouvement, permettant ainsi de passer d'une image à une autre de manière cohérente.

La seconde est une méthode de correction de courbe 2D, sous forme d'optimisation numérique, qui préserve les propriétés géométriques et la sémantique des coups de pinceau transformés par un mouvement. Cette contribution permet la réutilisation des coups de pinceaux présents dans les images clés pour la stylisation des images intermédiaires.

La troisième est une méthode permettant d'ordonner de manière cohérente deux ensembles de coups de pinceau transformés, chacun provenant à la base d'une image clé. Elle permet de réduire les interférences entre les deux ensembles en tenant compte des contraintes induites par la simulation de peinture. L'intérêt de cette méthode est illustré par la création d'un ensemble de coups de pinceau candidats pour remplir une image intermédiaire, à partir des propagations en avant et en arrière des coups de pinceau des images clés.

La dernière contribution est une méthode permettant de générer des images intermédiaires respectant une quantité de matière spécifique sur le support par composition et sélection de coups de pinceau candidats. Elle permet d'atteindre un niveau de cohérence temporelle et une apparence proches des techniques traditionnelles.

Chapitre 2

Fondements théoriques

Ce chapitre présente les bases théoriques liées à mes travaux de thèse. Il comprend un état de l’art axé sur les méthodes d’animation 2D, qu’elles soient traditionnelles ou non, une section décrivant le rendu à base de marques, et notre positionnement scientifique. De plus, ce chapitre permet d’approfondir les enjeux de la création semi-automatique d’animations 2D.

2.1 État de l’art

2.1.1 Le mouvement

Dans une animation, le mouvement est une illusion créée par une succession d’images fixes. Le mouvement est ainsi perçu comme la mise en relation du contenu de l’animation dans le temps. Cette perception du mouvement, que notre cerveau interprète, est le « mouvement apparent ». Les méthodes de génération d’animations 2D par ordinateur définissent le mouvement comme une transformation pouvant être représentée en 2D par des champs de vecteurs ou plus généralement par une fonction mathématique telle qu’une rotation ou translation.

• **Sources de mouvement** Certaines méthodes de stylisation d’animations prennent directement une description du mouvement comme entrée. BÉNARD et al. [13] utilisent en entrée un champs vectoriel extrait depuis une animation 3D représentant le mouvement d’objets animés projetés en 2D. L’exploitation des données 3D produit un mouvement précis et cohérent mais implique cependant la création d’une animation 3D sous jacente qui ajoute une contrainte supplémentaire. Les champs vectoriels peuvent aussi être créés de manière procédurale comme le montre BRIDSON et al. [16] qui utilisent la simulation d’un phénomène physique pour produire une carte de mouvement. Le mouvement peut aussi être dessiné. HU et al. [50] proposent une méthode capable de générer un champ vectoriel de fluide à partir d’un dessin. La méthode utilise un réseau de neurones entraîné à reproduire le mouvement de fluide à partir de traits.

Ces deux derniers types de méthodes sont limités par la complexité des mouvements générés. Il est difficile de décrire des mouvements complexes, tel que le mouvement de la marche d’un personnage, sans passer par des données 3D. En

effet, le problème d'occlusion fait nécessairement référence à la notion de profondeur de la scène et survient lorsqu'il y a un changement de visibilité d'éléments de la scène. C'est pourquoi ces types de méthodes se limitent à des cas d'utilisation simple.

Mouvement correspondances / matching

Un autre moyen de parvenir à reproduire le mouvement apparent est d'établir des correspondances entre deux images et d'estimer un mouvement qui assure la transition entre les deux images. Ce type d'approche soulève deux problématiques principales. Tout d'abord, les méthodes déterminent comment réaliser ces correspondances et identifient les problèmes qui surviennent. Ensuite, elles résolvent le problème de l'estimation du mouvement par le choix de la formulation, la représentation et le calcul. Les correspondances de pixels, de régions, de formes ou de caractéristiques dans l'image peuvent être réalisées par les deux types d'approches suivantes.

• **l'approche basée image :** l'estimation de flux optique à partir d'une vidéo permet d'obtenir un champs de mouvement. Le but de cet état de l'art n'est pas de faire l'inventaire de toutes ces méthodes mais de les mettre en perspective avec notre contexte de génération d'animations. ZHAI et al. [120] proposent un état de l'art sur les différentes façons d'aborder le calcul de flux optique. Avec l'émergence des méthodes d'apprentissage automatique, l'estimation de mouvement est également faite à l'aide de réseaux de neurones [88]. LI et al. [68] utilisent un croquis représentant l'état intermédiaire entre deux images clés pour calculer le flux de mouvement. Le principe est de transformer le croquis et les images clés dans un espace de paramètres commun à l'aide d'un réseau de neurones de type *transformers*, puis d'éliminer les espaces blancs du croquis. Ensuite un nouveau réseau de neurones permet d'extraire le flux de mouvement de l'espace des paramètres et fait la correspondance entre les images transformées du croquis et de chacune des images clés.

Les méthodes se basent sur le fait que les deux images mises en correspondance soient très proches ou successives dans le temps, ce qui n'est pas le cas lorsque l'on cherche à estimer le mouvement entre deux images clés dans un contexte de création pose à pose (1.3.3). De plus, les images utilisées pour estimer le mouvement varient en fonction du style artistique. Selon les méthodes, certains éléments du style peuvent entraver le calcul du mouvement en réduisant les correspondances possibles et la qualité du mouvement estimé.

• **l'approche basée marque :** lorsque l'image est définie par des marques, on peut établir des correspondances marque à marque entre deux images pour estimer un mouvement. Cependant, il n'y a pas de garantie d'avoir le même nombre de marques entre les deux images. Certaines correspondances peuvent être erronées ou des marques peuvent ne pas avoir de correspondances. Pour résoudre ce problème, il est possible de regrouper les marques d'un dessin ou d'une peinture sous forme de graphes ou de chaînes et faire ainsi apparaître des relations entre les structures. Ces structures définissent alors une caractéristique ou une forme de

l'image et sont construites en utilisant un ensemble de règles topologiques. La mise en correspondance de graphes est faite par élimination à la suite de vérifications de règles [61] ou à travers une descente de graphe où une fonction de similarité est mise à jour à chaque déplacement dans le graphe [109]. Les deux méthodes échantillonnent les marques pour garantir un nombre similaire de points et puis interpolent ces points pour produire de nouvelles marques représentant le mouvement point à point d'un graphe à l'autre.

Des descripteurs de forme sont utilisés pour établir des correspondances de régions entre deux images tel que la distance de Hausdorff. HUDON et al. [52] utilisent la méthode des *contextes de formes* de BELONGIE et al. [10] pour déterminer si deux marques de deux images clés différentes représentant des contours, appartiennent à la même région. Le principe des contextes de formes repose sur des comparaisons d'ensemble de points auxquels pour chaque point est associées des informations concernant son voisinage. La méthode est limitée aux contours proches de la forme à faire correspondre pour pallier et réduire l'impact des occlusions de l'animation sur les résultats. Une multitude de descripteurs de forme peuvent être utilisés [56].

Ces méthodes sont efficaces pour établir des correspondances lorsque les marques représentent des contours mais ne permettent pas de résoudre les cas où les marques sont utilisées pour décrire l'intérieur d'une forme avec toutes ses nuances de remplissage. Elles sont adaptées au dessin au trait lorsque que le contour est défini par un trait propre et clair. Elles ne résolvent pas parfaitement le problème de correspondance de marque de manière automatique et créent de mauvaises associations entre marques. C'est pourquoi la plupart de ces méthodes sont guidées par un utilisateur.

•**Transformation et Déformation** Lorsqu'il n'est pas possible de déterminer le mouvement de tous les pixels, les correspondances seules de pixels clairsemés, entre deux images successives, ne permettent pas de générer un mouvement réaliste. À partir des correspondances établies, il est possible de calculer une transformation affine en résolvant un système d'équations linéaires. Cette transformation affine peut ensuite être interpolée et appliquée à l'image initiale pour créer un effet de mouvement [115]. Cependant la transformation affine ne permet pas de représenter les déformations les plus complexes causées par le mouvement, telles que les effets d'étirement largement utilisés par les artistes pour accentuer la perception de la vitesse dans un mouvement. Dans le cas où les correspondances sont des régions de pixels, l'utilisation d'un maillage permet d'estimer une déformation complexe permettant de passer d'une région à une autre. LEE et al. [63] représentent les déformations de type étirement ou torsion sous forme quadratique grâce à l'analyse modale. Des paramètres physiques sont associés au maillage et appliquent des forces à chaque sommet qui déforment le maillage. Au lieu de formaliser mathématiquement une déformation complexe, SÝKORA et al. [98] proposent de la discrétiser en une succession de transformations rigides locales. Une grille régulière englobe la région, chaque cellule est ensuite déformée de manière itérative en préservant la rigidité de la forme sous-jacente ce qui permet d'assurer la consistance du dessin.

Les méthodes d'alignement d'images utilisées pour la génération d'animations ne parviennent pas à gérer les occlusions lorsque la structure du maillage est maintenue. Le problème se pose lorsqu'il est nécessaire de modifier la structure du maillage, pour résoudre les problèmes d'occlusions, car il faut alors garantir la cohérence des connexions dans le nouveau maillage. Elles sont également dépendantes de la résolution du maillage utilisé pour décrire la déformation finement, ce qui peut induire un coût de calcul élevé.

Stylisation par le mouvement

Le mouvement joue un rôle essentiel dans la stylisation de l'animation, offrant la possibilité de créer des effets visuels uniques et distinctifs qui renforcent le style artistique de l'animateur [112]. Nous avons vu que le mouvement se représente sous la forme de vecteurs décrivant une advection ou sous la forme de transformations permettant de passer d'une image à l'autre. La stylisation par le mouvement consiste à modifier les valeurs de la transformation ou des vecteurs pour obtenir un effet artistique. En s'appuyant sur ce concept, les mouvements présents dans une vidéo peuvent être déformés [63] ce qui permet de souligner une caractéristique du mouvement telle que la vitesse et de donner un aspect cartoon à la vidéo. WANG et al. [107] proposent un filtre permettant de créer un effet d'anticipation et de suivis de formes animées renforçant également l'aspect cartoon du mouvement. Ce filtre correspond à l'inverse du filtre de Marr [74] et peut être généralisé en 2D dans le cas d'une vidéo où l'on vient utiliser un maillage et convoluer chaque vertex. LIU et al. [70] proposent de mettre en évidence des mouvements invisibles ou non perceptibles dans des vidéos en les amplifiant. Pour cela ils identifient des groupes de pixels ayant un mouvement similaire, associent un coefficient amplifiant leur mouvement et reconstruisent la vidéo à partir des groupes formés. Les occlusions occasionnant une superposition sont gérées par un ordonnancement en couche des groupes de pixels partageant un mouvement similaire, et les trous sont complétés par de la synthèse de pixels.

Le mouvement est aussi une caractéristique fondamentale de la cohérence temporelle de chaque animation [12]. Une forte cohérence de mouvement peut donner l'impression que l'animation est construite à partir de scènes 3D. C'est pourquoi certaines méthodes mettent l'accent sur l'aspect *fait à la main* du rendu des animations. DELANOY et al. [28] proposent d'augmenter la rigidité apparente du mouvement pour renforcer l'aspect 2D (i.e. fait à la main) de vidéos stylisées. Chaque image de l'entrée vidéo est segmentée en régions où une trajectoire est calculée représentant le mouvement de chaque région. Ensuite les trajectoires des pixels appartenant aux régions sont modifiées en fonction des trajectoires calculées pour chaque région et de son voisinage. La vidéo est à nouveau rendu avec le nouveau mouvement optique obtenu. Cette méthode permet de réduire la cohérence de mouvement de l'animation en la rigidifiant pour lui permettre de se rapprocher des niveaux de cohérence de mouvement de celles faites à la main.

Le mouvement est donc une caractéristique fondamentale pour la réalisation d'une animation. Il est représenté par des vecteurs ou des transformations et impacte la stylisation de part sa cohérence. La création d'animations 2D ne se limite

pas seulement au mouvement, mais implique également la stylisation.

2.1.2 Stylisation basée synthèse de texture

Le périmètre de la stylisation basée synthèse de texture part de la stylisation de vidéos jusqu'à la création d'images, dans un médium artistique spécifique. La stylisation par ordinateur peut être faite de manière automatique ou guidée par l'artiste. Le style artistique peut être donné en exemple ou généré de toute pièce. On retrouve principalement deux types de méthodes, les approches basées synthèse de texture et les approches à base de marques. La synthèse de texture est un domaine qui regroupe les méthodes les plus performantes pour résoudre les problèmes de transfert de style d'une image stylisée à une autre image qui ne l'est pas. Cependant, les méthodes de transfert de style se confrontent au problème de la cohérence temporelle lorsqu'elles sont appliquées à la stylisation d'animations.

• **Analogie d'images, basée tuile** HERTZMANN et al. [48] introduisent le principe d'«Analogie d'images». Appliqué directement une séquence d'images, la méthode produit un scintillement perceptible au cours de l'animation. HASHIMOTO et al. [45] proposent de le résoudre en définissant une règle basée sur le mouvement des pixels tout au long de l'animation. Le voisinage de chaque pixel permet d'établir un coefficient de cohérence propre à ce pixel indiquant s'il y a de grands changements entre deux images successives. En fonction d'un seuil de cohérence, le pixel est simplement copié depuis l'image précédente ou recalculé depuis l'exemple définissant la stylisation. BÉNARD et al. [13] reprennent le concept d'analogie d'images et l'applique à la création d'animation pose à pose. Ils remplacent l'algorithme par celui de BARNES et al. [3] qui consiste à mettre en relation des morceaux d'images en comparant des mesures de similarité du voisinage de chaque morceau. Il permet de formaliser le problème en une optimisation numérique cherchant à faire correspondre des régions de pixels de l'image stylisée avec celle où la stylisation doit être appliquée. Plutôt que d'utiliser uniquement les données RGB des images, ils guident la stylisation de l'animation sous jacente par l'utilisation de textures de vitesse représentant le mouvement des pixels, et introduisent un terme sur les occlusions et sur l'orientation de la stylisation pour le contrôle de la cohérence temporelle.

Du point de vue de la synthèse de textures, ces méthodes souffrent de problèmes de distorsion dans le cas où des motifs sont visibles dans la stylisation, ce qui est le cas pour toute stylisation où les marques sont apparentes comme par exemple avec le style peinture. Ceci s'explique par le fait qu'aucune contrainte n'est mise en place pour en assurer la cohérence, ou en préserver la logique lors de la reproduction du style. En résultent également l'apparition de détails ou motifs non présents dans la stylisation d'origine. Pour limiter ces effets, FIŠER et al. [35] contribuent à améliorer l'approche en favorisant un usage unique des couple de tuile (i.e. groupe de pixels) provenant de la stylisation. Ils proposent également de se servir de données relatives à l'éclairage d'une scène 3D pour mieux guider la reproduction du style en fonction d'un éclairage cible, dans l'esprit de SLOAN et al. [93]. Ils montrent ainsi qu'avec les bonnes textures pour guider la synthèse de texture il est possible de conserver la finesse des détails de la stylisation.

FIŠER et al. [36] proposent une méthode basée sur la précédente [35] et démontrent son application à la stylisation de vidéo de portrait. Ils guident la recherche de tuiles par l'utilisation de textures calculées directement depuis le flux optique vidéo. Partant du principe qu'un artiste utilise une unique stylisation à chaque région, ils génèrent une segmentation des images vidéos et de l'exemple. Mises en correspondance, ces régions permettent de restreindre la recherche de tuiles de stylisation. Cependant leur méthode souffre d'un manque de diversité de tuiles disponibles pour la stylisation d'une région et ne parvient pas à recréer le rendu des animations traditionnelles. Ils limitent cet effet en lissant la stylisation au niveau de la recherche de tuiles par l'utilisation de tuiles flous ce qui a pour effet de créer un scintillement de la stylisation au cours de l'animation. Ceci est motivé par le constat de FIŠER et al. [34] qui suggèrent que la cohérence temporelle des animations traditionnelles n'est préservée que dans les basses fréquences de l'image.

Les méthodes présentées reposent sur des hypothèses concernant le type de données utilisées. Les textures guidant la génération sont calculées depuis la géométrie 3D pour BÉNARD et al. [13] ou supposent que le flux vidéo représente un portrait ce qui permet d'extraire la déformation du visage dans le cas de la méthode de FIŠER et al. [36].

JAMRIŠKA et al. [53] proposent une méthode permettant de mieux reproduire le rendu d'animations traditionnelles sans utiliser d'hypothèses sur les données. Les textures guides sont générées directement depuis le flux vidéo sans formaliser ce qui est représenté. Ils gardent les mêmes textures mais les génèrent différemment à l'exception d'une nouvelle qui représente les arrêtes de l'image. Cette texture réduit la variance de la stylisation autour des contours ce qui va à contre courant de FIŠER et al. [36] mais se justifie par l'application de la méthode à la création d'animation pose à pose. Un artiste peint avec le médium de son choix chaque image clé de la vidéo. Les textures guides sont calculées pour chaque image clé. La diversité des tuiles utilisés pour reproduire la stylisation est alors assurée par le nombre d'images clés considérées en exemple. Cependant, utiliser plusieurs images clés introduit une inconsistance au niveau des structures et des couleurs du style car un artiste ne reproduit pas exactement de la même manière ou avec la même logique deux images clés. Chaque image clé guide la génération indépendamment des autres ce qui donne une version stylisée de l'image intermédiaire pour chaque image clé. Les différentes stylisations d'une même image intermédiaire sont donc ensuite mélangées. Le mélange s'effectue pixel à pixel, et seul le pixel généré par l'algorithme de correspondance de tuiles possédant la plus petite erreur est conservé.

Les approches par apprentissage automatique

Des méthodes utilisent les réseaux de neurones pour styliser un flux vidéo. Ces réseaux capturent le style donné dans l'exemple et le reproduisent dans une image source. Lors de leur entraînement, les réseaux de neurones à bases de convolution (CNN) capturent le style de l'exemple en mesurant les caractéristiques de l'image après l'application de différents filtres [38]. Ce type de réseau parvient à capturer les

caractéristiques spatiales et globales du style depuis l'image donnée en exemple. Afin d'améliorer le contrôle utilisateur, ces réseaux peuvent également utiliser des caractéristiques spatiales pré-établies comme des textures guides [39]. D'autres réseaux de neurones capturent le style de l'image donnée en exemple par un champ aléatoire de Markov (*MRF*) [66]. Ce type de réseau apprend à reproduire le style en calculant la couleur de chaque pixel en fonction de leur voisinage proche. De manière similaire aux méthodes basées sur l'optimisation numérique, les réseaux de neurones préservent la cohérence temporelle de la stylisation en incluant dans leur fonction de perte (*loss function*) un terme pénalisant les changements de trajectoire des pixels du flux optique vidéo [83, 84]. Pour une étude plus poussée du domaine du transfert de style par réseau de neurones, je vous propose la revue de JING et al. [55].

Le réseau de neurones est utilisé pour résoudre un problème dont il existe des exemples offrant une solution plausible. Il est alors entraîné à résoudre le problème à partir de ces exemples. LOFTSDOTTIR et al. [71] proposent une méthode permettant de générer des *sprites* d'animations 2D à l'aide d'un unique réseau de neurones. Cette génération se base sur deux images clés et sur le croquis de toutes les images intermédiaires. Le contrôle de la génération par l'utilisateur est introduit grâce à la création de croquis, qui sont utilisés à la fois comme exemples pour former le jeu de données d'entraînement du réseau et comme données d'entrée pour la méthode. Cette méthode tire profit des capacités génératives des réseaux de neurones utilisant une optimisation statistique du problème (*VAE*). Elle est principalement limitée par les caractéristiques du jeu de données utilisé pour entraîner le réseau. Le jeu de données comprends des animations avec un haut taux d'images par seconde, ce qui entraîne le réseau sur des changements subtils de mouvement entre deux images et contient uniquement des animations de marches, de courses et de sauts d'animaux.

LI et al. [68] décomposent la création des images intermédiaires à partir d'un seul croquis représentant le milieu de la séquence. Tout d'abord, le croquis est simplifié pour obtenir des lignes claires. Ensuite un réseau de neurones génératif de type *Transformeur* [106] permet de générer les données qui n'existent pas dans la base d'entraînement en fonction des deux images clés de la séquence et permet d'établir des correspondances entre le milieu de la séquence et chaque image clé en fonction du flux optique estimé. L'analyse du flux optique permet d'établir un masque d'occlusion et de distordre l'image médiane vers les deux images clés (en avant, et en arrière). Ces informations sont utilisées par un nouveau réseau pour déterminer un schéma de mélange afin de générer les autres images intermédiaires composant la séquence, et pour permettre de résoudre à la fois les problèmes d'occlusion et les défauts visuels engendrés par les erreurs d'estimation du flux de mouvement. Enfin, les images intermédiaires nouvellement créées sont modifiées de sorte à améliorer la cohérence temporelle de la séquence par un nouveau réseau spécialement entraîné avec une base de données d'animations 3D. Lorsque l'artiste ne recrée pas exactement les contours présents dans les images clés dans le croquis utilisé en entrée, l'estimation du flux optique fait par le réseau de neurones produit un mouvement erroné et conduit à des incohérences de mouvement au niveau des contours. Cette méthode produit des résultats convaincants

pour la génération d'animations 2D de type cartoon tout en proposant un contrôle par l'exemple. Cependant le réseau de neurones estimant le flux de mouvement ne parvient pas à produire une estimation correcte lorsque les régions mises en correspondances représentent des nuances de détails sans traits clairs apparents ce qui dégrade la stylisation. C'est notamment le cas avec le style peinture.

2.1.3 Stylisation à base de marques

Les méthodes à base de marques pour la génération d'animations 2D utilisent différentes données en entrée. Il y a celles qui partent depuis une animation 3D sous jacente [100], celles qui partent d'images vidéos [80] et celles qui partent d'une ou de plusieurs images clés [32]. Les marques peuvent être distribuées dans l'espace 2D, sur un support, ou directement dans l'espace 3D, sur une surface par exemple. Les méthodes proposées résolvent trois grands problèmes : l'échantillonnage des marques, l'occlusion des marques en fonction du mouvement et leur cohérence dans le temps.

Déduction des traits depuis l'image

Les approches automatiques proposent divers stratégies afin reproduire l'image uniquement avec des marques. La répartition des marques peut être stochastique [44] ou régulière en suivant une grille [47]. Ces deux méthodes de stylisation posent les bases de la déduction des paramètres des marques en fonction des caractéristiques et des attributs sous-jacents d'une scène pouvant être représentée par une image, de la 3D ou une vidéo. Le concept est étendu à la stylisation de vidéos par HERTZMANN et al. [49]. Ils proposent de styliser les images de vidéo en réalisant du *paint over*, une méthode d'animation très utilisée dans la peinture sur verre. Chaque nouvelle image utilise la répartition et les caractéristiques des marques de l'image précédente et seules les parties mouvantes sont générées par dessus. Le flux optique de la vidéo peut également être utilisé pour déformer les marques de l'image précédente afin de les faire suivre le mouvement. HAYS et al. [46] proposent également de réutiliser les marques des images précédentes et développent un nouvel algorithme de mise à jour. De plus ils structurent chaque image en couches de coups de pinceau ce qui leur permet de représenter les détails à différentes échelles. La disposition de marques dans chaque couche est inspirée des travaux de LITWINOWICZ [69] et utilise les caractéristiques sous-jacentes de la scène pour déduire le positionnement et l'orientation de chaque marque. L'idée générale est de détecter une arêtes représentant une haute fréquence dans l'échelle choisie et d'y positionner une marque. Celles créées sur un gradient au dessus d'un seuil fixé sont considérées comme directrices et sont consistantes au fil de la séquence vidéo. Toutes les marques embarquent le mouvement associé au pixel à leur point d'encrage dans l'image et sont advectées uniformément en conséquence. La cohérence temporelle de l'animation est favorisée par l'ajout de contraintes : les marques se chevauchant voient leur opacité réduite en fonction de la couche à laquelle elles appartiennent et de leur statut si elles sont directrices ou non, et enfin les marques gardent une orientation cohérente d'image en image car leur orientation est calculée depuis leur marque directrice à partir de laquelle elles sont créées. Les marques s'éloignant de leur marque directrice voient leur

opacité réduites, et celles avec une opacité nulle sont supprimées. Les nouvelles marques décrivant de nouvelles formes sont graduellement de plus en plus visibles lorsqu'elles sont proches d'une arête. De cette façon le problème d'occlusion est lissé dans le temps. Cependant cette méthode souffre grandement de la précision de l'estimation du flux optique de mouvement de la vidéo qui conditionne finalement combien de temps une marque suit l'arête qu'elle décrit.

Les méthodes automatiques de génération de marques pour la stylisation de vidéos manquent de contrôle artistique et sont réparties en deux catégories : celles purement automatiques et celles manuelles [80]. Bien qu'il soit possible d'ajouter ou de supprimer manuellement des marques sur une image générée, cela reste fastidieux de le faire de manière cohérente sur une séquence entière. O'DONOVAN et al. [80] proposent une méthode de stylisation d'images et de vidéos avec un algorithme de génération de marques contrôlable par un artiste. Tout d'abord, leur génération automatique de marques consiste à placer des points d'ancrage des marques dans chaque image de sorte à respecter une liste de règles de positionnement. Ensuite chaque marque se voit rajouter progressivement de nouveaux points de contrôle de sorte à faire grossir la marque jusqu'à ce qu'une règle de placement soit brisée. Les nouveaux points de contrôle ajoutés à une marque sont déterminés grâce à la minimisation d'une fonction énergie qui contrôle la taille, la couleur, leur orientation et leur cohérence d'image en image. Les marques sont ensuite propagées d'image en image en fonction du mouvement associé au pixel du point d'ancrage. Cette conception permet à l'utilisateur de guider la génération en venant contraindre la minimisation par l'ajout de marque spécifique. Ces marques contraignent le positionnement, la densité et l'orientation des autres marques.

Le principe de déduction des attributs et caractéristiques des marques depuis la scène sous-jacente est également utilisé dans les méthodes où le style est défini par l'exemple. BEN-ZVI et al. [11] proposent une méthode qui permet de faire correspondre une marque créée automatiquement à une marque faite par un artiste. Pour cela l'artiste fournit un ensemble de marques dans un dictionnaire. L'algorithme détecte alors les contours des images de la vidéo à styliser puis génère des courbes, à partir des contours identifiés, qui seront mises en correspondance avec des marques depuis le dictionnaire d'exemples. Ce type d'approche ouvre la voie aux méthodes d'apprentissage automatique apprenant à disposer des marques en fonction des attributs propres à l'image de la même manière qu'un ensemble de stylisation donné en exemple [79]. Pour pouvoir disposer les marques en fonction des attributs de la scène, les approches utilisant les réseaux de neurones se basent sur le rendu différentiel dont le but est de déterminer les paramètres optimaux des marques qui permettent de faire converger le rendu vers l'image souhaitée [122]. Basé sur le rendu différentiel, HUANG et al. [51] proposent un réseau de neurones d'apprentissage par renforcement (*Deep reinforcement learning*) capable de décomposer une image en une centaine de marques. À la différence des approches utilisant un algorithme glouton [44, 47, 69], le réseau peut effectuer des prédictions et est capable de choisir une marque en fonction des cent prochaines.

Pose à pose

Dans les méthodes s'inspirant de la technique *pose à pose*, la disposition des marques est réalisé par un artiste dans les images clés. Pour générer par ordinateur les images intermédiaires il faut résoudre un problème d'interpolation entre les deux images clés tout en conservant le style donné en exemple.

• **Dessin au trait** Le problème d'interpolation des images clés est abordé par des approches effectuant directement une interpolation des marques. Les marques sont mises en correspondance entre deux images clés (Section 2.1.1). Une fois les correspondances de marques établies, les marques sont échantillonnées et les sommets de chaque marque sont ensuite mis en correspondance pour pouvoir être déplacés en fonction d'une trajectoire faisant le lien d'une image clé à une autre [109]. Ce type d'approche se prête particulièrement bien au dessin au trait et à la réalisation de croquis car seul les lignes principales sont représentées, assurant alors un nombre de trait similaire dans les deux images clés. La déformation des marques en fonction du déplacement de ses sommets peut briser des connexions entre marques et modifier la topologie du dessin. C'est pourquoi des approches utilisent les relations entre marques comme contexte à préserver lors de cette opération [119].

Les parties cachées présentes dans les images intermédiaires sont gérées par l'utilisation de surfaces additionnelles déposées dans les images clés par un artiste qui occultent des régions de l'image en fonction de la profondeur attribuée à chaque marque. Ces surfaces peuvent être représentées par des enchaînements de traits décrivant une forme fermée. L'établissement de ces surfaces peut se faire de manière automatique en fermant une courbe ouverte [31], ou par le dessin en faisant intervenir l'utilisateur [119]. Elles peuvent également être représentées par des annotations indiquant alors la direction et la largeur de la zone [54]. Le problème des méthodes précédentes est que pour établir une correspondance marque à marque, les images clés doivent contenir des contours définies par une seule marque à la fois. Pour permettre un nombre de marques différent dans les images clés, EVEN et al. [32] réalisent les mises en correspondances entre structures de marques présents des deux images clés. Ces structures sont mises en correspondance par l'animateur et une transformation par la méthode ARAP (Section 2.1.1) vient décrire la transition d'une structure à l'autre. Les images intermédiaires sont générées en transformant les marques d'une structure à l'autre en interpolant la transformation. L'interpolation des images intermédiaires est l'ensemble des interpolations de chaque structures mises en correspondance dans les images.

Une autre façon d'aborder le problème est d'utiliser les marques des images clés pour créer un panel de nouvelles marques capturant le style des images clés et composant les images intermédiaires. Les marques peuvent être créées depuis les images clés par transformation puis disposées à l'aide d'un guide ou d'une heuristique, comme une carte de densité [18]. Mes travaux de thèse s'inspirent de cette approche, mais se distinguent par l'extension de l'approche au style de peinture. J'aborde également les défis liés au remplissage et aux nuances de détails à l'intérieur des formes peintes dans les images clés.

Représentation des marques dans l'espace 3D

Au lieu d'être représentées dans l'espace 2D du support, les marques peuvent être représentées en 3D et fixées sur des surfaces. Ainsi, elles suivent le mouvement des objets 3D auxquels elles sont rattachées, ce qui permet aux marques de suivre parfaitement le mouvement décrit par la scène. Étant donné que les marques sont rasterisées à l'écran, leur répartition sur les surfaces 3D influence leur répartition 2D à l'écran après la projection. Cette projection se fait par projection orthogonale ou en perspective. Les marques sont réparties sur les surfaces 3D à l'aide d'approches stochastiques [75] ou à partir de l'espace image, puis sont rattachées à la surface par *ray casting* [104]. Ce type d'approche permet également à la stylisation de rester stable et de présenter peu de changements au fil du temps [75]. La profondeur, obtenue par la rasterisation des marques, permet de rendre à l'écran la marque appartenant à l'objet visible. Cependant, les méthodes représentant les marques dans l'espace 3D rencontrent des changements de stylisation abrupts lorsque les marques disparaissent ou apparaissent à la suite d'un changement de visibilité d'un objet de la scène. Afin de minimiser les artefacts visuels lors de ces transitions, le changement de visibilité est fondu en mélangeant les marques de l'image courante avec celles de l'image précédente [104].

BASSETT et al. [4] combinent la représentation des marques dans l'espace 3D avec le principe des images clés peintes par un artiste. La position et la forme des marques de deux images clés est interpolée en fonction de divers événements de scène comme l'animation d'une surface se déformant ou des changements d'éclairage. Cette méthode offre la possibilité de contrôler la stylisation 2D en fonction des techniques d'animations 3D mais rajoute tout un processus de création d'animations 3D qui prend du temps à réaliser. Dans la poursuite de ces travaux mais en rupture avec le principe des images clés, TODO et al. [100] proposent une méthode pour réduire le temps passé à peindre en automatisant la génération des marques à partir d'exemples. Elle consiste à dissocier les éléments spécifiques au style liés à la géométrie 3D, à l'éclairage et à la vue, des exemples fournis par l'utilisateur, puis les transforme en attributs de marque variant dans l'espace image, créant ainsi un champ multi-dimensionnel pour chaque type d'attribut. Ces champs sont ensuite lissés par optimisation afin de minimiser les changements de la valeur des attributs d'image en image et assurer une cohérence temporelle de la stylisation. Une fois les positions des marques déterminées, elles sont synthétisées avec des attributs extraits du champ de marque généré.

Les méthodes optant pour une attache des marques à la surface d'objets 3D obtiennent des stylisations 2D. Cependant les stylisations produites ne permettent pas d'obtenir les niveaux de cohérence temporelle des techniques traditionnelles à cause du mouvement 3D des marques dans la scène.

Ordonnement des marques

L'ordonnement des marques est peu abordé dans les méthodes de stylisation à base de marques. L'ordonnement pose des problèmes à différents niveaux. Par exemple, lorsque que deux marques se chevauchent sur le support il faut déterminer l'apparence finale du support. Cela se produit si les marques sont

déplacées en fonction d'un mouvement, ou peut survenir si elles sont déposées par une fonction heuristique avec une distribution aléatoire. La gestion de l'opacité de chaque marque permet de traiter ce problème [41, 104]. Lorsque les marques sont attachées à des surfaces 3D, elles peuvent se chevaucher une fois projetées à l'écran. Leur rendu nécessite un traitement qui peut se faire en analysant la profondeur du point d'attache à la surface à laquelle appartient la marque [89]. On retrouve également la notion d'ordonnement pour traiter différentes échelles de détails. Des couches regroupant alors des marques de différentes tailles sont ordonnées de sorte à faire apparaître les détails les plus grossiers aux plus fins [46]. Les problèmes d'ordonnement dépendent du médium utilisé. Les médiums ayant une interaction plus prononcée entre chaque marque, comme la peinture, posent davantage de défis pour les méthodes qui les utilisent par rapport à celles utilisant un médium tel que le dessin au trait, car la surface occultée est moindre. L'ordonnement est également un enjeu lorsqu'il s'agit de produire une stylisation naturelle respectant les codes d'un style traditionnel, comme la peinture traditionnelle chinoise, qui suit des règles spécifiques concernant la manière de peindre les formes [118].

2.2 Positionnement scientifique

2.2.1 Les enjeux de la génération d'images intermédiaires

La génération d'images intermédiaires prend en entrée deux images (ou poses) clés, comme dans l'approche traditionnelle *pose à pose*.

La génération d'images intermédiaires est un problème d'interpolation qui doit reproduire à la fois le mouvement de l'animation et le style artistique des images clés. Le mouvement peut être calculé à partir des images clés ou donné en entrée. Les marques des images intermédiaires peuvent être créées de toute pièce, créées à partir de celles dans les images clés ou simplement réutilisées depuis les images clés. La cohérence temporelle est directement liée à la génération d'images intermédiaires et guide les choix techniques. Par exemple si les marques sont représentées en 3D alors le niveau de cohérence de mouvement des marques sera naturellement haut car calqué sur le mouvement des objets auxquels les marques sont attachées. Les animations peuvent présenter différents niveaux de cohérence temporelle. Les animations 3D ont une forte cohérence de mouvement. Les animations de type "cartoon" ont une forte continuité temporelle à l'inverse des animations traditionnelles. Dans ces dernières, chaque image est faite à la main et présente des imperfections uniques qui sont à l'origine du scintillement de la stylisation au cours de l'animation.

L'approche la plus courante pour l'interpolation d'images clés consiste à utiliser la mise en correspondance des marques entre les images clés. Cette approche est particulièrement adaptée lorsque l'animation est principalement basée sur les contours. Cependant, pour les animations de type peinture, les images clés peuvent présenter des nombres de coups de pinceau différents, et il est nécessaire de reproduire les détails et les nuances à l'intérieur des formes. Cela implique l'utilisation

d'une fonction de remplissage complexe qui ne peut se baser entièrement sur la correspondance des contours.

Afin de reproduire le même niveau de contrôle sur la stylisation que la technique du pose à pose, la génération d'images intermédiaires nécessite un processus itératif de raffinement pour la création de l'animation souhaitée. Pour y parvenir, je propose une méthode où n'importe quelle image intermédiaire générée peut devenir une nouvelle image clé, ce qui permet au processus de génération de pouvoir recommencer jusqu'à produire l'intégralité des images de l'animation. Dans l'esprit de la technique pose à pose, avoir une méthode itérative permet à l'artiste d'améliorer le résultat de la génération et lui permet ainsi de se rapprocher progressivement de l'animation souhaitée.

Certaines approches de synthèse de texture permettent de définir un style en utilisant les outils des artistes mais ne permettent pas de modifier le résultat à l'aide de ces mêmes outils. C'est pourquoi mes travaux se sont naturellement orientés vers le rendu à base de marques, qui offre ce contrôle recherché.

Les approches similaires

Cette section présente les méthodes les plus proches scientifiquement de mes travaux de thèse. Elles ont déjà été abordées dans l'état de l'art, c'est pourquoi nous nous concentrons ici sur comment ces méthodes répondent au problème de la génération d'images intermédiaires.

La première méthode similaire à la nôtre est celle proposée par O'DONOVAN et al. [80]. Cette méthode semi-automatique permet de générer des images stylisées à partir d'un flux vidéo et des marques de contrôle dessinées par un artiste. Les marques de contrôle sont dessinées par l'artiste sur les images clés et guident l'orientation des marques générées sur l'image clé. Ces marques de contrôle sont ensuite propagées grâce au flux optique de la vidéo et guident la génération automatique des images suivantes dans la séquence. Cette méthode ne permet pas de définir un style artistique par l'exemple et de le reproduire sur les images générées. Le style artistique de cette méthode est défini par un ensemble de paramètres qui guident la génération des marques, ce qui lui permet d'obtenir un rendu de type peinture (Figure 2.1). Dans la méthode que nous proposons, le style artistique est défini implicitement par toutes les marques des images clés créées par l'artiste. Cet ensemble de marques définissant le style est réutilisé pour la génération des images intermédiaires, ce qui nous permet de reproduire fidèlement le style de l'artiste sans le formaliser explicitement. Notre méthode peut également utiliser un flux optique vidéo comme mouvement d'entrée, mais pas seulement. Nous démontrons que cela est possible avec des mouvements obtenus par la 3D ou le dessin.



FIGURE 2.1 – Stylisation semi-automatique d’images basée sur un algorithme génératif de marques, contraint par les marques ajoutées manuellement par l’utilisateur. O’DONOVAN *et al.* [80].

La méthode de CHEN *et al.* [18] s’inscrit totalement dans l’approche *pose à pose*. De manière similaire à leur méthode, celle que je propose calcule l’interpolation des images clés en remplissant le support de marques de sorte à respecter une carte de densité définissant la quantité de peinture idéale pour chaque image intermédiaire (Figure 2.2). Les deux méthodes procèdent à la propagation des marques dans les images clés en fonction d’un mouvement, afin de générer des marques candidates. Ces marques candidates sont évaluées par la suite et certaines rempliront le support. La première différence entre leur approche et la nôtre réside dans la correction que nous apportons à la forme des marques lorsqu’elles sont transformées par le mouvement. Cette correction vise à préserver leur sémantique et à conserver leurs propriétés géométriques. De leur côté, ils choisissent d’introduire une perturbation aléatoire pour reproduire l’imprécision des traits dans les dessins. La seconde différence concerne le mouvement et la propagation des marques puisque nous utilisons un champ vectoriel 2D tandis qu’eux font le choix d’enregistrer la déformation par la méthode *ARAP* [98]. Cette différence a des implications en termes de souplesse, notamment concernant l’espacement des images clés et les types de mouvement possibles à réaliser.

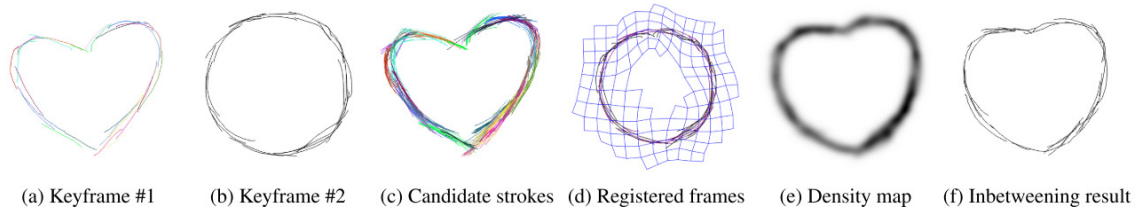


FIGURE 2.2 – Étapes de l'interpolation d'images clés. *CHEN et al. [18]*.

En combinant toutes les problématiques évoquées précédemment, aucune méthode ne permet d'assister un artiste dans la génération d'animation traditionnelle 2D de type peinture, de la réalisation d'images clés donnant le style en exemple à l'édition avec ces mêmes outils d'images intermédiaires générées. Cela nous permet d'explorer de nouvelles problématiques concernant l'ordonnancement des coups de pinceau, le remplissage du support en fonction d'une quantité de peinture cible ou encore l'utilisation de champs vectoriels pour le mouvement.

Chapitre 3

Génération d'images intermédiaires

Publication associée :

- Nicolas Barroso, Amélie Fondevilla, David Vanderhaeghe. "Automatic Intermediate Frames for Stroke-based Animation". *Journées Françaises d'Informatique Graphique (JFIG 2021)*, Nov 2021, Sophia Antipolis, France. HAL-Id : hal-03454288
- Nicolas Barroso, Amélie Fondevilla, David Vanderhaeghe. "Stroke based Painterly Inbetweening". *43rd Annual Conference of the European Association for Computer Graphics (Eurographics 2022 Posters)*, Apr 2022, Reims, France. HAL-Id : hal-03631518

3.1 Fondements pratiques

Cette section pose les concepts nécessaires pour comprendre les méthodes à base de marques depuis l'acquisition et la représentation des marques jusqu'à la création de l'animation finale. De plus, il précise le contexte dans lequel les résultats présentés en section 3.5.4 ont été créés.

3.1.1 Représentation des coups de pinceau

Dans la littérature scientifique, lorsqu'une marque est représentée par une courbe, on la désigne par le terme anglais *stroke*. Cependant, dans la suite de ce manuscrit, dans un souci de simplicité et afin de souligner la généralisation de mes travaux à d'autres styles artistiques, nous gardons le terme de *marque* pour exprimer le coup de pinceau.

Acquisition

L'acquisition des marques dessinées par l'artiste s'effectue par l'utilisation d'une souris ou d'une tablette graphique. Lors du geste de l'artiste, la pression, l'orientation et la pression sont enregistrées le long de la marque.

Paramètres

La structure de données d'une marque contient une courbe, représentée dans sa forme discrète par une polyligne qui décrit le chemin du geste fait l'artiste. Chaque

point de la polyligne est un échantillon comprenant la position 2D, la pression et l'orientation du pinceau virtuel. Dans ce manuscrit, la marque est exprimée comme suit : soit la marque S représentée par une courbe définie par la fonction $f(x)$ avec $x \in [0, 1]$ comme paramètre de la courbe, le j^{eme} point de la polyligne est $f(x_j)$ avec x_j comme paramètre de courbe au point j .

La structure enregistre également l'état du pinceau au moment où la marque est peinte. Cet état prend trois valeurs. Les deux premières valeurs sont *rempli* et *nettoyé* et indiquent que la quantité de peinture sur le pinceau a été mise à jour avant de peindre la marque. Ces deux premiers états précisent le type de marque déposées, respectivement celles qui ajoutent de la peinture sur le support et celles qui l'étalent. Le dernier état *inchangé* indique que le pinceau n'a pas été modifié avant de peindre la marque. La marque dépose ou étale la peinture sur le support en fonction des marques précédentes dans l'ordre de dessin.

$$S \left\{ \begin{array}{l} f(x) \left\{ \begin{array}{l} position \\ pression \\ orientation \end{array} \right. \\ \text{état du pinceau} \end{array} \right.$$

Le chemin emprunté par la marque est également représenté sous sa forme continue par une courbe paramétrique obtenu à partir d'un *fit de Bézier* de la polyligne. La représentation paramétrique permet d'échantillonner la marque en tout point de sa courbe. Cette représentation est avantageuse car elle permet de se libérer de la contrainte d'un nombre de points fixés à l'avance pour représenter la courbe.

Calcul de la représentation paramétrique

La représentation paramétrique de la courbe de la marque est obtenue à partir de l'approximation de sa polyligne par des morceaux de Bézier de degrés trois. L'approximation de la polyligne est un processus itératif qui se déroule jusqu'à obtenir une erreur satisfaisant un seuil fixé [30]. Cette erreur correspond à la distance maximale sur le support entre deux points partageant la même valeur du paramètre de courbe, l'un sur la polyligne et l'autre sur la courbe paramétrique. L'erreur est minimisée en résolvant un système d'équations linéaires par la méthode des moindres carrés où les points de contrôle des morceaux de Bézier sont les variables du système. En pratique nous utilisons un seuil d'erreur fixé à $0.1mm$ pour un support de l'ordre d'un millier de centimètres carrés.

3.1.2 Simulation de peinture

J'ai implémenté la génération d'images intermédiaires dans un moteur de simulation de peinture interne à l'équipe *STORM* du laboratoire *IRIT*. Ce moteur utilise le moteur graphique *Radium* [78] et une implémentation du modèle CHU et al. [21] pour la simulation de peinture. Ce modèle permet de représenter les échanges bidirectionnels de matière entre le pinceau virtuel et le support. Cette simulation pose des contraintes évidentes sur l'ordonnancement des marques et

représente un enjeu pour la méthode proposée. En effet, la plupart des méthodes de génération de marques ne prennent pas en compte les contraintes amenées par un modèle de simulation physique. En intégrant ces contraintes de manière à respecter les techniques de peinture manuelles, ces travaux proposent une méthode compatible avec les modèles physiques de simulation de peinture. Le simulateur de peinture simule également l'éclairage du support et tient compte du type de support utilisé (toile, papier, etc...). Le grain du support et les conditions d'éclairage font partie d'un ensemble de paramètres de rendu qui modifient l'aspect de la peinture et des marques.

3.1.3 Composition d'une animation

Dans l'esprit des méthodes traditionnelles, les animations sont structurées en calques de peinture. Chaque calque est indépendant et la peinture contenu dans le calque ne se mélange pas avec les autres calques. Ces calques se superposent selon un ordre choisi par un artiste et sont rendus pour obtenir l'image finale. Chaque calque possède ses propres paramètres de simulation de peinture et ses propres images clés. Les images clés peuvent être situées à des endroits différents entre chaque calque.

Pour un calque donné, les images d'une séquence d'animation sont structurées en une liste de marques. Cette liste de marques contient toutes les caractéristiques des marques, la configuration du pinceau et les paramètres de rendu. Cela inclut, pour chaque marque, les paramètres du pinceau virtuel tels que la quantité de matière ou la taille, ainsi que la liste des échantillons de la marque comprenant chacun la position, la pression et l'orientation du pinceau virtuel. La même structure de données est utilisée pour les images générées automatiquement par ordinateur. C'est ce qui permet à l'artiste d'éditer les images générées à l'aide des mêmes outils que ceux utilisés pour la création des images clés.

3.1.4 Représentation du mouvement

Le mouvement de l'animation est représenté par une série de champs vectoriels 2D, où chaque champ vectoriel correspond à une image de l'animation. Cela implique que le nombre d'images dans la séquence est déterminé par la discrétisation du mouvement. Un champ vectoriel 2D décrit le mouvement d'une image à l'autre. Ce mouvement peut se faire dans le sens de l'animation ou dans le sens inverse (voir Figure 3.1). Le champ vectoriel possède une largeur et une hauteur, et il est similaire à une image (Figure 3.1). Chaque pixel du champ vectoriel est composé de deux vecteurs 2D. Le premier, \vec{u} , décrit le déplacement dans l'espace image du pixel p de l'image au temps t_i à l'image au temps t_{i-1} . Le deuxième, \vec{v} , décrit le déplacement dans l'espace image du pixel p de l'image au temps t_i à l'image au temps t_{i+1} .

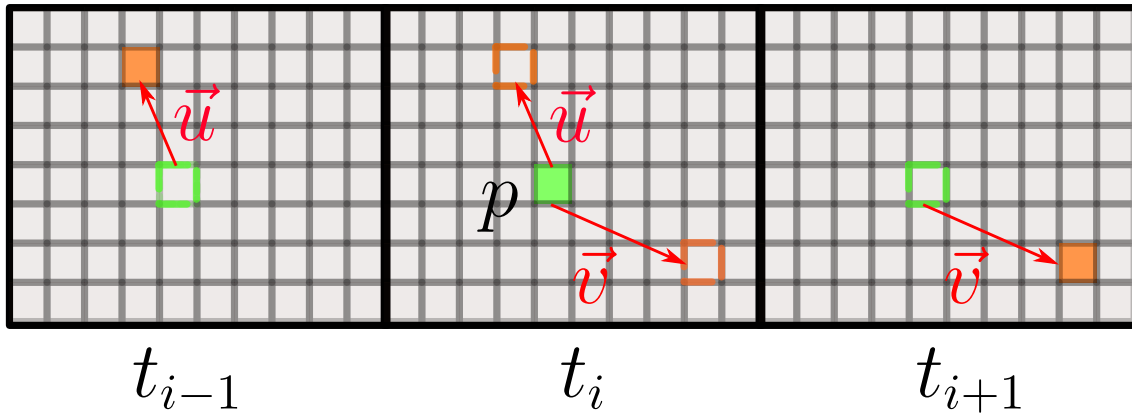


FIGURE 3.1 – Encodage du mouvement en avant représenté par le vecteur \vec{v} et le mouvement en arrière \vec{u} d'un pixel du champs vectoriel.

Résolution et précision

Les champs vectoriels ont une résolution prédéfinie et peuvent être interpolés pour décrire le mouvement de n'importe quel plan 2D, même à une résolution différente. À titre d'exemple, les animations générées dans la partie 3.5.4 utilisent des champs vectoriels de $1000px$ par $600px$ pour décrire le mouvement d'un support de $42cm$ par $25.2cm$, ce qui équivaut à une définition de $0.42mm$ par pixel. La résolution et le ratio d'aspect du mouvement est indépendant de ceux du support. L'artiste peut modifier la mise en correspondance entre le support et le mouvement. L'échantillonnage du mouvement se fait par interpolation bi-linéaire et mise à l'échelle pour correspondre à la résolution du canevas.

3.2 Vue d'ensemble

La méthode présentée dans ce chapitre permet de générer des images intermédiaires à partir d'images clés données en exemple et une description du mouvement (Figure 3.2). En pratique, cette méthode permet à l'artiste de dessiner une image toutes les cinq à dix images de l'animation, tandis que les images intermédiaires générées automatiquement fournissent des intermédiaires plausibles. Le style est capturé à partir des images clés. Ce type de contrôle correspond à ce que les artistes professionnels maîtrisent le mieux : dessiner efficacement pour transmettre leur propre style. La méthode proposée est compatible avec les styles composés de marques pouvant être rendus par ordinateur. Nous illustrons son fonctionnement avec une mise en application avec le style peinture. La méthode proposée ne nécessite pas de compétences supplémentaires de la part des artistes professionnels et s'intègre dans un pipeline de production.

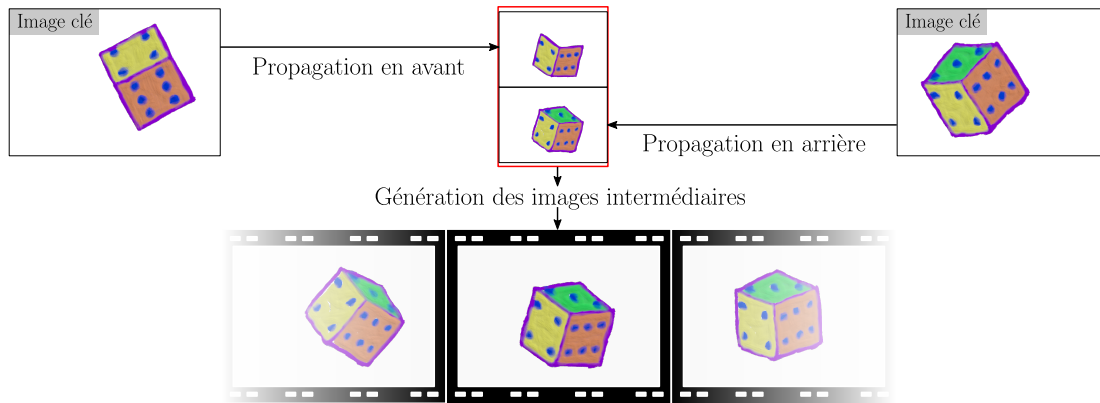


FIGURE 3.2 – Notre méthode génère des images intermédiaires en utilisant la propagation des marques provenant des images clés. Un champ de mouvement fourni en entrée guide le processus de propagation. La description de haut niveau des marques générées permet à l’artiste de modifier les images intermédiaires.

3.2.1 Principe de la méthode

Pour guider la génération d’une image intermédiaire, nous avons besoin d’un champ de mouvement fourni par l’utilisateur. Nous présentons trois exemples de sources de champ de mouvement (3.3.1) : le mouvement rendu à partir d’une scène en 3D qui représente le mouvement prévu de l’animation, le champ de mouvement extrait d’une vidéo et le champ de mouvement esquissé défini par un ensemble de courbes dessinées par l’artiste. Nous montrons ensuite les limites du mouvement avec le style peinture et proposons une solution pour traiter le mouvement avant de l’utiliser (3.3.2). Nous utilisons le mouvement pour :

- propager les marques des images clés afin de créer des marques temporellement cohérentes au moment des images intermédiaires (3.4.1). Lors de la propagation, les marques des images sont soumises à diverses déformations que nous traitons. Ces marques propagées depuis les deux images clés forment deux ensembles ordonnés de marques candidates pour chaque image intermédiaire qui sont fusionnés (3.4.2).
- advecter les quantités de matière des images clés pour estimer la quantité de matière attendues dans les intermédiaires (3.5.1).

Ces ensembles fusionnés de marques candidates servent alors à peindre automatiquement les images intermédiaires en fonction de la quantité de matière attendue (3.4.2). Pour cela chaque marque est évaluée afin de déterminer si sa contribution est positive (3.5.3) pour remplir l’image intermédiaire. A la fin du processus, toutes les images intermédiaires sont générées à partir de la propagation et la modification des marques des images clés. Chaque image intermédiaire est générée dans un style artistique similaire à celui donné dans les images clés.

3.2.2 Flux de travail

La méthode que nous proposons se positionne comme une approche hybride entre les deux méthodes manuelles de création d’animations traditionnelles. Pour créer une animation avec notre méthode, l’animateur a plusieurs options. Tout

d'abord, il peut peindre directement les deux images clés puis utiliser notre méthode pour générer les intermédiaires. Parfois il peut s'avérer difficile de déterminer à l'avance la position de la deuxième image clé et de peindre les éléments de scène de manière cohérente, c'est à dire en tenant compte du mouvement utilisé. L'animateur peut alors utiliser notre méthode pour propager en avant les marques de la première image clé en fonction du mouvement et se servir des images générées comme guide pour peindre la deuxième image clé (Figure 3.3). Ensuite il régénère les intermédiaires avec notre méthode qui tient compte des deux images clés. Ce type d'approche hybride permet à l'artiste de travailler avec la technique *pose à pose* ou *sous la caméra* et offre de nouvelles possibilités d'édition et de création.

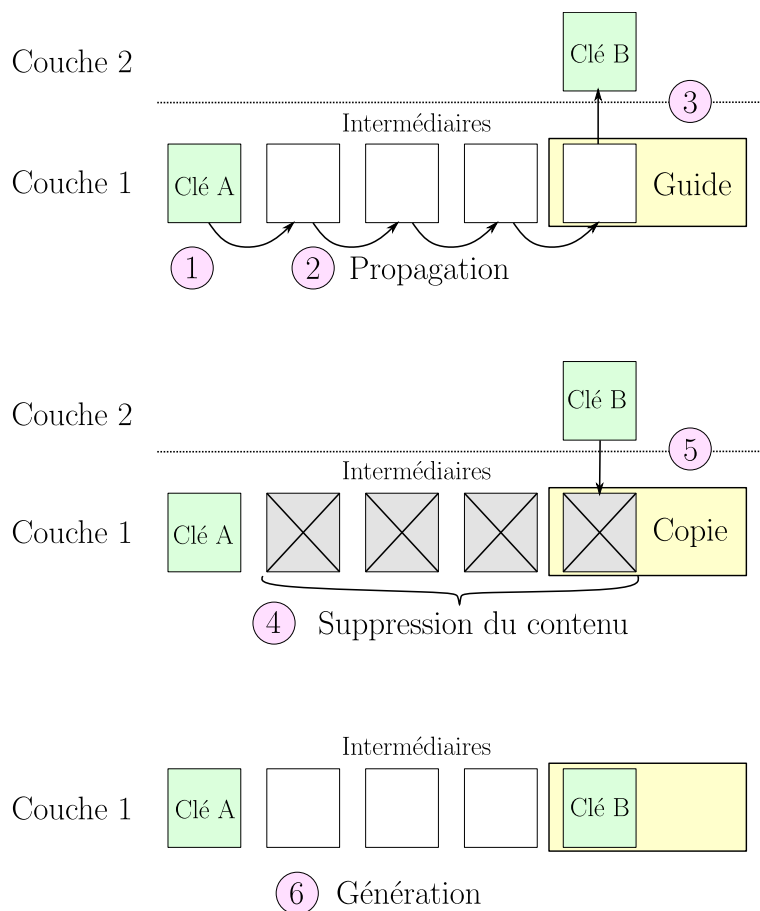


FIGURE 3.3 – Décomposition du flux de travail. Étape 1 : dessin de la première image clé *A*. Étape 2 : propagation en avant du contenu de la première image clé *A* jusqu'à atteindre la position de la seconde image clé *B*. Le contenu obtenu au temps de l'image clé *B* sert de guide pour la peindre. Étape 3 : dessin de la seconde image clé *B* à partir du guide, dans un autre calque. Étape 4 : suppression des contenus générés par la propagation. Étape 5 : copie de l'image clé *B* dans le calque principal. Étape 6 : génération des intermédiaires à partir des deux images clés *A* et *B*.

Les méthodes manuelles de créations d'animations 2D utilisent des calques superposables pour traiter les mouvements complexes comme les objets qui se chevauchent. Notre méthode permet à l'artiste d'animer indépendamment chaque calque, en utilisant ses propres images clés et champs de mouvement.

3.3 Mouvement de l'animation

Les champs de mouvement fournis décrivent comment le contenu de l'image clé doit se déplacer d'une image à l'autre. Chaque champ de mouvement est défini dans l'espace de l'image indépendamment des marques des images clés. Les champs de mouvement déterminent le *spacing* de l'animation (Section 1.3.1). Dans le cadre de notre méthode, ils sont un outil de contrôle indirect sur la stylisation. Nous identifions trois propriétés que les champs de mouvement en entrée doivent respecter afin d'être pleinement exploitables par notre méthode.

1. Les champs de vecteurs doivent décrire le mouvement des objets représentés dans la scène.
2. Chaque champ de vecteurs doit représenter un seul mouvement correspondant à un objet ou un groupe d'objets partageant un mouvement similaire.
3. Le mouvement représenté par le champ de vecteurs ne doit pas se limiter à la silhouette de l'objet en mouvement.

3.3.1 Différentes sources de mouvement

Nous proposons trois sources pour la création des champs de mouvement :

Champs vectoriels dessinés

Nous proposons une méthode pour obtenir des champs de mouvement en utilisant des courbes paramétriques dessinées par l'artiste pour représenter le mouvement entre les images (Figure 3.4). Chaque courbe paramétrique définit le mouvement d'un pixel au cours du temps, du début à la fin de la courbe. Des marqueurs temps sont ensuite disposés sur la courbe et définissent la position du pixel au temps des images de la séquence. Le mouvement d'un pixel correspond alors à une suite de translation, passant d'une image à l'autre. L'utilisation de multiples courbes paramétriques permet de définir un champ vectoriel épars où seulement quelques pixels encore une information de mouvement. Mais pour obtenir un champ vectoriel utilisable par notre méthode, le mouvement des pixels de l'image doit être extrapolé. L'extrapolation permet de diffuser les valeurs de mouvement connues dans l'image ce qui génère un champ vectoriel dense où chaque pixel encode un mouvement (3.3.2).

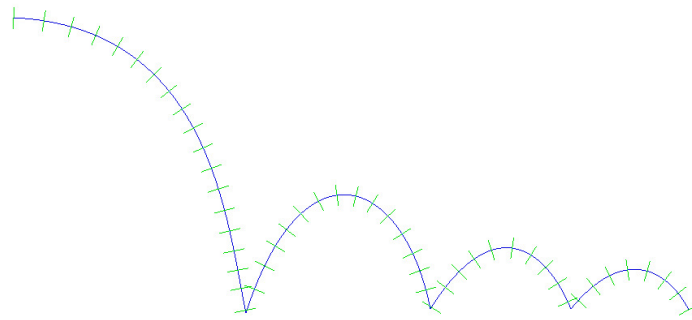


FIGURE 3.4 – Mouvement d'un pixel dessiné au cours des cinquante images de la scène *Balle rebondissante* 3.5.4, de gauche à droite. Les marqueurs temps correspondent aux traits de couleur verte et indiquent la position du pixel au temps de l'image.

Champs vectoriels rendus depuis la 3D

Le rendu des scènes 3D animées permet d'exporter le mouvement des objets composant la scène. Ce mouvement est obtenu en calculant la différence de position à l'écran des objets entre l'image précédente et l'image courante. Pour ce faire, on utilise les matrices de transformation modèle, vue et projection (*MVP*), de chaque objet pour l'image précédente et l'image courante. Dans le cadre de nos travaux, nous désactivons l'anti-crénelage pour éviter un lissage des mouvements aux extrémités des objets. Ce lissage a un effet négatif sur la qualité du mouvement et introduit des valeurs de mouvement erronées. La configuration de la scène 3D offre la possibilité de rendre uniquement certains éléments plutôt que d'autres, facilitant ainsi la création de champs de mouvement personnalisés par calque. Dans nos travaux, nous avons exploité les fonctionnalités de *Blender* [24] pour tirer parti de cet avantage et obtenir nos champs de mouvement à partir des scènes 3D (Figure 3.5). Cependant, les champs de mouvement provenant de scènes 3D définissent uniquement le mouvement dans la zone rendue de l'objet animé. Cela pose un problème en style peinture lorsqu'on utilise le champ de mouvement pour déplacer les marques, car celles qui s'étendent au-delà de la zone de l'objet peuvent ne pas être correctement déplacées. Par conséquent, le champ de mouvement rompt avec la troisième propriété et nécessite une extrapolation afin d'avoir un mouvement dans l'intégralité de l'espace image (Section 3.3.2).



FIGURE 3.5 – Champ vectoriel utilisé pour la scène du *Ballon de plage* (Section 3.5.4), obtenu avec le logiciel *Blender*.

Champs vectoriels extraits depuis une vidéo

L'extraction d'un flux optique dense à partir d'une vidéo définit le mouvement de chaque pixel dans l'espace image. Ce champ de mouvement peut être utilisé tel quel pour notre méthode. Il fournit des informations sur le mouvement de toute la scène, mais ne permet pas d'isoler le mouvement d'éléments distincts dans la scène, ce qui ne permet pas de respecter la deuxième propriété. L'utilisation du flux optique vidéo rend délicate la séparation du mouvement en calque. L'extraction du mouvement d'un objet de la scène est possible en utilisant des méthodes de segmentation de flux optique [120]. Cependant, nous n'utilisons pas de segmentation du flux optique dans nos résultats car nous montrons les limites de notre méthode avec un flux optique dense. La troisième propriété ne peut également pas être respectée car le mouvement extrait représente le mouvement des formes présentes dans la vidéo et ce mouvement s'arrête à leur contour. Nous avons utilisé l'algorithme de FARNEBÄCK [33] implémenté dans *OpenCV* pour extraire le flux de mouvement de vidéos (Figure 3.6).



FIGURE 3.6 – Champ vectoriel estimé depuis le flux vidéo de l’animation *Sintel*.

3.3.2 Extrapolation du mouvement

Les champs de mouvement éparés ou incomplets nécessitent une extrapolation afin de fournir des vecteurs de mouvement à chaque pixel du champ, permettant ainsi de déplacer les marques peintes, même si elles s’étendent dans des régions où l’information de mouvement est absente. Dans ce but, nous effectuons un pré-traitement du mouvement pour le rendre compatible avec notre méthode de génération d’images intermédiaires. Le mouvement produit à partir de sa version incomplète vérifie la troisième propriété sans rompre avec les deux premières.

Pour ce faire nous utilisons une diffusion bi-harmonique de BAXTER et al. [7] que nous adaptons à notre problème. L’idée est d’extrapoler dans l’espace image les valeurs des vecteurs de mouvement connus, encodés dans chaque pixel. La méthode de diffusion est basée sur la construction d’un maillage, dans lequel des sommets représentent les valeurs connues ou inconnues.

Tout d’abord, le processus d’extrapolation construit un maillage 2D triangulaire T qui subdivise le plan image. Le maillage est construit à l’aide de la bibliothèque Python *Triangle* [87] en utilisant l’algorithme proposé par SHEWCHUK [90]. Cette triangulation fournit un ensemble de sommets qui échantillonnent visuellement l’espace de manière régulière (Figure 3.7a) en posant une contrainte de distance entre chaque sommet. L’impact de la répartition des sommets sur la qualité de l’extrapolation est négligeable par rapport à celui concernant le nombre de sommets utilisés. Ensuite, le processus associe les vecteurs de mouvement connus aux sommets les plus proches puis calcule la diffusion bi-harmonique qui permet d’obtenir les vecteurs de mouvement des autres sommets (Figure 3.7b).

La diffusion repose sur une optimisation numérique où la somme des normes des différence entre chaque couple de matrice jacobienne de triangles adjacents du maillage est minimisée. Cela correspond à la minimisation de l’énergie suivante :

$$\sum_{j,k \in \zeta(T)} \|J(j) - J(k)\|^2 + \epsilon \sum_{i \in F} \|J(i)\|^2,$$

où :

$$J(f) = \begin{bmatrix} V(f_a) - V(f_c) \\ V(f_b) - V(f_c) \end{bmatrix}^{-1} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} \omega(f_a) \\ \omega(f_b) \\ \omega(f_c) \end{bmatrix}$$

F représente l'ensemble des faces, V les sommets, ω les vecteurs de mouvement associés aux sommets, c'est-à-dire les variables de la minimisation. Une face f possède trois indices, f_a , f_b et f_c , qui indexent V et ω . $\zeta(T)$ représente les paires de toutes les faces dans T partageant une arête.

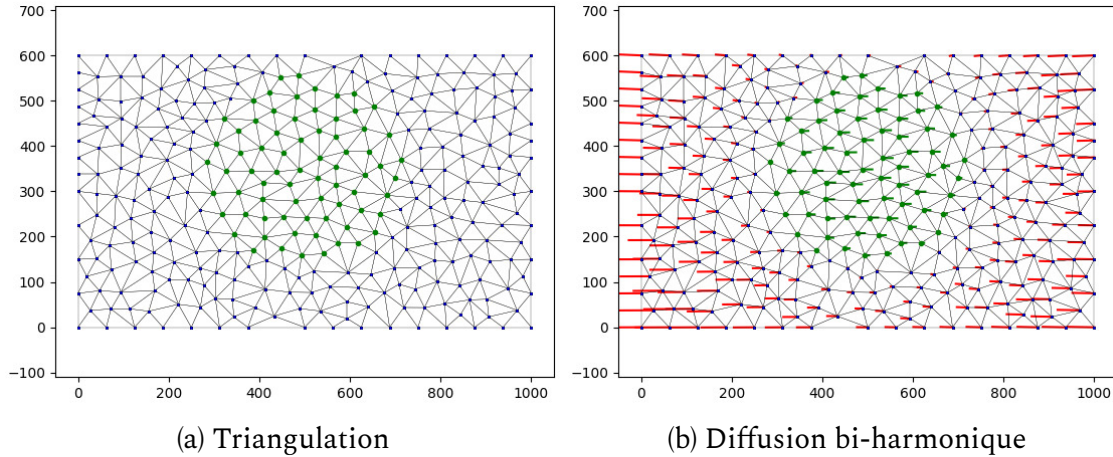


FIGURE 3.7 – Extrapolation par diffusion bi-harmonique du mouvement de la scène du *Ballon de plage* (Section 3.5.4). Les points de couleur verte représentent le mouvement connu associé aux sommets. Les lignes rouges représentent la direction et l'intensité des vecteurs de mouvement extrapolés.

Pour calculer la matrice jacobienne d'un triangle, le processus implique l'inversion d'une matrice. Cependant, si les sommets formant le triangle sont colinéaires, cette matrice devient non inversible. Le deuxième terme de la minimisation garantit la validité du système [7]. Une fois les variables et les contraintes du système établies, celui-ci est résolu à l'aide de la méthode des moindres carrés.

Après avoir diffusé les valeurs de mouvement à tous les sommets du maillage, le processus calcule un nouveau champ vectoriel décrit par un plan image de même dimension et résolution. Les valeurs de mouvement calculées aux sommets du maillage sont interpolées pour chaque pixel du plan image. Pour chaque pixel p , si les trois sommets de T qui encadrent p possèdent un vecteur de mouvement initialement connu (i.e. non nul dans le champ de mouvement d'origine), alors p conserve la valeur présente initialement avant le processus d'extrapolation. Dans le cas contraire, le vecteur de mouvement du pixel p du plan image est calculé en utilisant une interpolation barycentrique à partir des vecteurs de mouvement associés aux sommets de T .

Limites : L'extrapolation par diffusion bi-harmonique d'un mouvement permet de générer des champs vectoriels cohérents et exploitables pour l'advection de marques, particulièrement lorsque le mouvement correspond à une transformation simple telle qu'une translation, une rotation ou un zoom. Cette extrapolation

est effectuée de manière globale sur l'ensemble de l'espace image du champ vectoriel. Cependant lorsque deux mouvements représentant deux objets distincts sont extrapolés, le mouvement créé comporte une ou plusieurs zones dont le mouvement ne décrit plus le mouvement de l'un des objets. Ces zones sont considérées comme des singularités de l'extrapolation (Figure 3.8). Ces zones de mouvements extrapolés représentent une jonction à forte turbulence entre deux mouvements différents qui déforment les marques lors de l'advection. Nous introduisons les poids de confiance comme un outil pour atténuer cet impact.

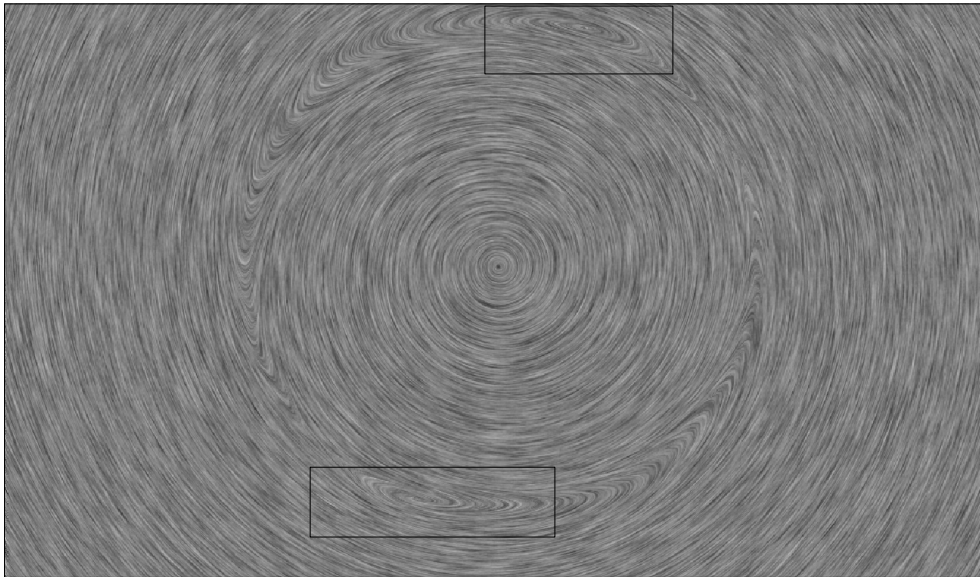


FIGURE 3.8 – Représentation sous forme de flux (*line integral convolution* [17]) d'un champs de mouvement extrapolé. Le champ de mouvement contient la rotation dans le sens horaire d'un disque situé au centre et la rotation externe dans le sens anti-horaire d'un tore. Cette représentation met en évidence les zones dont le flux de mouvement extrapolé ne correspond pas au mouvement initial des objets. La version en couleur de ce champ de mouvement est disponible sur la figure 3.10.

3.3.3 Poids de confiance

Pour chaque vecteur du champ de mouvement, une valeur de confiance lui est associée créant ainsi un masque de confiance dans l'espace image. Dans un premier temps, cette valeur est binaire et prend la valeur un si le pixel dans le champ de mouvement correspond à un mouvement donné en entrée avant le processus d'extrapolation ou prend la valeur zéro sinon. Ensuite, un filtre *Gaussien* est appliqué sur le masque binaire et permet d'adoucir la transition entre les zones d'origines et les zones extrapolées. La puissance du filtre définit la largeur de la zone de transition où la confiance est non nulle entre le mouvement extrapolé et le mouvement d'origine. Cette valeur de confiance est ensuite utilisée en tant que poids de mélange σ pour corriger la forme des marques (Section 3.4.1).

3.4 Génération de coups de pinceau

L'idée principal de la méthode est de générer un ensemble de marques candidates à partir des images clés. Les marques candidates sont ensuite évaluées et sélectionnées pour reproduire le style des images clés.

3.4.1 Propagation des coups de pinceau

Chaque image intermédiaire est encadrée dans la séquence par deux images clés que nous appelons l'image clé précédente et l'image clé suivante. Pour générer les marques candidates d'une image intermédiaire, les marques de l'image clé précédente sont advectées en utilisant le champ de mouvement vers l'avant, tandis que les marques de l'image clé suivante sont advectées avec le champ de mouvement arrière. Prenons en considération l'image clé précédente et le champ de mouvement vers l'avant pour le reste de cette section. Chaque étape d'advection déplace les points de polygone des marques d'une image à l'image suivante et produit des marques candidates pour cette image intermédiaire. Chaque marque de l'image clé est transformée en une marque candidate pour chaque image intermédiaire. Les marques candidates d'une image intermédiaire sont calculées, de manière itérative, à partir des marques candidates de l'image intermédiaire précédente dans la chronologie, ou à partir de l'image clé pour la première image intermédiaire.

De légères divergences dans le champ de mouvement ont un impact perceptible sur la courbe de la marque, entraînant la perte de la forme de courbe originale. Cependant par choix artistique, l'artiste peut souhaiter conserver la forme de courbe originale de la marque tout au long de l'animation, avec une transformation rigide uniquement. Pour contrôler la forme de courbe de la marque advectée, nous proposons un algorithme de régularisation. La régularisation permet à la forme de la courbe associée à la marque de suivre le mouvement tout en préservant la forme originale de la marque dessinée dans l'image clé.

Pour obtenir la courbe advectée et régularisée f_n de l'image intermédiaire t_i à partir de la courbe de la marque f_p provenant de l'image précédente t_{i-1} , nous utilisons deux courbes : la courbe advectée et la courbe de référence. La courbe de référence correspond à la version de courbe qui a été advectée de manière rigide de sorte à préserver la forme originelle de la courbe f_p .

f_p : la courbe avant advection,
 f_r : la courbe de référence, obtenue par transformation rigide,
 f_a : la courbe advectée,
 f_n : la courbe advectée régularisée, la courbe finale.

La régularisation est contrôlée par un poids de mélange permettant de mélanger la forme entre la courbe advectée et la courbe de référence. Le poids de mélange d'une marque est défini comme la fonction $\sigma(x)$, où $x \in [0, 1]$ représente le paramètre de la courbe. La fonction $\sigma(x)$ est calculée en faisant correspondre la valeur de confiance du champ de mouvement sous-jacent utilisé pour l'advection (voir Section 3.3.3) aux points de la polygone de f_p sur le support. Nous permettons

également à l'artiste de modifier le poids de mélange au niveau de la marque pour permettre un contrôle précis et détaillé du processus de régularisation. Lorsque $\sigma(x)$ vaut zéro, cela correspond à l'endroit de la marque où l'on s'attend à une transformation rigide de la courbe. Lorsque $\sigma(x)$ vaut un, cela correspond à l'endroit de la marque où l'on s'attend à une advection modifiant la courbe selon le mouvement.

Calcul de l'advection par le mouvement

La courbe advectée f_a est définie comme suit :

$$f_a(x_j) = f_p(x_j) + M(t_{i-1} \rightarrow t_i, f_p(x_j))$$

où $M(t_{i-1} \rightarrow t_i, p)$, le vecteur de mouvement de t_{i-1} à t_i à la position du pixel p .

Calcul de la déformation rigide

La courbe de référence f_r est calculée en appliquant une transformation 2D rigide Y à f_p :

$$f_r(x_j) = Y(f_p(x_j))$$

La transformation Y est calculée par la méthode des moindres carrés en suivant l'approche de décomposition en valeurs singulières de SORKINE-HORNUNG et al. [95].

$$Y = \arg \min_Y \sum_j w_j \|Y(f_p(x_j)) - f_a(x_j)\|^2$$

où les poids $w_j = \lfloor \sigma(x_j) \rfloor$, de sorte que seuls les points où nous prévoyons une advection soient utilisés pour définir Y , permettant ainsi à f_r de suivre ces points. Si aucun point de ce type n'existe, nous utilisons $w_j = 1$ pour tous les points de manière à ce que f_r suive f_a , tout en utilisant une transformation rigide.

Correction de la forme des coups de pinceau propagés

Nous appliquons une correction de chaque marque en fonction des poids de mélange qui sont déterminés par le mouvement utilisé. Les parties de la courbe f_n utilisant un mouvement à faible confiance doivent reproduire les caractéristiques de la courbe de référence f_r qui préserve la forme de la courbe de l'image précédente f_p . Les autres parties de la courbe f_n utilisant un mouvement à forte confiance lors de l'advection doivent reproduire les caractéristiques de la courbe advectée, potentiellement transformée par le mouvement. Pour reproduire les caractéristiques d'une courbe nous nous basons sur : la longueur de la courbe, plus précisément la longueur entre deux échantillons à pas constant le long de la courbe, la position de la courbe sur le support et la courbure locale le long de la courbe. Nous cherchons une interpolation entre f_a et f_r dont l'interpolant $\sigma(x)$ varie le long de la courbe f_n . Pour y parvenir nous utilisons une optimisation non linéaire :

Étant donné qu'une courbe est représentée par une polyligne pouvant contenir des centaines de points, nous utilisons la forme paramétrique de la courbe d'une marque 3.1.1 afin de réduire le nombre de variables à optimiser. Nous formulons le problème comme une optimisation des points de contrôle d'une spline composée de morceaux de Bézier de degré trois et utilisons une combinaison de trois fonctions d'énergie à minimiser. Ces énergies sont mesurées le long de la courbe pour K points d'échantillonnage répartis uniformément $x_k, k \in [1, K]$.

La première fonction d'énergie correspond au terme de fidélité aux données E_f pour la courbe linéaire mélangée f_b , ce qui garantit qu'en l'absence d'autres contraintes, l'optimisation convergera spatialement vers une interpolation linéaire entre la courbe advectée et la courbe rigide en fonction de $\sigma(x)$.

$$f_b(x) = f_r(x) + \sigma(x)(f_a(x) - f_r(x))$$

$$E_f = \frac{1}{K} \sum_{k=1}^K \|f_b(x_k) - f_n(x_k)\|^2$$

Cependant, la courbe résultante de l'interpolation linéaire n'est pas suffisante. Elle introduit des défauts qui ne sont pas présents dans la courbe de référence (Figure 3.9-c). Nous proposons d'ajouter deux nouvelles contraintes, toujours sous forme de fonctions d'énergie, dans l'optimisation.

Les deux prochaines fonctions d'énergie E_c et E_l pénalisent les variations entre la courbe de référence et la courbe régularisée en minimisant, respectivement, la différence de courbure et la différence de distance entre chaque paire d'échantillons consécutifs.

$$E_c = \frac{1}{K} \sum_{k=1}^K (1 - \sigma(x_k)) (\gamma_r(x_k) - \gamma_n(x_k))^2$$

avec γ_λ la courbure de f_λ , $\lambda \in [r, n]$.

$$E_l = \frac{1}{K-1} \sum_{k=1}^{K-1} \left(1 - \frac{\sigma(x_k) + \sigma(x_{k+1})}{2} \right) (\delta_r - \delta_n)^2$$

avec

$$\delta_\lambda = \|f_\lambda(x_k) - f_\lambda(x_{k+1})\|$$

Nous utilisons l'algorithme de Levenberg-Marquardt [73] pour résoudre l'optimisation et utilisons les points de contrôle de la courbe advectée f_a comme initialisation car ils fournissent déjà une approximation de la courbe régularisée. L'énergie minimisée est une combinaison linéaire des trois énergies présentées :

$$E = E_f + \alpha E_c + \beta E_l$$

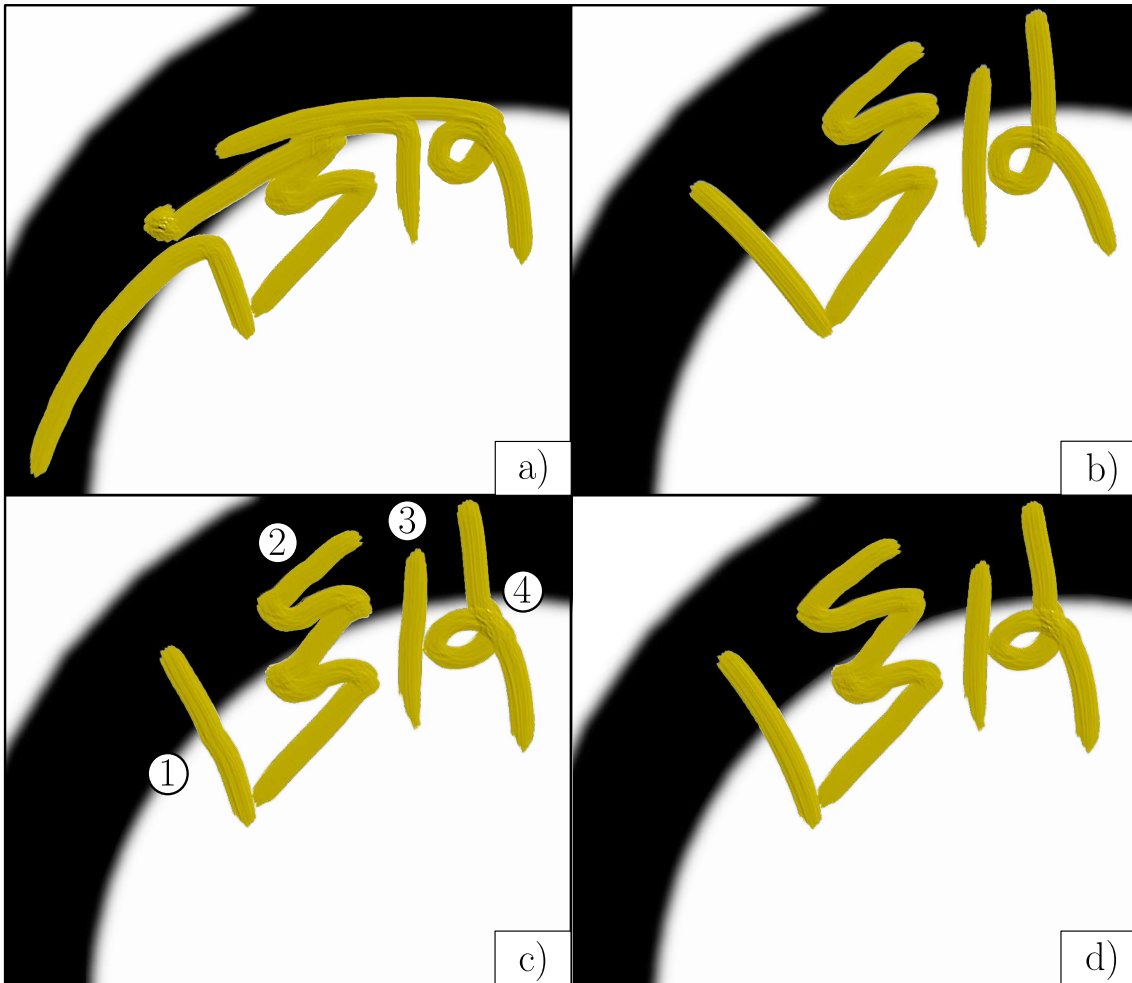


FIGURE 3.9 – Différences entre une interpolation linéaire et notre régularisation de courbes. En a) les courbes advectées f_a non régularisées présentant une forte déformation sur la partie supérieure des courbes, où les poids de confiance tendent vers zéro lorsque l'on s'éloigne de la forme principale blanche. Le résultat attendu ici est un mélange entre a) et les courbes de référence f_r b) obtenues par advection rigide. On souhaite conserver les parties inférieures des courbes de a) et les parties supérieures de b). En c) Les courbes mélangées linéairement f_b , et en d) les courbes régularisées par notre optimisation. Le champ de mouvement utilisé est celui de la figure 3.10. L'intensité du fond correspond aux poids de mélange utilisés $\sigma(x)$ qui varient de zéro à un. Les défauts liés au mélange linéaire sont mis en valeur dans c) avec des cassures au niveau de la courbe (1, 4), un tassement (2) et une elongation (3).

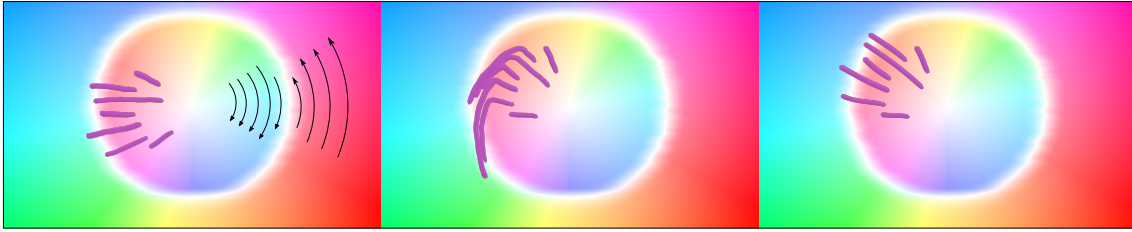


FIGURE 3.10 – Lorsque les marques (en violet) d’une image clé (à gauche) sont advectées par un champ de mouvement, les marques advectées vers l’avant (au milieu) présentent des déformations importantes. Notre régularisation des marques (à droite) capture le mouvement tout en préservant la forme des marques de l’image clé, selon un poids de mélange. Le champ de mouvement sous-jacent (représenté par la couleur de fond et les flèches) illustre une rotation centrale dans le sens horaire et une rotation externe dans le sens anti-horaire. Les poids de mélange (indiqués dans la Figure 3.9) sont nuls dans la région de transition entre les deux rotations.

Nos résultats sont produit avec $\alpha = 25$ et $\beta = 50$. La Figure 3.10 montre comment le processus de régularisation gère les marques fortement déformées par l’application du mouvement.

De la même manière, un second ensemble de marques régularisées sont calculées à partir de l’image clé suivante en utilisant le champ de mouvement vers l’arrière. À la fin de l’étape d’advection et de régularisation des marques, nous avons généré deux listes de marques candidates, provenant chacune d’une image clé, pour chaque image intermédiaire. Ces listes de marques candidates ont le même ordre de dessin que leur image clé d’origine.

3.4.2 Organisation des marques candidates

La méthode de génération des images intermédiaires consiste à tirer parti des marques candidates issues de l’image clé précédente et de l’image clé suivante, car ces deux sources se complètent mutuellement. En effet, si une partie d’un objet est uniquement peinte dans une image clé donnée, cela peut indiquer une apparition ou une occlusion. La représentation de cette partie dépend alors exclusivement des marques issues de cette image clé. Par conséquent, il est essentiel d’utiliser les marques candidates provenant des deux images clés. Ces deux listes rassemblent les marques candidates qui viendront composer l’image intermédiaire par un processus de sélection (Section 3.4.2).

Bien que chaque liste soit ordonnée de manière cohérente, l’ensemble des marques candidates, lui ne l’est pas. L’ordre des marques a un impact significatif sur l’apparence du résultat final lors de la combinaison des deux listes. Pour obtenir un ensemble cohérent des marques candidates, nous fusionnons les deux listes de marques candidates issues de la propagation et définissons un ordre final pour l’ensemble. Lorsqu’aucune attention particulière est consacrée à l’ordre des marques candidates, c’est à dire que les marques sont mélangées en fonction de leur ordre de dessin interne à chaque image clé, le mélange de couleurs qui en résulte produit

des résultats très éloignés des images clés utilisées. La façon dont chaque marque agit sur le support dépend de la quantité et de la couleur de la peinture dans le pinceau ainsi que sur le support avant de dessiner la marque. Une marque peut commencer avec un outil fraîchement rempli d'une couleur, fraîchement nettoyé pour les traces, ou conserve son état et correspond à la fin du dessin des marques précédentes (Section 3.1.1). Dans ce dernier cas, la couleur de l'outil dépend de l'interaction précédente du pinceau avec le support lors de la simulation de la peinture.

Chaque marque d'une image clé possède un indice d'ordre de dessin, qui correspond à son indice dans la liste des marques. Pour faciliter le mélange des deux listes de marques, le processus d'ordonnancement attribue une position de dessin en faisant correspondre les indices des marques sur l'intervalle $[0, 1]$, indépendamment dans chaque liste de marques. Trier les marques candidates par leur position de dessin n'est pas suffisant pour fusionner les deux listes car les marques avec un état d'outil «inchangé» voient leur couleur modifiée (Fig. 3.11-a). Par conséquent, nous définissons des groupes de marques pour créer des lots de marques cohérentes dans chaque liste, puis fusionnons les deux listes en préservant les groupes et l'ordre des groupes.

Nous proposons un mécanisme de groupement automatique basé sur l'état de l'outil des marques. La création des groupes dans chaque liste de marques candidates se fait de la manière suivante. Chaque marque ayant un état d'outil «nettoyé» ou «rempli» démarre un nouveau groupe, tandis que celles avec l'état «inchangé» ne le font pas. Si nécessaire, l'artiste peut ajuster manuellement le groupement automatique et l'indice de groupe pour affiner la manière dont les groupes des images clés sont entrelacés. Ensuite, nous attribuons à chaque groupe un ordre unique en fonction de la position d'ordre de dessin de la première marque du groupe, et ce dans les groupes des deux listes de marques candidates. Cet ordre du groupe permet de fusionner les deux listes de marques candidates et de les ordonner par groupe. L'ordre de dessin des marques dans un groupe est le même que celui dans l'image clé et l'état de l'outil de ces marques est le même que dans l'image clé (Fig. 3.11-b).

Ce processus définit la liste fusionnée des marques candidates ainsi que l'ordre de dessin pour l'image intermédiaire. Il rassemble les marques dépendantes dans un même groupe et permet une fusion groupe à groupe qui préserve l'état du pinceau virtuel de chaque marque au moment où elles ont été peintes par l'artiste.

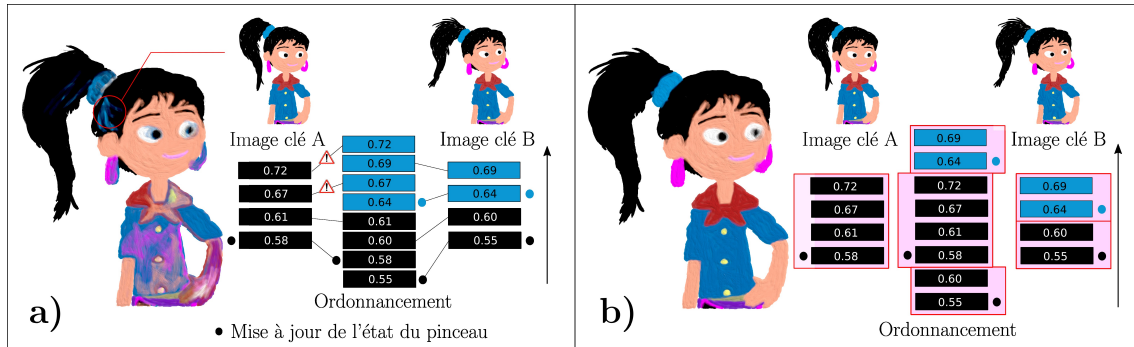


FIGURE 3.11 – Une comparaison entre l’ordre des marques candidates basé sur leur position de dessin individuelles (a) et l’ordre des groupes de marque basé sur leur position de groupe (b). Ordonner les marques en fonction de leur position de dessin peut influencer l’état de l’outil utilisé pour la peinture, entraînant des différences de couleur et de quantité de matière pendant la simulation. Les images clés sont montrées au temps t_{10} et t_{20} , et la trame intermédiaire est montrée au temps t_{11} .

3.5 Sélection des marques candidates

Une fois que chaque image intermédiaire possède son propre ensemble ordonné de marques candidates, il est nécessaire d’en sélectionner une partie pour la composition. La première intuition est d’utiliser toutes les marques candidates disponibles pour composer l’image intermédiaire. Cependant, ce choix ne permet pas de reproduire un style similaire à celui des images clés, car l’image intermédiaire contient deux fois plus de marques qu’une seule image clé. Un excès significatif de matière devient clairement perceptible sur le support (Figure 3.12). Pour pallier ce problème, nous guidons la sélection des marques en utilisant une carte spécifiant la quantité optimale de matière nécessaire pour la création de l’image intermédiaire dans un style proche de celui des images clés.

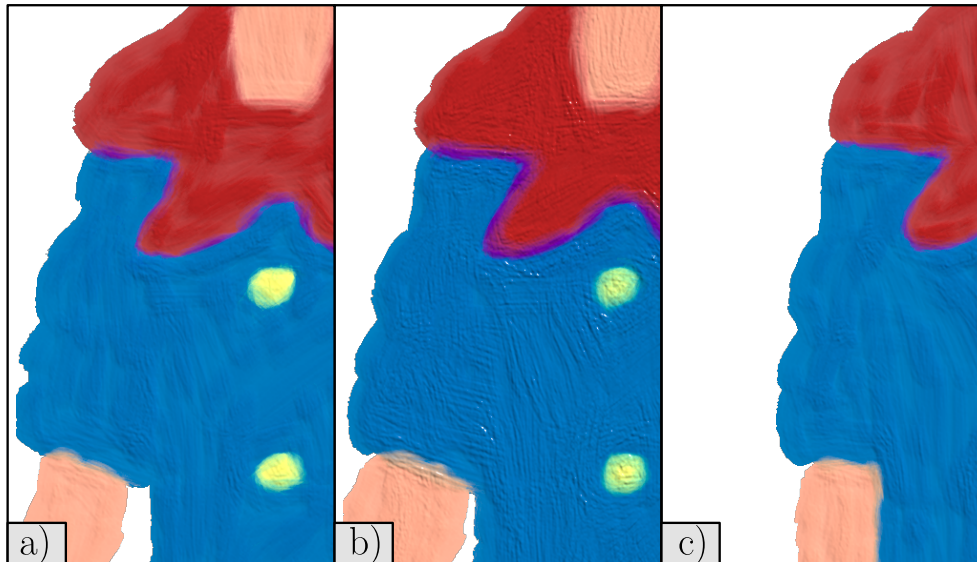


FIGURE 3.12 – La sélection de toutes les marques des deux images clés entraîne un surplus de peinture visible qui ne permet pas de reproduire le style des images clés a) et c), respectivement aux temps t_{10} et t_{20} . L'image intermédiaire b) est au temps t_{11} .

3.5.1 Génération des cartes de quantité de matière

Nous utilisons les cartes de quantité de matière utilisées par les modèles de simulation de peinture pour guider la création des images intermédiaires. L'objectif est d'assurer que, une fois rendues, les images intermédiaires présentent une quantité de matière similaire à celle des images clés, en évitant tout excès ou manque de matière. Chaque carte de quantité de matière est représentée par une image où chaque pixel encode la quantité de matière. Pour accomplir cela, les cartes de quantité de matière des deux images clés sont advectées vers les images intermédiaires à travers une série d'advections successives, en utilisant les champs de mouvement sous-jacents. Ce processus se déroule de manière similaire à l'advection des marques (Section 3.4.1). L'advection se fait par correspondance en cherchant la nouvelle quantité de matière du pixel p dans l'image intermédiaire t_i depuis son correspondant dans l'image précédente t_{i-1} en utilisant le mouvement $M(t_i \rightarrow t_{i-1}, p)$, si l'advection se fait vers l'avant. De cette façon nous évitons d'avoir plusieurs correspondants pour un même pixel si nous cherchions le correspondant des pixels de t_{i-1} dans t_i avec $M(t_i \rightarrow t_{i+1}, p)$.

$$\begin{aligned} A_i^0(p) &= A_{i-1}^0(p + M(t_i \rightarrow t_{i-1}, p)) \\ A_i^N(p) &= A_{i+1}^N(p + M(t_i \rightarrow t_{i+1}, p)) \end{aligned}$$

avec A_i^0 et A_i^N qui sont respectivement la quantité de matière advectée depuis l'image clé précédente t_0 et l'image clé suivante t_N , jusqu'à l'image intermédiaire t_i . Nous définissons la carte de quantité de matière cible T_i comme le maximum pixel par pixel de la quantité de matière des deux images clés. Le maximum agit comme un opérateur d'union qui combine les informations apportées par les deux images clés (Figure 3.13).

$$T_i = \max(A_i^0, A_i^N)$$

À la fin de la propagation de la quantité de matière, nous obtenons une carte de quantité de matière cible pour chaque image intermédiaire.

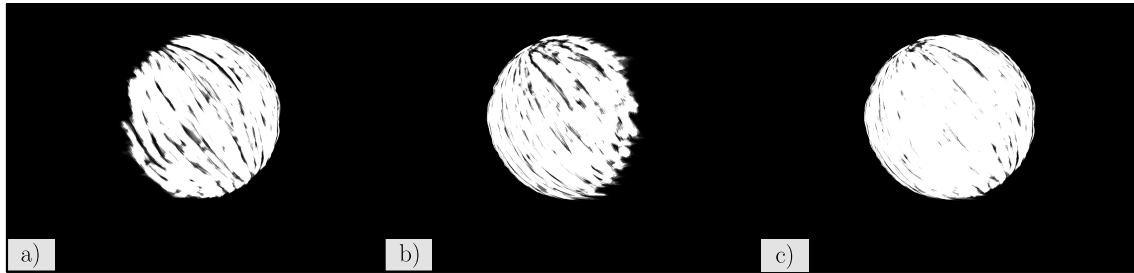


FIGURE 3.13 – Les quantités de peinture propagées, en avant provenant de t_{i-1} (a), en arrière provenant de t_{i+1} (b), apportent les informations nécessaires pour reconstituer la carte de quantité de matière cible (c). Ces cartes proviennent de la scène *Ballon de plage* 3.5.4 au temps t_{14} utilisant les images clés t_{10} et t_{20} .

3.5.2 Heuristique de sélection des marques

Cette sélection des marques utilise un score qui évalue la contribution d'une marque à se rapprocher de la carte de quantité de matière cible. Si le score dépasse zéro, la marque est intégrée dans l'image intermédiaire ; sinon, elle est écartée. Le processus de calcul de ce score est expliqué plus en détail dans la section 3.5.3.

La sélection des marques s'inspire des algorithmes de type glouton, où toutes les combinaisons d'un ensemble sont testées une à une afin de trouver l'optimum global. Dans notre cas, l'optimum global correspond à la combinaison de marques candidates qui respecte au mieux la carte de quantité de matière cible. Le nombre de possibilité pour le trouver correspond à 2^n avec n le nombre de marques. Ce qui avec un exemple comprenant seulement 100 marques candidates représente un nombre colossal de possibilités. Trouver l'optimum global n'est pas une option. C'est pourquoi nous avons conçu un algorithme qui, peu importe l'ordre choisi, aboutit à un optimum local correspondant à un résultat visuellement satisfaisant.

Les marques sont donc testées individuellement, mais de manière interdépendantes, car le résultat de chaque marque à tester dépend de celles qui ont été acceptées précédemment. Les marques candidates sont organisées en groupes dans un ordre cohérent. La première intuition ici est de procéder à l'évaluation des marques en suivant l'ordre de leur dessin initial. Cette approche offre l'avantage de maintenir des évaluations cohérentes tout au long du processus. Cependant, cette approche confère également une grande similarité entre deux images intermédiaires successives générées. On perçoit alors visuellement les marques se déplacer sur le support lorsque l'animation est jouée, ce qui donne l'impression d'un mouvement parfait avec une forte continuité temporelle. Or nous recherchons justement les niveaux de cohérence temporelle liés aux méthodes manuelles afin d'obtenir un résultat comparable à celui obtenu lorsque toutes les images sont faites à la main. Nous avons besoin de reproduire le scintillement d'une image à l'autre induit par une stylisation à faible continuité temporelle et des contours imprécis d'une stylisation à faible cohérence de mouvement. L'ordre aléatoire des

évaluations des marques offre la possibilité de diversifier les résultats, générant ainsi deux listes de marques très distinctes pour deux images intermédiaires successives tout en conservant le style des images clés. Toutefois, cette approche peut remettre en cause la validité du test en ne reproduisant pas un environnement similaire à l'image clé au moment du test de la marque candidate.

Cohérence des évaluations

Lorsqu'une marque est évaluée, elle est retirée de l'ensemble des marques candidates et insérée dans la liste composant l'image intermédiaire, à la même position et dans le même groupe qui lui ont été attribués dans l'ensemble. En fonction du résultat de l'évaluation la marque est acceptée ou rejetée, ce qui en cas de rejet entraîne sa suppression de liste de l'image intermédiaire. Pour que chaque évaluation reste valide, il est essentiel de maintenir la cohérence de cette liste, avant et après chaque évaluation en s'assurant que les règles de construction des groupes soient respectées (Section 3.4.2). Tout au long du processus de génération, de multiples insertions et suppressions de marques ont lieu dans la liste qui constitue l'image intermédiaire et altèrent la composition des groupes. Ces multiples insertions posent un problème en ce qui concerne l'état du pinceau virtuel, impactant notamment la couleur des marques. Ce problème est similaire à celui observé lors de la fusion des listes de marques candidates issues des propagations (voir Figure 3.11). Pour mieux illustrer le problème, voici un exemple pertinent :

Problème d'insertion Une marque ne mettant pas à jour l'état du pinceau est évaluée, puis insérée dans la liste de l'image intermédiaire à la même position et au sein de son propre groupe. Dans le cas où cette marque est la première de son groupe à être évaluée, son groupe dans l'image intermédiaire devient invalide. La première marque d'un groupe doit forcément modifier l'état du pinceau (état nettoyé ou état rempli) pour obtenir un groupe valide et des marques cohérentes. Cette situation provoque un rendu de la marque par la simulation de peinture qui ne correspond plus à celui de la marque dans l'image clé d'origine, ce qui engendre une perte de cohérence dans l'évaluation. Pour maintenir la cohérence des évaluations, nous proposons une méthode. L'idée est de garantir la cohérence des groupes de marques lors des évaluations, à savoir qu'un groupe n'est considéré comme cohérent que si la première marque de ce groupe met à jour l'état du pinceau. Pour ce faire, nous modifions l'état du pinceau virtuel des marques appartenant au même groupe que la marque testée, tout en respectant l'ordre établi. De cette façon, quelle que soit l'étape de la génération, la première marque du groupe met à jour l'état du pinceau virtuel, tandis que toutes les marques suivantes ne le font pas.

Limite Cette façon de faire n'est pas parfaite. Si l'on évalue une marque avec un ordre de dessin antérieur à une autre marque déjà évaluée et acceptée, on peut se poser la question si l'évaluation qui a été précédemment faite reste valide. En effet l'ajout de la nouvelle marque de la séquence modifie l'impact sur le support de la marque précédemment acceptée. Les changements concernant la précédente évaluation sont mineurs, car l'évaluation de la nouvelle marque tient compte de ces variations lors du choix de l'accepter ou de la rejeter. Cette approche permet de

profiter du caractère aléatoire de la sélection tout en maintenant des évaluations des marques qui ont du sens une fois rendues.

3.5.3 Évaluation des coups de pinceau

L'évaluation du score d'une marque repose sur deux critères : la comparaison de la quantité de matière présente sur le support par rapport à la carte de quantité cible $T(p)$ avant l'ajout de la marque évaluée dans la liste finale de l'image intermédiaire $C(p)$, et après $U(p)$. Ces différences de quantité de matière sont calculées d'abord pixel par pixel, puis les pixels ayant subi des variations à la suite de l'ajout de la marque sont additionnés.

$$\Delta_C(p) = |C(p) - T(p)|$$

$$\Delta_U(p) = |U(p) - T(p)|$$

Ensuite la variation des ces deux différences par rapport à la carte de quantité de matière cible constitue le score $\Delta(p)$ de la marque :

$$\Delta(p) = \Delta_C(p) - \Delta_U(p) + \lambda$$

Un score positif signifie que les changements dans la quantité de matière entre C et U se rapprochent de la carte de quantité de matière cible T . Nous ajoutons une prime λ au score si le pixel n'est pas encore couvert alors que de la peinture est attendue, c'est-à-dire $C(p) = 0 \wedge T(p) > 0$. La valeur de λ est constante et dépend du modèle de simulation de peinture utilisé pour l'évaluation. Nous la fixons de manière empirique à deux fois la quantité de matière moyenne des pixels couverts du support dans notre simulateur de peinture.

Découpage de la trace

Calculer le score des marques en tant que somme des scores des pixels pour accepter les marques avec un score positif et rejeter les autres peut entraîner un débordement excessif de peinture dans les résultats (Fig.3.14-c). Une marque peut avoir des morceaux de sa courbe avec un score élevé tandis que d'autres parties ont un score faible. Pour mieux faire correspondre la quantité de matière du support avec la carte de quantité de matière cible, nous proposons un mécanisme pour découper une marque en morceaux en fonction des scores des pixels qui appartiennent à la surface rendue de la marque.

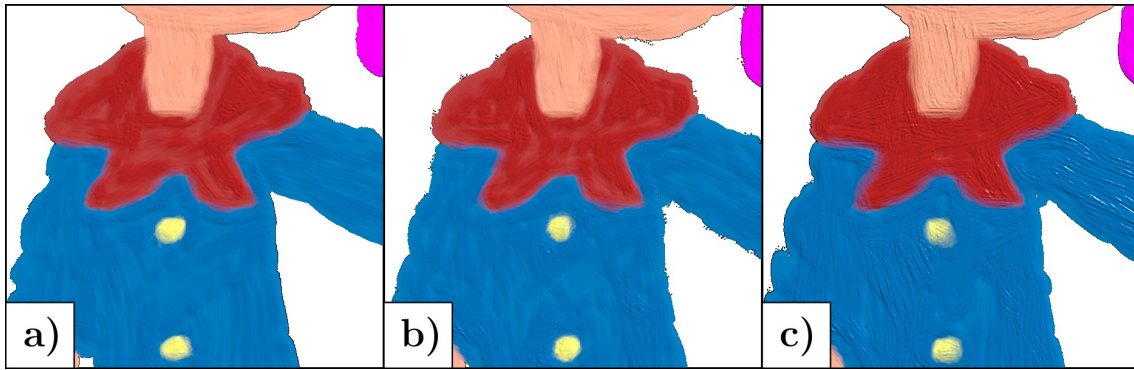


FIGURE 3.14 – Une comparaison entre la génération d’images intermédiaires avec et sans l’utilisation de la découpe des marques. La figure (a) présente l’une des deux images clés d’entrée utilisées pour la génération des marques, tandis que les figures (b) et (c) montrent les résultats de la génération pour la même image intermédiaire. Dans la figure (b), la découpe des marques est activée, ce qui conduit à une quantité de matière plus cohérente avec l’image clé. En revanche, dans la figure (c), la découpe des marques est désactivée, ce qui entraîne un surplus de peinture dans certaines régions.

Algorithme Pour découper une marque en fonction de son impact sur le support, l’objectif est d’associer les scores des pixels à la courbe. Pour y parvenir, il est nécessaire d’assigner le paramètre x correspondant à l’abscisse curviligne de la courbe $f(x)$ à chaque pixel. La première étape consiste donc à construire un maillage correspondant au rendu de la marque. Le paramètre de courbe est ensuite incorporé dans les données liées aux sommets du maillage. La seconde étape consiste à rendre le maillage dans le même espace que le rendu de la marque. Les paramètres de courbe associés aux sommets sont interpolés pendant le rendu, générant ainsi une texture de même dimension que le support. La nouvelle texture obtenue permet d’associer les paramètres de courbe à chaque pixel de la simulation et d’y intégrer le score de la marque.

Évaluation par morceaux Ensuite, la courbe de la marque est divisée en segments uniformément répartis le long de la courbe. Chaque segment représente une plage de valeurs du paramètre de courbe x . Le nombre de segments varie en fonction de la taille de la marque afin de maintenir une précision de découpe acceptable. En procédant ainsi, chaque segment regroupe un ensemble de pixels dont le paramètre de courbe x appartient à l’intervalle des valeurs du segment. Le score d’un segment est calculé comme la somme de tous les scores des pixels faisant partie du segment. Les segments avec un score positif indiquent les parties de la courbe de la marque à conserver. Les segments valides, pris successivement, forment les nouvelles marques qui seront ajoutées à la liste finale de l’image intermédiaire.

Ajout des morceaux retenus Les nouvelles marques créées, formées à partir des morceaux valides, possèdent les mêmes paramètres que la marque évaluée, à l’exception de celui concernant l’état du pinceau. Parmi ces nouvelles marques, celle qui correspond au début de la marque évaluée prend le même état du pinceau que

la marque d'où elle provient, c'est-à-dire celle qui a été évaluée. Les autres nouvelles marques ont leur état du pinceau défini comme «inchangé», afin de garder un rendu similaire à celui de la marque évaluée. Ces nouvelles marques sont ensuite insérées dans le même groupe et au même emplacement dans la liste finale de l'image intermédiaire, en suivant l'ordre temporel du début à la fin de la courbe de la marque évaluée.

Remarque Le maillage construit ne permet pas d'associer le paramètre de courbe au score de tous les pixels impactés par un changement de quantité de matière lors l'évaluation de la marque. C'est le cas pour les pixels subissant une variation de peinture due à l'insertion de la marque évaluée en milieu de séquence dans l'ordre de dessin. Dans cette situation spécifique, les marques ultérieures dans l'ordre de dessin peuvent être affectées par le dépôt en amont de peinture par la marque évaluée, créant à leur tour de nouvelles variations. Ce constat découle des échanges bidirectionnels de matière entre le support et le pinceau mais ne représente pas un problème pour la découpe de la marque.

Etude de l'impact La découpe des marques n'a pas d'impact perceptible sur les régions uniformément peintes, comme le montrent les figures 3.14-a et 3.14-b. Cela n'est pas le cas pour les régions peintes avec différentes couleurs (Figure 3.15) où la découpe devient visible. En pratique, ce problème ne se produit que dans de rares cas, comme lorsque la même région est peinte de deux manières différentes dans les images clés. Pour effectuer des mélanges de peinture entre deux régions, nous utilisons les marques d'étalement, qui, comme nous le verrons dans la section 3.5.4, ne sont pas découpées.

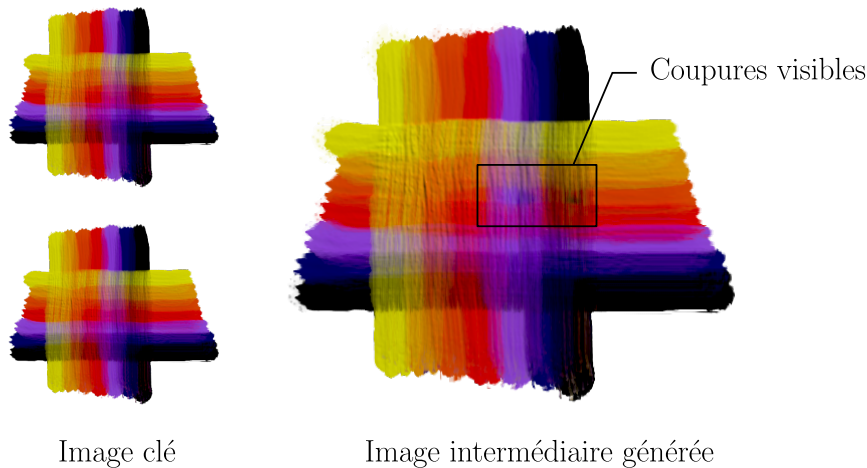


FIGURE 3.15 – Impact de la découpe des marques sur l'apparence des images intermédiaires générées lorsque des images clés avec des régions colorées complexes sont utilisées. À des fins d'illustration, les deux images clés sont identiques.

3.5.4 Les marques d'étalements

Les marques d'étalement sont utilisées par les artistes pour mélanger les couleurs et lisser les transitions de couleur sur le support. Une marque d'étalement

est définie comme telle dès lors qu'elle appartient à un groupe de marques, dont la première marque du groupe possède l'état du pinceau virtuel *nettoyé*. Ces marques nécessitent un traitement spécifique, car elles ne déposent pas plus de peinture sur le support qu'il n'y en a déjà. Jusqu'à présent, nous avons détaillé le processus de génération d'une image intermédiaire en considérant le cas général de marques ayant un état du pinceau «rempli» ou «inchangé». Cependant, pour les marques d'étalement, quelques ajustements sont effectués dans l'étape de formation de la liste des marques candidates, ainsi que dans le processus aléatoire qui sélectionne la marque à évaluer.

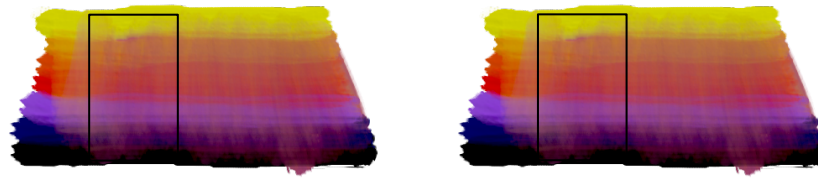
Dans nos expériences, nous constatons que lorsque toutes les marques d'étalement des deux images clés sont acceptées, elles créent un lissage excessif dans les images intermédiaires. C'est pourquoi nous concevons un mécanisme de sélection spécifique, sans découpe de traits pour les traits de lissage. Nous filtrons les marques d'étalement candidates pour ne considérer qu'un sous-ensemble d'entre elles provenant de l'image clé précédente et de l'image clé suivante, en fonction de la position de l'image intermédiaire t_i dans la séquence que nous générons. Pour garantir que plus l'image intermédiaire est proche d'une image clé, plus les marques d'étalement de cette image clé ont de chances d'être considérées comme candidates, nous sélectionnons les marques d'étalement selon le ratio ρ :

$$\rho = \frac{t}{N}$$

et prenons $(1 - \rho) \cdot \kappa_0$ marques d'étalement de l'image clé précédente et $\rho \cdot \kappa_N$ marques d'étalement de l'image clé suivante, où κ_i est le nombre de marques d'étalement dans l'image clé t_i .

L'ordre aléatoire des évaluations peut entraîner des situations où une marque d'étalement est évaluée avant que la peinture qu'elle est censée étaler ne soit déposée sur le support. Dans de tels cas, l'évaluation devient invalide. Afin de garantir la cohérence de chaque évaluation de marque d'étalement, nous proposons de les évaluer de manière aléatoire, mais seulement après que toutes les autres marques aient été évaluées. Il est important de rappeler que l'ordre aléatoire des évaluations de marque n'affecte en rien l'ordre de dessin établi lors de la création de la liste de marques candidates (Section 3.4.2). De cette façon nous nous assurons que lorsqu'une marque d'étalement est évaluée, elle étale correctement la peinture au moment où elle doit être appliquée selon l'ordre de dessin. Cela se fait tout en conservant les avantages du processus aléatoire.

Bien que la découpe des autres marques puisse améliorer la répartition générale de la peinture, les avantages potentiels de la découpe des marques d'étalement sont beaucoup moins significatifs, et les découper ajoute des discontinuités visibles sur le support (Figure 3.16). Les marques d'étalement sont acceptées ou rejetées dans leur intégralité sans passer dans le processus de découpe.



Images clés

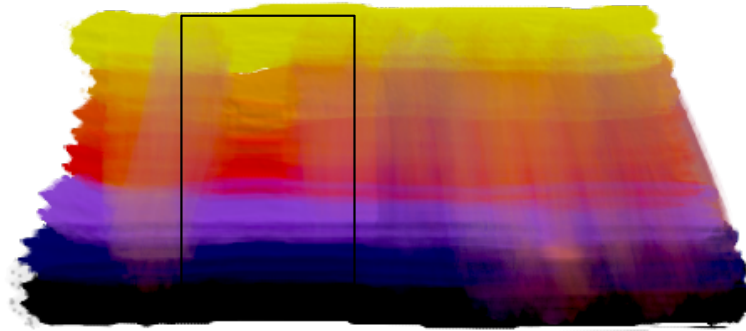


Image intermédiaire

FIGURE 3.16 – La découpe des marques d'étales ajoute des discontinuités visibles sur le support qui ne sont pas initialement présentes dans les images clés.

Chapitre 4

Synthèse des résultats obtenus

Ce chapitre propose une synthèse des résultats obtenus. Ces travaux de thèse explorent le domaine de la génération d'images à base de marques dans un style peinture, et mettent en lumière de nouveaux défis. Le champ de recherche demeure ouvert, et nous mettons en évidence les perspectives envisagées pour ces travaux.

4.1 Présentation des animations générées

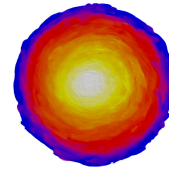
Notre approche dans la sélection des animations générées suit une progression logique, où chaque nouvelle scène représente un test de complexité croissante. Les premières animations générées servent de preuve de concept pour les mouvements simples tels que la translation, la rotation et la mise à l'échelle (zoom). Ensuite, nous avons créé des scènes avec des mouvements plus complexes afin de mettre à l'épreuve notre méthode. Nous avons également réalisé des animations comprenant des occlusions d'objets et avons augmenté le nombre de marque pour plus de complexité. Enfin, nous avons exploré différentes sources de mouvement, autres que les champs vectoriels rendus, dans le but d'élargir les cas d'utilisation de notre méthode et de la rendre pertinente pour une utilisation professionnelle. Un résumé sous forme de tableau récapitulant les informations des scènes est disponible en Section [4.1.1](#).

L'approche utilisée pour peindre les images clés est la suivante. L'artiste peut afficher une image sur le support et s'en servir comme guide pour peindre les images clés. Lorsque le champ vectoriel utilisé est rendu depuis une scène 3D ou extrait depuis une vidéo il est possible d'utiliser la colorisation de la scène comme guide. Lorsque le guide tient compte du mouvement, les images clés peuvent être peintes directement. Dans le cas contraire, on utilisera de manière intuitive la propagation en avant pour peindre la seconde image clé ([3.2.2](#)).

Pomme : cette scène représente une pomme translaturée en 2D verticalement et horizontalement. Le mouvement est obtenu par le rendu d'un plan animé sur *Blender*.

Pizza : cette scène représente une pizza *pepperoni* effectuant une rotation 2D. Le mouvement est obtenu par le rendu d'un disque animé sur *Blender*. Elle est décomposée en deux calques : l'un représente la pâte à pizza et l'autre la garniture.

Supernova : cette scène ne représente pas formellement un objet ou un concept. Elle met en évidence le comportement de la génération de d'images lorsque le mouvement est un zoom. Ce mouvement est obtenu par le rendu d'une sphère sur *Blender*. Cette scène contient également un nombre conséquent de marques d'étalement qui viennent fortement mélanger les couleurs et lisser le rendu.



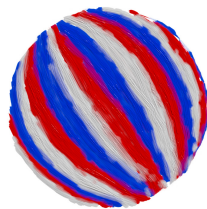
(a) Animation : Pomme

(b) Animation : Pizza

(c) Animation : Supernova

Ballon de plage : cette scène représente un ballon de plage effectuant une rotation 3D non alignée aux axes. Le mouvement est obtenu par le rendu d'une sphère sur *Blender*. Elle met en évidence le comportement de notre méthode lorsque des occlusions sont présentes au cours de l'animation.

Ellie : cette scène représente le personnage *Ellie* issue de l'animation *Sprite fright* de SIDDI et al. [92] effectuant plusieurs changement de position. Le mouvement provient de la scène 3D disponible dans les échantillons de *Blender*. Les images clés de l'animation comporte un millier de marque.



(a) Animation : Ballon de plage

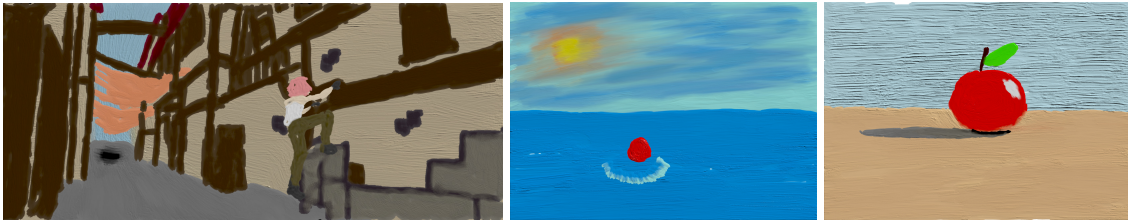
(b) Animation : Ellie

Alley : cette scène provient du court métrage *Sintel*, une animation 3D réalisée par LEVY et al. [64]. Elle utilise le flux de mouvement extrait directement de la vidéo de l'animation. Elle est composée de quatre calques, tous utilisant le même mouvement. Le calque dédié à l'animation du personnage est édité à la main après sa génération, afin d'enlever les parties où la peinture n'a pas pu suivre correctement le mouvement en raison d'imprécisions. Aucun ajout de peinture n'a été effectué lors de l'édition.

Balle rebondissante : cette scène illustre une balle rebondissant sur une surface d'eau. Elle est constituée de trois calques, dont celui de la balle rebondissante utilisant un mouvement conçu par le dessin (3.3.1). L'arrière-plan du ciel est statique

et n'a pas été généré. Il a été peint une seule fois et sa reproduction a été appliquée à toutes les autres images. En revanche, le calque représentant la surface de l'eau est généré sans mouvement, créant ainsi un effet de *faux-fixe*. Un effet d'éclaboussures d'eau et d'ombrage a été ajouté par-dessus la génération afin de donner davantage de vie à l'animation.

Pomme éclairée : cette scène représente le mouvement de l'ombre d'une pomme et des reflets spéculaires de sa surface lorsqu'une lumière est en mouvement. Elle est constituée de quatre calques. L'arrière plan est représenté par des coups de pinceau très larges déposant une quantité importante de peinture dans le but d'obtenir une texture différente au rendu, tandis que le sol, toujours dans le même calque, est lissé par des coups de pinceau d'étalement. Ce calque est généré sans mouvement en entrée donnant également un effet de *faux-fixe*. Le mouvement de l'ombre et des reflets spéculaires ont été rendus sur *Blender*. Le mouvement de l'ombre est obtenu par le mouvement d'un plan en rotation se déformant en longueur et en largeur de sorte à simuler le mouvement réel d'une ombre soumise aux mêmes conditions d'éclairage de la scène. Le mouvement des reflets spéculaires est obtenu en projetant le maillage d'un plan sur la surface d'une pomme 3D. Le plan projeté glisse sur la surface de la pomme de sorte à obtenir un mouvement acceptable se rapprochant des reflets spéculaires sur une pomme dans les mêmes conditions d'éclairage.



(a) Animation :
Alley

(b) Animation : Balle re-
bondissante

(c) Animation : Pomme
éclairée

4.1.1 Tableau des scènes

Scène ID	Calque ID	Images	Images clés	Strokes	Source du champ vectoriel	Extrapolation	Edition
Pomme	Principal	52	0-10-20-30-40-51	110	Rendu	Oui	Non
Pizza	Pâtes à pizza	52	0-10-20-30-40-51	260	Rendu	Oui	Non
	Garniture	52	0-10-20-30-40-51	150	Rendu	Oui	Non
Supernova	Principal	52	0-15-25-38-51	500	Rendu	Oui	Non
Ballon de plage	Balle	52	0-10-20-30-40-51	600	Rendu	Oui	Non
	Contours	52	0-10-20-30-40-51	100	Pas de mouvement	Non	Non
Ellie	Principal	60	0-10-20-30-40-50 60	1300	Rendu	Oui	Non
Alley	Arrière plan	50	1-10-20-30-40-50	400	Extraction vidéo	Non	Non
	Personnage	50	1-5-10-15-20-25 30-35-40-45-50	225	Extraction vidéo	Non	Oui
	Sol	50	1-6-11-16-21-26 31-36-41-46-50	200	Extraction vidéo	Non	Non
	Architecture	50	1-6-11-16-21-26 31-36-41-46-50	550	Extraction vidéo	Non	Non
Balle rebondissante	Ciel	51	Aucune	0	Pas de mouvement	Non	Oui
	Mer	51	0-10-20-30-40-50	100	Pas de mouvement	Non	Non
	Balle	51	0-5-10-21-22-23 30-35-41-45-50	30	Dessiné en 2D	Oui	Non
	Effets	51	Aucune	0	Pas de mouvement	Non	Oui
Pomme éclairée	Arrière plan	52	0-10-20-30-40-51	135	Pas de mouvement	Non	Non
	Premier plan	52	0-10-20-30-40-51	320	Pas de mouvement	Non	Non
	Ombres	52	0-5-10-15-20-25 30-35-40-45-51	120	Rendu	Non	Non
	Spéculaire	52	0-5-10-15-20-25 30-35-40-45-51	40	Rendu	Non	Non

4.1.2 Impact du choix du simulateur

Toutes les animations générées l'ont été en utilisant le même modèle de simulation de peinture (Section 3.1.2). Notre méthode est compatible avec d'autres modèles utilisant les notions de quantité de matière et modélisant les échanges bidirectionnels de matière entre le support et le pinceau, comme les modèles physiques. Néanmoins, les résultats ne peuvent être garantis avec les modèles qui n'intègrent pas les quantités de matière et nécessitent de repenser l'évaluation des marques.

4.2 Implémentations

La méthode de génération d'images intermédiaires est implémentée en *C++* dans le simulateur de peinture. La partie graphique utilise *OpenGL* pour effectuer les calculs sur le processeur graphique *GPU*. Les marques (Section 3.4.1) ainsi que les cartes de quantité de matière (Section 3.5.1) sont advectés sur le *GPU* en utilisant le mécanisme des *compute shaders*. Les marques sont transmises au *GPU* par paquet. En revanche, la régularisation de la courbe des marques (Section 3.4.1) est effectuée sur le processeur centrale *CPU* et utilise l'implémentation d'*Eigen* pour les opérations vectorielles [42]. Les listes de marques candidates sont ordonnées sur *CPU* (Section 3.4.2). La sélection des marques (Section 3.4.2) est réalisée de manière concurrentielle en utilisant plusieurs processus *CPU*. Le processus *CPU* principal se charge de recevoir les requêtes graphiques et de les traiter en utilisant le *GPU*. Les autres processus ont pour responsabilité la génération d'une image intermédiaire, ainsi que la gestion de la logique et l'envoi de requêtes graphiques.

4.2.1 Performance

Le temps de calcul de notre génération d'images intermédiaires est conséquent. A titre d'exemple, des séquences avec des images clés composées d'une centaine de marques prennent seulement quelques minutes par image intermédiaire pour être calculées. Cependant, pour des séquences d'images clés plus complexes composées d'un millier de marques, le temps de calcul peut atteindre jusqu'à une heure par image intermédiaire. Bien que la création d'images clés soit effectuée en temps réel par l'artiste, le calcul des images intermédiaires est un processus hors ligne dans sa forme actuelle. Ces temps de calculs ne tiennent pas compte de l'extrapolation d'un mouvement incomplet. En effet, cette partie de nos travaux s'utilise comme outil en tant que pré-traitement.

L'analyse des performances de la méthode met en évidence un goulot d'étranglement au niveau de la simulation de peinture. Le temps de calcul de notre méthode est directement proportionnel au nombre de marques candidates utilisées. En plus d'être de type glouton, notre algorithme de sélection des marques nécessite le rendu d'une liste croissante de marques à étape d'évaluation d'une marque. Par conséquent, le coût de simulation augmente de manière linéaire lorsque chaque marque évaluée est acceptée afin de composer l'image intermédiaire finale. Des mesures relatives aux temps de calcul des différentes étapes de notre génération d'images intermédiaires sont présentées dans le tableau 4.1.

4.2.2 Mesure des temps de calcul

Etape	Total	Moyen / image	Moyen / marque
Génération	171 min	19 min	
- Propagation	12,9 s	1,4 s	
- - Advection des marques	1,5 s	83,9 ms	
- - Régularisation	10,9 s	0,6 s	0,7 ms
- - Advection des quantités	0,3 s	18,2 ms	
- Ordonnancement	1 ms	<0 ms	
- Sélection	168 min	18,7 min	
- - Simulation	154 min	17,1 min	1,0 s
- - Evaluation	12,9 min	1,4 min	88,9 ms
- - Découpe	58,3 s	6,5 s	6,8 ms

TABLE 4.1 – Mesures prises lors de la génération de la séquence *Ballon de plage* utilisant les images clés t_{10} et t_{20} , générant ainsi neuf images intermédiaires. Les temps de calcul ont été mesurés à partir d’une machine comprenant un processeur *Intel(R) Xeon(R) CPU E5-1630 v4 @ 3.70GHz* et une carte graphique *NVIDIA GeForce GTX 1080/PCIe/SSE2*.

4.3 Comparaison avec les animations traditionnelles

Les travaux de cette thèse utilise un démonstrateur de peinture. C’est la raison pour laquelle nous comparons nos résultats avec des animations traditionnelles utilisant la peinture telles que *Loving Vincent* de KOBIELA et al. [58] ou *La traversée* de MIAILHE et al. [77].



(a) Animation : Loving Vincent



(b) Animation : La Traversée

FIGURE 4.4 – Extraits d’animations traditionnelles. Images tirées des oeuvres [58, 77].

Nos résultats présentent une cohérence temporelle similaire à celle de ces deux animations traditionnelles mentionnées. Notre algorithme utilise une sélection aléatoire des marques pour introduire de légères variations dans chaque image intermédiaire. Cette approche contribue à briser la continuité temporelle et à renforcer la perception en 2D de l’animation, comme décrit par BÉNARD et al. [12],

en rompant avec la perception de marques synthétiques en mouvement. Cependant dans *Loving Vincent* ou *La Traversée*, seules les parties mobiles de l'image sont repeintes par-dessus l'image précédente. Ce constat découle de la technique utilisée «*sous la caméra*» et du temps que prendrait la production s'il fallait repeindre toute la toile à chaque image. Dans *Loving Vincent*, l'artiste peint sur une seule et unique toile avec de la peinture à l'huile, attendant que la peinture sèche avant de repeindre par-dessus. Dans *La Traversée*, l'artiste travaille sur des calques en verre, ce qui lui permet de repeindre uniquement les parties mobiles et de retirer facilement la peinture entre chaque image à produire. Non seulement notre méthode permet de repeindre aisément l'intégralité du support à chaque image, mais elle permet également de faire varier et de reproduire la cohérence de ces animations grâce à son utilisation par calque. Cela donne à l'artiste la liberté de créer des calques statiques ou mobiles, comme dans notre scène animée *Balle rebondissante* (voir la section 3.5.4).

Avant de produire le film animé *Loving Vincent*, les scènes ont été filmées avec des acteurs en chair et en os. Ensuite, chaque plan a été peint à l'huile par un artiste. Les prises de vues réelles servent alors de guide à l'artiste et fournissent des informations sur le mouvement. Cette méthode est similaire à celle que nous avons utilisée pour créer notre scène *Alley* (Section 3.5.4), où nous avons montré qu'il était possible d'utiliser le flux optique d'une vidéo pour générer une animation. Cela renforce la validation de notre flux de travail. En effet on peut utiliser partiellement ou totalement notre méthode pour un usage flexible de la technique d'animation traditionnelle «*sous la caméra*» ou de la technique de *pose à pose*.

Comme le montre *Loving Vincent*, il n'est pas toujours aisé d'obtenir un mouvement pour la réalisation de l'animation. Parfois, cela requiert des compétences que les animateurs 2D ne possèdent pas. Pour illustrer qu'il est possible de créer le mouvement par le dessin et rester dans le cadre d'un flux de travail exclusivement 2D, nous avons conçu la scène de la *Balle rebondissante*. Cependant, bien que cette façon de procéder soit pratique pour la création de mouvements simples et uniformes, ce n'est pas le cas pour les mouvements de scène plus complexes. La création du mouvement par cette méthode implique de prédire le résultat de l'extrapolation du mouvement et nécessite plusieurs itérations avant d'obtenir un mouvement satisfaisant. Nuancé dans ses possibilités, notre méthode offre plusieurs outils pour concevoir une animation 2D entièrement par ordinateur, sans avoir à recourir à une scène 3D préalablement rendue ou à des prises de vue réelles.

4.4 Application aux cas problématiques d'animations

Les problèmes d'occlusions entre plusieurs objets avec des mouvements différents sont gérés par l'utilisation de calques par les artistes. Nous suivons la même approche en utilisant un mouvement par calque mais nous abordons néanmoins le problème dans l'extrapolation du mouvement (Section 3.3.2).

Notre méthode permet de reproduire les mouvements simples comme la translation ou la rotation. L'extrapolation du mouvement permet de peindre sans se

soucier des limites des formes et l'algorithme de génération des images intermédiaires permet d'obtenir une quantité de matière proche de celles des images clés en disposant et découpant de manière logique les marques sur le support. Nous avons également évalué notre méthode avec la scène *Supernova* (3.5.4), qui implique un effet de zoom. Cette situation représente un défi en raison des changements rapides d'échelle et de perspective, pouvant conduire à des zones non remplies de peinture sur le support. Dans ce cas, la simple propagation en avant des marques de la première image clé ne suffit pas à garantir un remplissage correct du support des images intermédiaires, lorsque le zoom amplifie la proportion d'un objet sur le support. Notre méthode surmonte cette difficulté en utilisant la propagation des marques des deux images clés pour maintenir la cohérence des formes dans chaque image intermédiaire, proposant alors assez de marques pour remplir le support. Cette résolution du problème suppose que les images clés et le mouvement capturent les informations de mise à l'échelle. Cette conception permet à notre algorithme de s'adapter en douceur aux changements d'échelle de la scène en exploitant le champ de mouvement bidirectionnel, ce qui lui permet de gérer le zoom avant et arrière de manière transparente.

Les animateurs 2D utilisent souvent une technique appelée le *faux-fixe* pour créer des effets stylisés. Le *faux-fixe* fait référence à la technique consistant à dessiner plusieurs fois des parties fixes, qui produit un léger tremblement ou une fluctuation dans les lignes et donne une impression de mouvement et d'animation malgré le caractère statique des dessins. C'est typiquement ce que cherche à reproduire notre méthode, mais au lieu de réaliser le *faux-fixe* sur des lignes, notre méthode le produit sur toute l'image. On peut considérer, par extension, que l'objectif de produire un *faux-fixe* sur toute l'image équivaut à simuler les imperfections qui se produiraient si une même image était dessinée à la main plusieurs fois. Dans cette optique, notre méthode génère des images intermédiaires en se basant sur deux images clés, qui une fois propagées par le mouvement représentent le même état de la scène. Elle reproduit ces imperfections entre chaque paire d'images intermédiaires grâce aux légères variations induites par le caractère aléatoire de la sélection lors de l'évaluation des marques issues de la propagation des images clés. Nous montrons l'effet *faux-fixe* dans les scènes *Pomme éclairée* et *Balle rebondissante*.

De la même manière que pour la gestion du zoom, notre méthode, de part son algorithme, tire parti des informations présentes dans les deux images clés pour résoudre les problèmes d'auto-occlusion illustrés dans la scène *Ballon de plage*. Cependant, la gestion de l'auto-occlusion suppose un ajout et un retrait précis de la peinture sur le support au cours de l'animation. Ces deux aspects dépendent de la qualité de la carte de quantité de matière cible, laquelle repose sur la précision du champ de mouvement. Lorsque le champ de mouvement est extrapolé, sa précision dépend alors de la finesse du maillage utilisé. Dans l'exemple du *Ballon de plage*, les ajouts ou retracts de peinture ont lieu sur le bord de la surface de l'objet. Toutefois, lorsque la précision du champ de mouvement est insuffisante, des sauts de peinture apparaissent au cours de l'animation, principalement au niveau des zones du support où la peinture est advectée en dehors de la surface de l'objet. Ces sauts de peinture sont perçus à chaque transition d'une image clé à une image

intermédiaire, mettant ainsi à jour la quantité de matière sur le support.

4.5 Limitations de la méthode

Impact de la qualité du champ de mouvement La qualité de nos résultats dépend largement du champ de mouvement fourni en entrée. Les champs de mouvement rendus capturent avec précision les vecteurs de mouvement de vitesse, mais ils génèrent des champs de mouvement non bijectifs car ils ne tiennent pas compte de l'accélération. L'utilisation d'un mouvement non bijectif ne permet pas de superposer parfaitement les propagations des marques ou des quantités de matière aux temps des images intermédiaires, ce qui engendre un dédoublement de certaines zones. Ce désagrément est observable au niveau des yeux du personnage *Ellie* de la scène du même nom en figure 4.5, qui effectuent de brusques changements avec de fortes accélérations. De plus, ces champs de mouvement ne représentent que le mouvement à l'intérieur des formes des objets, ce qui nécessite une extrapolation.

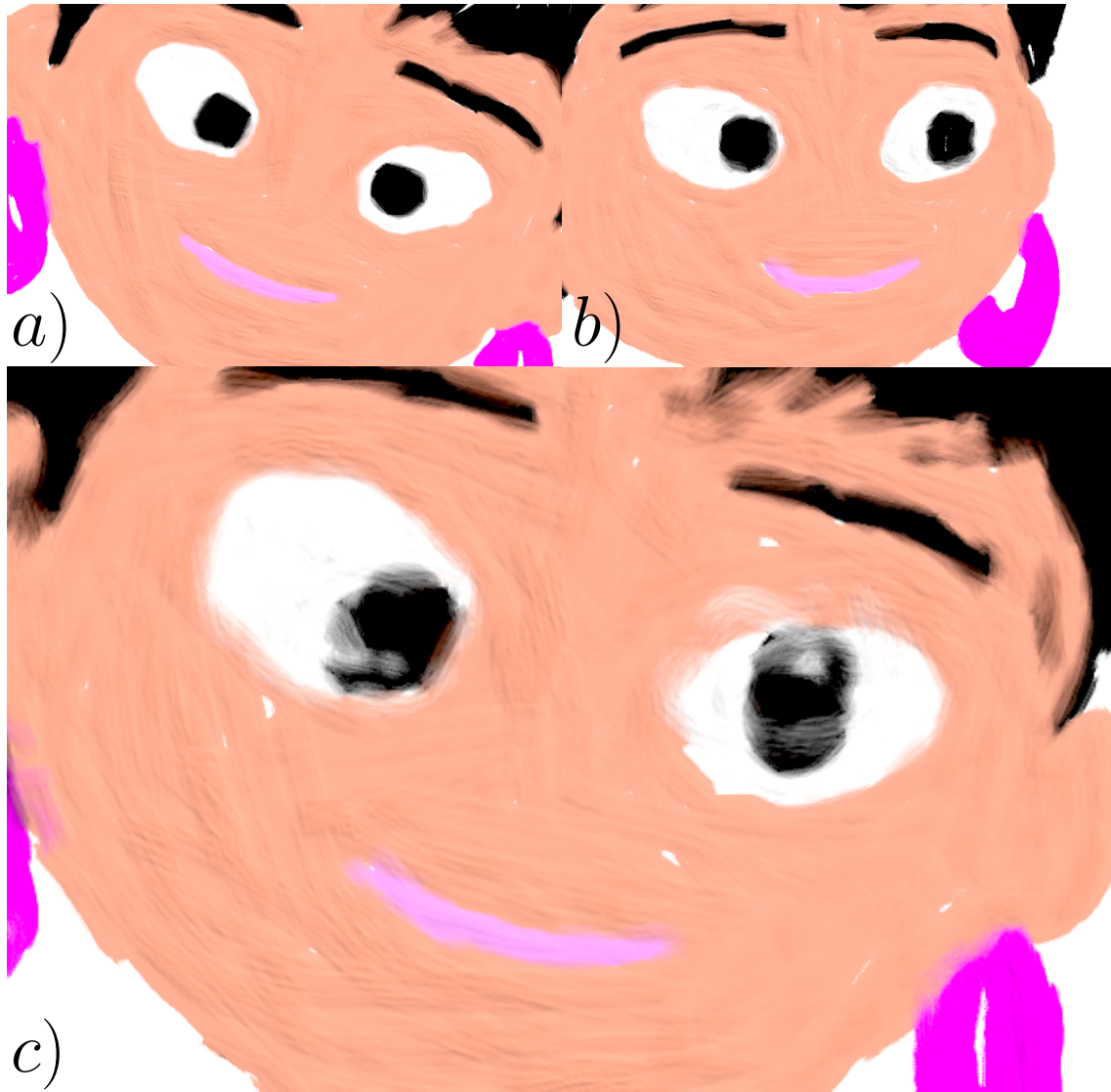


FIGURE 4.5 – Dédoublément des yeux et pupilles à cause de l'utilisation d'un mouvement non bijectif. Les images clés a) et b) sont respectivement aux temps t_{40} et t_{50} tandis que l'image intermédiaire c) est au temps t_{48} .

Une mauvaise qualité du champ de mouvement peut également découler de la méthode d'estimation, notamment lors de l'utilisation d'un flux optique vidéo. Les méthodes d'estimation de mouvement ne permettent pas d'obtenir un mouvement parfait lorsqu'il s'agit de capturer le mouvement d'une scène entière, notamment dans les régions présentant des changements de visibilité [120]. Il suffit d'un pixel encodant un mouvement erroné pour advecter de la mauvaise façon le point associé à la polygone d'une marque sur le support. Notre méthode est sensible aux bruits du champ de mouvement utilisé, en particulier au niveau des contours d'un objet, là où l'estimation du mouvement a le plus de chance de produire une erreur. Ce problème est illustré dans notre scène *Alley* et mis en évidence sur la Figure 4.6.



FIGURE 4.6 – Génération du personnage de la scène *Alley* avec un champ de mouvement comportant beaucoup d'imprécisions. La sortie de la génération est éditée à la main afin de retirer la peinture qui n'a pas été correctement advectée. Enfin le calque généré est composé avec les autres pour former l'image finale.

Une mauvaise advection se traduit par un dédoublement des formes due à un mouvement non bijectif et des déformations qui ne respectent pas la direction ar-

tistique, engendrées par des valeurs erronées dans les vecteurs du champ de mouvement. Cependant, ces imperfections du champ de mouvement, qui engendrent une mauvaise advection, ne remettent pas en cause notre approche. Elles nécessitent généralement une édition manuelle après la génération. Plus l'espacement des images clés est grand et plus la qualité du mouvement aura un impact sur la génération. L'advection successive accumule les erreurs d'imprécision d'image en image. Pour limiter cette accumulation, l'artiste peut resserrer l'espacement des images clés et se servir de la génération comme base ou guide pour peindre une nouvelle image clé. Notre méthode n'est pas parfaite, s'utilise de manière itérative et la qualité du mouvement est l'élément clé pour obtenir un bon résultat.

Trous de peinture Les images intermédiaires générées par notre méthode peuvent contenir des zones vides où la matière est absente. Cette absence de matière se produit lorsque aucune marque candidate n'est disponible pour remplir la zone. L'advection peut créer des cassures dans la couverture de matière sur le support en perturbant la répartition des marques. Les trous de peinture qui en résultent sont fins et peu visibles une fois la simulation de peinture effectuée. Cela devient un problème lorsque les marques sont advectées en fonction d'un mouvement dont la transformation n'est pas capturée par les images clés. Par exemple, lorsque le mouvement correspond à un zoom avant suivi d'un zoom arrière, avec deux images clés identiques, les deux images clés ne capturent pas la transformation. Cela entraîne inévitablement des zones sans peinture, car il n'y a pas suffisamment de marques candidates pour remplir la nouvelle surface de l'objet dans les images intermédiaires générées, une fois le zoom avant effectué. La disposition des images clés dans la séquence est un facteur important qui joue sur ce phénomène. De manière générale, notre méthode permet d'obtenir une couverture suffisante et ne nécessite pas d'intervention manuelle dans la mesure où la transformation engendrée par le mouvement est capturée par les deux images clés.

Impact de l'ajustement des polygones Notre algorithme alterne entre la représentation des marques sous forme discrète et sous forme paramétrique lors de la propagation d'une image à l'autre (Section 3.4.1). Cela engendre une accumulation des erreurs d'ajustement des polygones, associées aux marques, par des morceaux de Bézier lors d'une propagation d'une image à l'autre. En conséquence, les courbes des marques ont tendance à se lisser au fil de leur propagation. En pratique, ce problème dépend de la précision de l'ajustement de courbe, ce qui nécessite de trouver un compromis entre des performances acceptables et un ajustement parfait. Dans notre cas, nous n'observons pas de lissage de courbe lors de la propagation sur une dizaine d'images, grâce aux paramètres d'ajustement que nous avons utilisés (Section 3.1.1). Dans l'hypothèse où l'on générerait un nombre plus important d'images intermédiaires, ce phénomène peut avoir un impact visuel. Il n'est pas possible d'advecter une courbe sous forme paramétrique, car cela ne permettrait pas d'associer correctement le mouvement à la courbe, ni de calculer les nouveaux points de contrôle à partir du mouvement.

Limitation des performances sur le flux de travail Notre méthode se présente comme un outil destiné aux animateurs et artistes professionnels du milieu de

la création d'animation 2D. Les animateurs dessinent les images clés manuellement puis génèrent les images intermédiaires automatiquement par ordinateur. Ils peuvent ainsi vérifier la cohérence de la séquence et décider soit d'éditer les images intermédiaires, soit de modifier ou de créer de nouvelles images clés, puis répéter le processus de génération automatique. Notre méthode est donc adaptée à une utilisation itérative, voire interactive. Cependant, les performances que nous obtenons à ce jour (Section 4.2.1) ne permettent pas de l'envisager ainsi, mais plutôt de manière hors ligne. Nous proposons de répartir le nombre de marques dans de multiples calques et d'utiliser la méthode par calque pour réduire les temps de calcul. Nous ne remettons pas en cause notre approche, car elle présente un fort potentiel d'amélioration, et nous considérons ces travaux comme une base solide pour de futures avancées.

Sémantique intergroupe L'algorithme d'ordonnancement que nous proposons permet de former des groupes de marques cohérentes et indissociables entre elles mais ne tient pas compte de la sémantique associée à chaque groupe. Nous avons opté pour un entrelacement des groupes en fonction de la position de chaque groupe (Section 3.4.2) ce qui produit un résultat satisfaisant lorsque les images clés suivent la même logique de construction. Dans le cas contraire, certains groupes, provenant de différentes images clés mais qui partagent la même sémantique, peuvent se retrouver, après la génération de l'image intermédiaire, ordonnés de manière différente dans la séquence. Un exemple de ce problème est illustré figure 4.7.

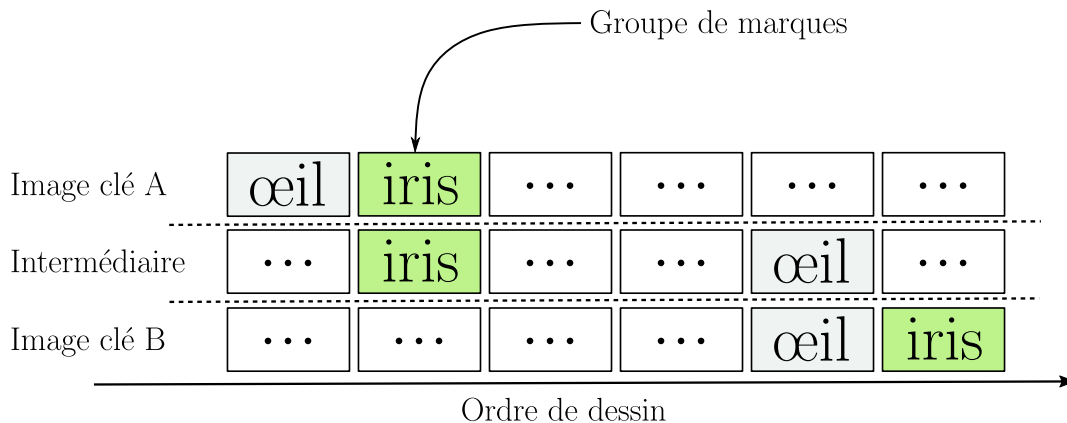


FIGURE 4.7 – Dans cet exemple, les marques du groupe *œil* de l'image clé *A* et les marques du groupe *iris* de l'image clé *B* ont toutes été refusées lors de l'évaluation. L'image intermédiaire ainsi générée présente un ordonnancement des groupes de marques, d'abord *iris* puis *œil*, qui n'est pas présent dans les images clés. Cet ordonnancement ne respecte pas la façon dont ont été construites les images clés.

4.6 Perspectives des travaux menés

Approche génétique Les algorithmes génétiques peuvent offrir de meilleures performances que les approches gloutonnes pour la génération d'images intermédiaires à partir d'un ensemble de marques candidates. Ce type d'approche est

adapté à notre contexte car les marques peuvent être analogues à une population avec des caractéristiques très diverses. Cette approche permet de créer de nouvelles marques, par opération de combinaison et de mutation, à partir de celles des images clés. De cette façon, l'esprit de notre méthode est conservé car le style défini dans les images clés est reproduit dans les images intermédiaires à partir d'éléments et d'informations tirés des marques directement faites par l'artiste. Le domaine génétique peut être représenté par l'ensemble des traits, paramètres et caractéristiques des marques des images clés. La fonction heuristique, ou d'ajustement, doit être repensée afin de ne pas reproduire les limitations actuelles. Il faudra alors veiller à préserver les variations légères entre deux images intermédiaires successives pour conserver l'aspect fait main de la génération par ordinateur. Si l'approche génétique ne permet d'obtenir ces variations, il est possible d'envisager d'introduire des modifications de la courbe des marques par du bruit.

Apprentissage automatique Pour améliorer notre méthode, il serait possible d'explorer une approche plus avancée pour la création des groupes ou l'ordonnement des marques des images clés. Une piste potentielle est d'utiliser des techniques d'apprentissage automatique pour identifier les marques ayant des propriétés sémantiques similaires, tels que ceux décrivant des contours des formes, des effets d'éclairages ou des éléments de premier plan et d'arrière-plan. En effet, lorsque l'artiste peint ou dessine, il dépose les marques en suivant une logique qui lui est propre, mais qui, le plus souvent, suit des codes ou des motifs. Avoir une connaissance a priori de ces informations permettrait de traiter les marques différemment en fonction des groupes sémantiques auxquels elles peuvent appartenir et de proposer une solution à l'ordonnement intergroupe, conduisant ainsi à une animation plus fidèle et à une génération plus efficace.

Le mouvement L'obtention du mouvement par le dessin 2D peut être poussée au delà de ce que nous proposons afin d'élaborer des mouvements plus complexes. Pour rappel, notre définition du mouvement par le dessin repose sur la description du mouvement 2D d'un pixel au fil du temps dans l'espace du support à l'aide d'une courbe. Bien que transcrite en 2D, l'information transmise par l'artiste peut être représentée sous différentes formes (Figure 4.8). On peut dès lors imaginer de nouvelles transitions entre ces différentes façons de représenter l'information de mouvement au fil du temps en introduisant à la fois de nouvelles annotations et de nouveaux opérateurs [26].

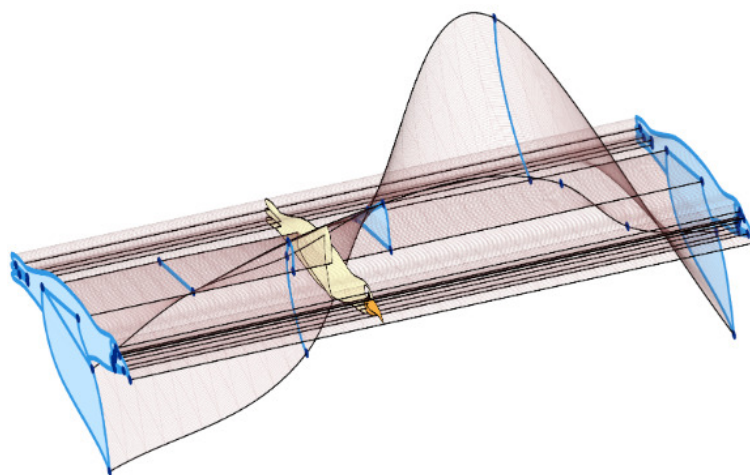


FIGURE 4.8 – Représentation temporelle annotée du mouvement d'un oiseau.
DALSTEIN et al. [26].

Chapitre 5

Conclusion

Nous pouvons repositionner notre méthode par rapport aux critères des enjeux de l'automatisation des techniques traditionnelles.

La méthode permet tout d'abord de reproduire le rendu des animations réalisées avec les techniques traditionnelles. Un risque résidait dans la possibilité que la réutilisation des coups de pinceau des images clés rende perceptible le mouvement des marques sur le support. Grâce à notre algorithme de composition et de sélection des marques, le déplacement des marques n'est pas visible sur le support. Cela nous permet d'atteindre un niveau de continuité temporelle de la stylisation similaire à celui des techniques traditionnelles. Ainsi, chaque image donne l'impression d'avoir été repeinte à la main, un objectif que nous avons atteint.

L'efficacité de la méthode est démontrée dans le style peinture. L'heuristique de sélection des marques, basée sur la quantité de matière du médium, peut être étendue aux stylisations fonctionnant avec cette notion comme par exemple les stylisations utilisant le graphite ou le pastel. Dans le cas contraire, l'approche générale peut être adaptée à d'autres stylisations n'utilisant pas les quantités de matière en modifiant l'algorithme d'évaluation des marques.

L'un des objectifs était de rester fidèle aux techniques traditionnelles. Pour y répondre, notre méthode propose deux flux de travail intégrant les techniques *pose à pose* et *sous la caméra*. L'animation produite peut ensuite être peaufinée en insérant une ou plusieurs images clés, générées puis éventuellement modifiées manuellement par l'artiste. Cependant pour être utilisée, la méthode requiert l'utilisation d'un mouvement. L'artiste peut dessiner et extrapoler le mouvement avec notre méthode, mais cela reste limité à la création de mouvements simples. Pour les mouvements plus complexes, l'artiste devra s'éloigner des techniques traditionnelles et recourir à la 3D ou à la vidéo.

La méthode ne peut pas être utilisée en continu par un artiste en raison de ses performances actuelles. Cependant elle peut être intégrée dans un processus hors ligne, où les images clés sont préparées le jour pour générer les images intermédiaires la nuit. Les perspectives laissent entrevoir une amélioration des performances. Le gain de temps pour un artiste utilisant notre méthode dans sa forme actuelle et dans un contexte de production est à nuancer.

D'un point de vue scientifique, nous proposons la première méthode à base de marques réalisant une interpolation entre deux images clés et une reproduction du style donné en exemple avec une approche de remplissage des formes. La méthode reproduit les détails à l'intérieur des formes grâce à la réutilisation des marques des images clés et assure une quantité de matière cohérente sur le support. Notre méthode enrichit l'état de l'art en matière de génération d'animations 2D à base de marques dans un style peinture où les techniques traditionnelles étaient jusqu'ici exclusivement utilisées pour les styles du dessin au crayon.

Au cours de ces travaux, plusieurs outils ont été développés pour étayer les nouveaux horizons qu'ouvre la méthode de génération d'animations traditionnelles par ordinateur. La méthode proposée dans ces travaux offre aux artistes la flexibilité des techniques traditionnelles de *pose à pose* et *sous la caméra* pour générer des animations 2D avec une apparence et la sensation des animations réalisées à la main. Cette approche est compatible avec les flux de travail d'animation en 2D pure et mixte avec de la 3D. Nous démontrons que le mouvement peut être dessiné et extrapolé afin de l'utiliser dans notre génération, et que la combinaison de scènes 3D peut améliorer la précision du mouvement et les possibilités de création. Ces travaux de thèse abordent une nouvelle approche de l'interpolation d'images qui prend en compte les contraintes de remplissage. Nous explorons le domaine et levons les problématiques à travers le médium de la peinture. Cela représente une avancée dans l'adaptation du style de l'animation traditionnelle à l'animation générée par ordinateur et ouvre de nouvelles possibilités pour les animateurs.

Chapitre 6

Annexes

6.1 Vue d'ensemble du système

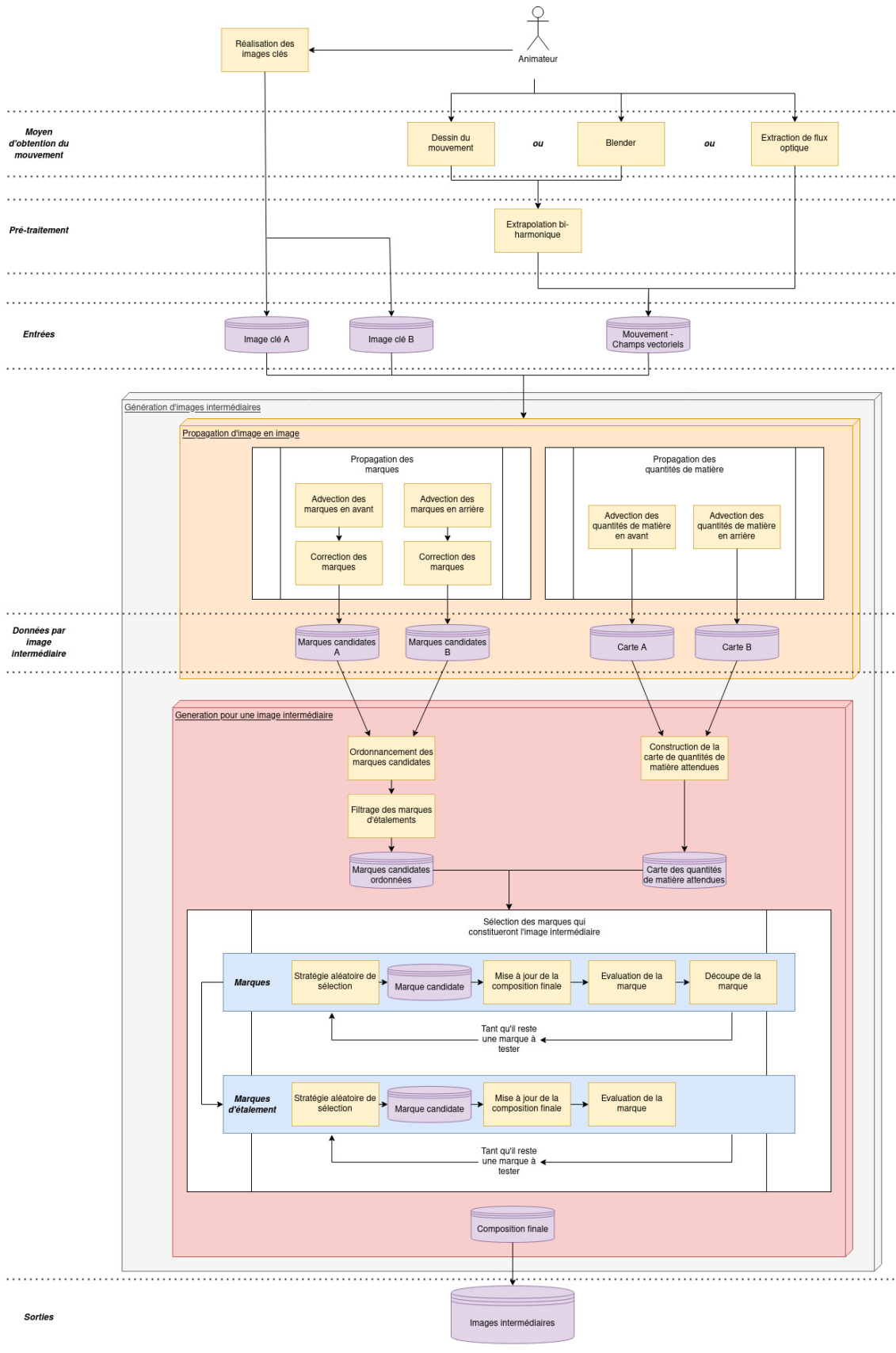


FIGURE 6.I – Pipeline de la méthode

Bibliographie

- [1] Ryoichi ANDO et Reiji TSURUNO. “Segmental Brush Synthesis with Stroke Images”. In : *Eurographics 2010 - Short Papers*. Sous la dir. de H. P. A. LENSCH et S. SEIPEL. The Eurographics Association, 2010. DOI : [10.2312/egsh.20101055](https://doi.org/10.2312/egsh.20101055) (cf. p. 12).
- [2] Michael ASHIKHMIN. “Synthesizing Natural Textures”. In : *Proceedings of the 2001 Symposium on Interactive 3D Graphics*. I3D '01. New York, NY, USA : Association for Computing Machinery, 2001, p. 217-226. ISBN : 1581132921. DOI : [10.1145/364338.364405](https://doi.org/10.1145/364338.364405) (cf. p. 15).
- [3] Connelly BARNES, Eli SHECHTMAN, Adam FINKELSTEIN et Dan B GOLDMAN. “PatchMatch : A Randomized Correspondence Algorithm for Structural Image Editing”. In : *ACM Trans. Graph.* 28.3 (juill. 2009). ISSN : 0730-0301. DOI : [10.1145/1531326.1531330](https://doi.org/10.1145/1531326.1531330) (cf. p. 31).
- [4] Katie BASSETT, Ilya BARAN, Johannes SCHMID, Markus GROSS et Robert W. SUMNER. “Authoring and Animating Painterly Characters”. In : *ACM Trans. Graph.* 32.5 (oct. 2013). ISSN : 0730-0301. DOI : [10.1145/2484238](https://doi.org/10.1145/2484238) (cf. p. 37).
- [5] Bill BAXTER, Vincent SCHEIB, Ming C. LIN et Dinesh MANOCHA. “DAB : Interactive Haptic Painting with 3D Virtual Brushes”. In : *Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques*. SIGGRAPH '01. New York, NY, USA : Association for Computing Machinery, 2001, p. 461-468. ISBN : 158113374X. DOI : [10.1145/383259.383313](https://doi.org/10.1145/383259.383313) (cf. p. 8).
- [6] W.V. BAXTER et M.C. LIN. “A versatile interactive 3D brush model”. In : *12th Pacific Conference on Computer Graphics and Applications, 2004. PG 2004. Proceedings*. 2004, p. 319-328. DOI : [10.1109/PCCGA.2004.1348363](https://doi.org/10.1109/PCCGA.2004.1348363) (cf. p. 8).
- [7] William BAXTER, Pascal BARLA et Ken ANJYO. “N-way morphing for 2D animation”. In : *Computer Animation and Virtual Worlds* 20.2-3 (2009), p. 79-87. DOI : [10.1002/cav.310](https://doi.org/10.1002/cav.310) (cf. p. 51, 52).
- [8] William BAXTER et Naga GOVINDARAJU. “Simple Data-Driven Modeling of Brushes”. In : *Proceedings of the 2010 ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games*. I3D '10. Washington, D.C. : Association for Computing Machinery, 2010, p. 135-142. ISBN : 9781605589398. DOI : [10.1145/1730804.1730826](https://doi.org/10.1145/1730804.1730826) (cf. p. 11, 12).

- [9] William BAXTER, Jeremy WENDT et Ming C. LIN. “IMPASTo : A Realistic, Interactive Model for Paint”. In : *Proceedings of the 3rd International Symposium on Non-Photorealistic Animation and Rendering*. NPAR '04. Annecy, France : Association for Computing Machinery, 2004, p. 45-148. ISBN : 1581138873. DOI : [10.1145/987657.987665](https://doi.org/10.1145/987657.987665) (cf. p. 10, 11).
- [10] S. BELONGIE, J. MALIK et J. PUZICHA. “Shape matching and object recognition using shape contexts”. In : *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24.4 (2002), p. 509-522. DOI : [10.1109/34.993558](https://doi.org/10.1109/34.993558) (cf. p. 29).
- [11] N. BEN-ZVI, J. BENTO, M. MAHLER, J. HODGINS et A. SHAMIR. “Line-Drawing Video Stylization”. In : *Computer Graphics Forum* 35.6 (2016), p. 18-32. DOI : <https://doi.org/10.1111/cgf.12729> (cf. p. 35).
- [12] Pierre BÉNARD, Adrien BOUSSEAU et Joëlle THOLLOT. “State-of-the-Art Report on Temporal Coherence for Stylized Animations”. In : *Computer Graphics Forum* 30.8 (2011), p. 2367-2386. DOI : [10.1111/j.1467-8659.2011.02075.x](https://doi.org/10.1111/j.1467-8659.2011.02075.x) (cf. p. 21, 30, 74).
- [13] Pierre BÉNARD, Forrester COLE, Michael KASS, Igor MORDATCH, James HEGARTY, Martin Sebastian SENN, Kurt FLEISCHER, Davide PESARE et Katherine BREEDEN. “Stylizing Animation by Example”. In : *ACM Trans. Graph.* 32.4 (juill. 2013). ISSN : 0730-0301. DOI : [10.1145/2461912.2461929](https://doi.org/10.1145/2461912.2461929) (cf. p. 27, 31, 32).
- [14] Adrien BOUSSEAU, Matt KAPLAN, Joëlle THOLLOT et François X. SILLION. “Interactive Watercolor Rendering with Temporal Coherence and Abstraction”. In : *Proceedings of the 4th International Symposium on Non-Photorealistic Animation and Rendering*. NPAR '06. Annecy, France : Association for Computing Machinery, 2006, p. 141-149. ISBN : 1595933573. DOI : [10.1145/1124728.1124751](https://doi.org/10.1145/1124728.1124751) (cf. p. 14).
- [15] Simon BRESLAV. “Towards Coherent Image Space Stylization of Animated 3D Shapes”. In : (2010) (cf. p. 22).
- [16] Robert BRIDSON, Jim HOURIHAM et Marcus NORDENSTAM. “Curl-Noise for Procedural Fluid Flow”. In : *ACM Trans. Graph.* 26.3 (juill. 2007), 46-es. ISSN : 0730-0301. DOI : [10.1145/1276377.1276435](https://doi.org/10.1145/1276377.1276435) (cf. p. 27).
- [17] Brian CABRAL et Leith Casey LEEDOM. “Imaging Vector Fields Using Line Integral Convolution”. In : *Proceedings of the 20th Annual Conference on Computer Graphics and Interactive Techniques*. SIGGRAPH '93. Anaheim, CA : Association for Computing Machinery, 1993, p. 263-270. ISBN : 0897916018. DOI : [10.1145/166117.166151](https://doi.org/10.1145/166117.166151) (cf. p. 53).
- [18] Jiazhou CHEN, Xinding ZHU, Pierre BÉNARD et Pascal BARLA. “Stroke Synthesis for Inbetweening of Rough Line Animations”. In : *Pacific Graphics Short Papers, Posters, and Work-in-Progress Papers*. Sous la dir. de Sung-hee LEE, Stefanie ZOLLMANN, Makoto OKABE et Burkhard WUENSCHÉ. The Eurographics Association, 2020. ISBN : 978-3-03868-120-5. DOI : [10.2312/pg.20201233](https://doi.org/10.2312/pg.20201233) (cf. p. 36, 40, 41).

- [19] N.S.-H. CHU et Chiew-Lan TAI. “An efficient brush model for physically-based 3D painting”. In : *10th Pacific Conference on Computer Graphics and Applications, 2002. Proceedings*. 2002, p. 413-421. DOI : [10.1109/PCCGA.2002.1167885](https://doi.org/10.1109/PCCGA.2002.1167885) (cf. p. 9).
- [20] N.S.H. CHU et Chiew-Lan TAI. “Real-time painting with an expressive virtual Chinese brush”. In : *IEEE Computer Graphics and Applications* 24.5 (2004), p. 76-85. DOI : [10.1109/MCG.2004.37](https://doi.org/10.1109/MCG.2004.37) (cf. p. 8).
- [21] Nelson CHU, William BAXTER, Li-Yi WEI et Naga GOVINDARAJU. “Detail-Preserving Paint Modeling for 3D Brushes”. In : *Proceedings of the 8th International Symposium on Non-Photorealistic Animation and Rendering*. NPAR '10. Annecy, France : Association for Computing Machinery, 2010, p. 27-34. ISBN : 9781450301251. DOI : [10.1145/1809939.1809943](https://doi.org/10.1145/1809939.1809943) (cf. p. 10, 43).
- [22] Nelson S.-H. CHU et Chiew-Lan TAI. “MoXi : Real-Time Ink Dispersion in Absorbent Paper”. In : *ACM Trans. Graph.* 24.3 (juill. 2005), p. 504-511. ISSN : 0730-0301. DOI : [10.1145/1073204.1073221](https://doi.org/10.1145/1073204.1073221) (cf. p. 10).
- [23] Malcolm Tunde COCKSHOT. “Wet and Sticky : a novel model for computer-based painting”. Thèse de doct. ProQuest Dissertations & Theses, 1991 (cf. p. 9).
- [24] Blender Online COMMUNITY. *Blender - a 3D modelling and rendering package*. Blender Foundation. Stichting Blender Foundation, Amsterdam, 2018. URL : <http://www.blender.org> (cf. p. 49).
- [25] Cassidy J. CURTIS, Sean E. ANDERSON, Joshua E. SEIMS, Kurt W. FLEISCHER et David H. SALESIN. “Computer-Generated Watercolor”. In : *Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques*. SIGGRAPH '97. USA : ACM Press/Addison-Wesley Publishing Co., 1997, p. 421-430. ISBN : 0897918967. DOI : [10.1145/258734.258896](https://doi.org/10.1145/258734.258896) (cf. p. 10, 11).
- [26] Boris DALSTEIN, Rémi RONFARD et Michiel van de PANNE. “Vector Graphics Animation with Time-Varying Topology”. In : *ACM Trans. Graph.* 34.4 (juill. 2015). ISSN : 0730-0301. DOI : [10.1145/2766913](https://doi.org/10.1145/2766913) (cf. p. 82, 83).
- [27] Doug DECARLO, Adam FINKELSTEIN, Szymon RUSINKIEWICZ et Anthony SANTELLA. “Suggestive Contours for Conveying Shape”. In : *ACM Trans. Graph.* 22.3 (juill. 2003), p. 848-855. ISSN : 0730-0301. DOI : [10.1145/882262.882354](https://doi.org/10.1145/882262.882354) (cf. p. 16).
- [28] Johanna DELANOY, Adrien BOUSSEAU et Aaron HERTZMANN. “Video Motion Stylization by 2D Rigidification”. In : *Proceedings of the 8th ACM/Eurographics Expressive Symposium on Computational Aesthetics and Sketch Based Interfaces and Modeling and Non-Photorealistic Animation and Rendering*. Expressive '19. Genoa, Italy : Eurographics Association, 2019, p. 11-19. DOI : [10.2312/exp.20191072](https://doi.org/10.2312/exp.20191072) (cf. p. 30).
- [29] Stephen DIVERDI, Aravind KRISHNASWAMY et Sunil HADAP. “Industrial-Strength Painting with a Virtual Bristle Brush”. In : *Proceedings of the 17th ACM Symposium on Virtual Reality Software and Technology*. VRST '10. Hong Kong : Association for Computing Machinery, 2010, p. 119-126. ISBN : 9781450304412. DOI : [10.1145/1889863.1889889](https://doi.org/10.1145/1889863.1889889) (cf. p. 8, 9).

- [30] David H DOUGLAS et Thomas K PEUCKER. “Algorithms for the reduction of the number of points required to represent a digitized line or its caricature”. In : *Cartographica : the international journal for geographic information and geovisualization* 10.2 (1973), p. 112-122. DOI : [10.3138/FM57-6770-U75U-7727](https://doi.org/10.3138/FM57-6770-U75U-7727) (cf. p. 43).
- [31] Even ENTEM, Amal Dev PARAKKAT, Marie-Paule CANI et Loïc BARTHE. “Structuring and Layering Contour Drawings of Organic Shapes”. In : *Proceedings of the Joint Symposium on Computational Aesthetics and Sketch-Based Interfaces and Modeling and Non-Photorealistic Animation and Rendering*. Expressive '18. Victoria, British Columbia, Canada : Association for Computing Machinery, 2018. ISBN : 9781450358927. DOI : [10.1145/3229147.3229155](https://doi.org/10.1145/3229147.3229155) (cf. p. 36).
- [32] Melvin EVEN, Pierre BÉNARD et Pascal BARLA. “Non-linear Rough 2D Animation using Transient Embeddings”. In : *Computer Graphics Forum* 42.2 (2023), p. 411-425. DOI : [10.1111/cgf.14771](https://doi.org/10.1111/cgf.14771) (cf. p. 34, 36).
- [33] Gunnar FARNEBÄCK. “Two-Frame Motion Estimation Based on Polynomial Expansion”. In : *Image Analysis*. Sous la dir. de Josef BIGUN et Tomas GUSTAVSSON. Berlin, Heidelberg : Springer Berlin Heidelberg, 2003, p. 363-370. ISBN : 978-3-540-45103-7. DOI : [10.1007/3-540-45103-X_50](https://doi.org/10.1007/3-540-45103-X_50) (cf. p. 50).
- [34] J. FIŠER, M. LUKÁČ, O. JAMRIŠKA, M. ČADÍK, Y. GINGOLD, P. ASENTE et D. SÝKORA. “Color Me Noisy : Example-based Rendering of Hand-colored Animations with Temporal Noise Control”. In : *Computer Graphics Forum* 33.4 (2014), p. 1-10. DOI : [10.1111/cgf.12407](https://doi.org/10.1111/cgf.12407) (cf. p. 32).
- [35] Jakub FIŠER, Ondřej JAMRIŠKA, Michal LUKÁČ, Eli SHECHTMAN, Paul ASENTE, Jingwan LU et Daniel SÝKORA. “StyLit : Illumination-Guided Example-Based Stylization of 3D Renderings”. In : *ACM Trans. Graph.* 35.4 (juill. 2016). ISSN : 0730-0301. DOI : [10.1145/2897824.2925948](https://doi.org/10.1145/2897824.2925948) (cf. p. 15, 31, 32).
- [36] Jakub FIŠER, Ondřej JAMRIŠKA, David SIMONS, Eli SHECHTMAN, Jingwan LU, Paul ASENTE, Michal LUKÁČ et Daniel SÝKORA. “Example-Based Synthesis of Stylized Facial Animations”. In : *ACM Trans. Graph.* 36.4 (juill. 2017). ISSN : 0730-0301. DOI : [10.1145/3072959.3073660](https://doi.org/10.1145/3072959.3073660) (cf. p. 32).
- [37] Leon A GATYS, Alexander S ECKER et Matthias BETHGE. “A neural algorithm of artistic style”. In : *arXiv preprint arXiv:1508.06576* (2015). DOI : [10.48550/arXiv.1508.06576](https://doi.org/10.48550/arXiv.1508.06576) (cf. p. 15).
- [38] Leon A. GATYS, Alexander S. ECKER et Matthias BETHGE. “Image Style Transfer Using Convolutional Neural Networks”. In : *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2016, p. 2414-2423. DOI : [10.1109/CVPR.2016.265](https://doi.org/10.1109/CVPR.2016.265) (cf. p. 32).
- [39] Leon A. GATYS, Alexander S. ECKER, Matthias BETHGE, Aaron HERTZMANN et Eli SHECHTMAN. “Controlling Perceptual Factors in Neural Style Transfer”. In : *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Juill. 2017. DOI : [10.48550/arXiv.1611.07865](https://doi.org/10.48550/arXiv.1611.07865) (cf. p. 33).

- [40] Amy GOOCH, Bruce GOOCH, Peter SHIRLEY et Elaine COHEN. “A Non-Photorealistic Lighting Model for Automatic Technical Illustration”. In : *Proceedings of the 25th Annual Conference on Computer Graphics and Interactive Techniques*. SIGGRAPH '98. New York, NY, USA : Association for Computing Machinery, 1998, p. 447-452. ISBN : 0897919998. DOI : [10.1145/280814.280950](https://doi.org/10.1145/280814.280950) (cf. p. 16).
- [41] Bruce GOOCH, Greg COOMBE et Peter SHIRLEY. “Artistic Vision : Painterly Rendering Using Computer Vision Techniques”. In : *Proceedings of the 2nd International Symposium on Non-Photorealistic Animation and Rendering*. NPAR '02. Annecy, France : Association for Computing Machinery, 2002, 83-ff. ISBN : 1581134940. DOI : [10.1145/508530.508545](https://doi.org/10.1145/508530.508545) (cf. p. 38).
- [42] Gaël GUENNEBAUD, Benoît JACOB et al. *Eigen* v3. 2010. URL : <http://eigen.tuxfamily.org> (cf. p. 73).
- [43] Chet S. HAASE et Gary W. MEYER. “Modeling Pigmented Materials for Realistic Image Synthesis”. In : *ACM Trans. Graph.* 11.4 (oct. 1992), p. 305-335. ISSN : 0730-0301. DOI : [10.1145/146443.146452](https://doi.org/10.1145/146443.146452) (cf. p. 10).
- [44] Paul HAEBERLI. “Paint by Numbers : Abstract Image Representations”. In : *Proceedings of the 17th Annual Conference on Computer Graphics and Interactive Techniques*. SIGGRAPH '90. Dallas, TX, USA : Association for Computing Machinery, 1990, p. 207-214. ISBN : 0897913442. DOI : [10.1145/97879.97902](https://doi.org/10.1145/97879.97902) (cf. p. 13, 34, 35).
- [45] R. HASHIMOTO, H. JOHAN et T. NISHITA. “Creating various styles of animations using example-based filtering”. In : *Proceedings Computer Graphics International 2003*. 2003, p. 312-317. DOI : [10.1109/CGI.2003.1214488](https://doi.org/10.1109/CGI.2003.1214488) (cf. p. 31).
- [46] James HAYS et Irfan ESSA. “Image and Video Based Painterly Animation”. In : *Proceedings of the 3rd International Symposium on Non-Photorealistic Animation and Rendering*. NPAR '04. Annecy, France : Association for Computing Machinery, 2004, p. 113-120. ISBN : 1581138873. DOI : [10.1145/987657.987676](https://doi.org/10.1145/987657.987676) (cf. p. 13, 34, 38).
- [47] Aaron HERTZMANN. “Painterly Rendering with Curved Brush Strokes of Multiple Sizes”. In : *Proceedings of the 25th Annual Conference on Computer Graphics and Interactive Techniques*. SIGGRAPH '98. New York, NY, USA : Association for Computing Machinery, 1998, p. 453-460. ISBN : 0897919998. DOI : [10.1145/280814.280951](https://doi.org/10.1145/280814.280951) (cf. p. 13, 34, 35).
- [48] Aaron HERTZMANN, Charles E. JACOBS, Nuria OLIVER, Brian CURLESS et David H. SALESIN. “Image Analogies”. In : *Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques*. SIGGRAPH '01. New York, NY, USA : Association for Computing Machinery, 2001, p. 327-340. ISBN : 158113374X. DOI : [10.1145/383259.383295](https://doi.org/10.1145/383259.383295) (cf. p. 15, 31).
- [49] Aaron HERTZMANN et Ken PERLIN. “Painterly Rendering for Video and Interaction”. In : *Proceedings of the 1st International Symposium on Non-Photorealistic Animation and Rendering*. NPAR '00. Annecy, France : Association for Computing Machinery, 2000, p. 7-12. ISBN : 1581132778. DOI : [10.1145/340916.340917](https://doi.org/10.1145/340916.340917) (cf. p. 34).

- [50] Zhongyuan HU, Haoran XIE, Tsukasa FUKUSATO, Takahiro SATO et Takeo IGARASHI. “Sketch2VF : Sketch-based flow design with conditional generative adversarial network”. In : *Computer Animation and Virtual Worlds* 30.3-4 (2019). e1889 cav.1889, e1889. DOI : <https://doi.org/10.1002/cav.1889> (cf. p. 27).
- [51] Zhewei HUANG, Wen HENG et Shuchang ZHOU. “Learning to Paint With Model-Based Deep Reinforcement Learning”. In : *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*. Oct. 2019. DOI : [10.48550/arXiv.1903.04411](https://doi.org/10.48550/arXiv.1903.04411) (cf. p. 35).
- [52] Matis HUDON, Mairéad GROGAN, Rafael PAGÉS, Jan ONDŘEJ et Aljoša SMOLIĆ. “2DToonShade : A stroke based toon shading system”. In : *Computers & Graphics : X 1* (2019), p. 100003. ISSN : 2590-1486. DOI : [10.1016/j.cagx.2019.100003](https://doi.org/10.1016/j.cagx.2019.100003) (cf. p. 29).
- [53] Ondřej JAMRIŠKA, Šárka SOCHOROVÁ, Ondřej TEXLER, Michal LUKÁČ, Jakub FIŠER, Jingwan LU, Eli SHECHTMAN et Daniel SÝKORA. “Stylizing Video by Example”. In : *ACM Trans. Graph.* 38.4 (juill. 2019). ISSN : 0730-0301. DOI : [10.1145/3306346.3323006](https://doi.org/10.1145/3306346.3323006) (cf. p. 32).
- [54] Jie JIANG, Hock Soon SEAH et Hong Ze LIEW. “Stroke-Based Drawing and Inbetweening with Boundary Strokes”. In : *Computer Graphics Forum* 41.1 (2022), p. 257-269. DOI : [10.1111/cgf.14433](https://doi.org/10.1111/cgf.14433) (cf. p. 22, 36).
- [55] Y. JING, Y. YANG, Z. FENG, J. YE, Y. YU et M. SONG. “Neural Style Transfer : A Review”. In : *IEEE Transactions on Visualization & Computer Graphics* 26.11 (nov. 2020), p. 3365-3385. ISSN : 1941-0506. DOI : [10.1109/TVCG.2019.2921336](https://doi.org/10.1109/TVCG.2019.2921336) (cf. p. 33).
- [56] Yoshihiro KANAMORI. “A comparative study of region matching based on shape descriptors for coloring hand-drawn animation”. In : *2013 28th International Conference on Image and Vision Computing New Zealand (IVCNZ 2013)*. 2013, p. 483-488. DOI : [10.1109/IVCNZ.2013.6727062](https://doi.org/10.1109/IVCNZ.2013.6727062) (cf. p. 29).
- [57] Matthew KAPLAN et Elaine COHEN. “A Generative Model For Dynamic Canvas Motion”. In : *Computational Aesthetics in Graphics, Visualization and Imaging*. Sous la dir. de Laszlo NEUMANN, Mateu SBERT, Bruce GOOCH et Werner PURGATHOFER. The Eurographics Association, 2005. ISBN : 3-905673-27-4. DOI : [10.2312/COMPAESTH/COMPAESTH05/049-056](https://doi.org/10.2312/COMPAESTH/COMPAESTH05/049-056) (cf. p. 11).
- [58] Dorota KOBIELA et Hugh WELCHMAN. *Loving Vincent*. BreakThru Productions, Trademark Films. 2017 (cf. p. 74).
- [59] Jonathan KONIECZNY, John HECKMAN, Gary MEYER, Mark MANYEN, Marty RABENS et Clement SHIMIZU. “Automotive spray paint simulation”. In : *International Symposium on Visual Computing*. Springer. 2008, p. 998-1007. DOI : [10.1007/978-3-540-89639-5_95](https://doi.org/10.1007/978-3-540-89639-5_95) (cf. p. 9).
- [60] Jonathan KONIECZNY et Gary MEYER. “Airbrush Simulation for Artwork and Computer Modeling”. In : *Proceedings of the 7th International Symposium on Non-Photorealistic Animation and Rendering*. NPAR '09. New Orleans, Louisiana : Association for Computing Machinery, 2009, p. 61-69. ISBN : 9781605586045. DOI : [10.1145/1572614.1572625](https://doi.org/10.1145/1572614.1572625) (cf. p. 9).

- [61] Alexander KORT. “Computer Aided Inbetweening”. In : *Proceedings of the 2nd International Symposium on Non-Photorealistic Animation and Rendering*. NPAR '02. Annecy, France : Association for Computing Machinery, 2002, p. 125-132. ISBN : 1581134940. DOI : [10.1145/508530.508552](https://doi.org/10.1145/508530.508552) (cf. p. 29).
- [62] J. LEE. “Simulating oriental black-ink painting”. In : *IEEE Computer Graphics and Applications* 19.3 (1999), p. 74-81. DOI : [10.1109/38.761553](https://doi.org/10.1109/38.761553) (cf. p. 9).
- [63] Sun-Young LEE, Jong-Chul YOON, Ji-Yong KWON et In-Kwon LEE. “CartoonModes : Cartoon Stylization of Video Objects through Modal Analysis”. In : *Graph. Models* 74.2 (mars 2012), p. 51-60. ISSN : 1524-0703. DOI : [10.1016/j.gmod.2012.02.001](https://doi.org/10.1016/j.gmod.2012.02.001) (cf. p. 29, 30).
- [64] Colin LEVY et Ton ROOSENDAAL. “Sintel”. In : *ACM SIGGRAPH ASIA 2010 Computer Animation Festival*. SA '10. Seoul, Republic of Korea : Association for Computing Machinery, 2010. ISBN : 9781450305259. DOI : [10.1145/1900264.1900346](https://doi.org/10.1145/1900264.1900346) (cf. p. 70).
- [65] John-Peter LEWIS. “Texture Synthesis for Digital Painting”. In : *SIGGRAPH Comput. Graph.* 18.3 (jan. 1984), p. 245-252. ISSN : 0097-8930. DOI : [10.1145/964965.808605](https://doi.org/10.1145/964965.808605) (cf. p. 11).
- [66] Chuan LI et Michael WAND. “Combining Markov Random Fields and Convolutional Neural Networks for Image Synthesis”. In : *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Juin 2016. DOI : [10.48550/arXiv.1601.04589](https://doi.org/10.48550/arXiv.1601.04589) (cf. p. 33).
- [67] Shaohua LI, Xinxing XU, Liqiang NIE et Tat-Seng CHUA. “Laplacian-steered neural style transfer”. In : *Proceedings of the 25th ACM international conference on Multimedia*. 2017, p. 1716-1724. DOI : [10.48550/arXiv.1707.01253](https://doi.org/10.48550/arXiv.1707.01253) (cf. p. 15).
- [68] Xiaoyu LI, Bo ZHANG, Jing LIAO et Pedro V. SANDER. “Deep Sketch-Guided Cartoon Video Inbetweening”. In : *IEEE Transactions on Visualization and Computer Graphics* 28.8 (2022), p. 2938-2952. DOI : [10.1109/TVCG.2021.3049419](https://doi.org/10.1109/TVCG.2021.3049419) (cf. p. 28, 33).
- [69] Peter LITWINOWICZ. “Processing Images and Video for an Impressionist Effect”. In : *Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques*. SIGGRAPH '97. USA : ACM Press/Addison-Wesley Publishing Co., 1997, p. 407-414. ISBN : 0897918967. DOI : [10.1145/258734.258893](https://doi.org/10.1145/258734.258893) (cf. p. 13, 34, 35).
- [70] Ce LIU, Antonio TORRALBA, William T. FREEMAN, Frédo DURAND et Edward H. ADELSON. “Motion Magnification”. In : *ACM Trans. Graph.* 24.3 (juill. 2005), p. 519-526. ISSN : 0730-0301. DOI : [10.1145/1073204.1073223](https://doi.org/10.1145/1073204.1073223) (cf. p. 30).
- [71] Dagmar LOFTSDOTTIR et Matthew GUZDIAL. “SketchBetween : Video-to-Video Synthesis for Sprite Animation via Sketches”. In : *Proceedings of the 17th International Conference on the Foundations of Digital Games*. FDG '22. Athens, Greece : Association for Computing Machinery, 2022. ISBN : 9781450397957. DOI : [10.1145/3555858.3555928](https://doi.org/10.1145/3555858.3555928) (cf. p. 33).

- [72] Jingwan LU, Fisher YU, Adam FINKELSTEIN et Stephen DIVERDI. “HelpingHand : Example-Based Stroke Stylization”. In : *ACM Trans. Graph.* 31.4 (juill. 2012). ISSN : 0730-0301. DOI : [10.1145/2185520.2185542](https://doi.org/10.1145/2185520.2185542) (cf. p. 15).
- [73] Donald W. MARQUARDT. “An Algorithm for Least-Squares Estimation of Nonlinear Parameters”. In : *Journal of the Society for Industrial and Applied Mathematics* 11.2 (1963), p. 431-441. DOI : [10.1137/0111030](https://doi.org/10.1137/0111030) (cf. p. 56).
- [74] D. MARR, E. HILDRETH et Sydney BRENNER. “Theory of edge detection”. In : *Proceedings of the Royal Society of London. Series B. Biological Sciences* 207.1167 (1980), p. 187-217. DOI : [10.1098/rspb.1980.0020](https://doi.org/10.1098/rspb.1980.0020) (cf. p. 30).
- [75] Barbara J. MEIER. “Painterly Rendering for Animation”. In : *Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques. SIGGRAPH '96*. New York, NY, USA : Association for Computing Machinery, 1996, p. 477-484. ISBN : 0897917464. DOI : [10.1145/237170.237288](https://doi.org/10.1145/237170.237288) (cf. p. 13, 37).
- [76] Nolan MESTRES, Romain VERGNE, Camille NOÛS et Joëlle THOLLOT. “Local Light Alignment for Multi-Scale Shape Depiction”. In : *Computer Graphics Forum* 40.2 (2021), p. 575-584. DOI : [10.1111/cgf.142656](https://doi.org/10.1111/cgf.142656) (cf. p. 16, 17).
- [77] Florence MIALHE et Marie DESPLECHIN. *La Traversée*. Gebeka Films. 2021 (cf. p. 74).
- [78] Charly MOURGLIA, Valentin ROUSSELLET, Loïc BARTHE, Nicolas MELLADO, Mathias PAULIN, David VANDERHAEGHE et al. *Radium-Engine*. Version doi-1.0.0. Juill. 2021. DOI : [10.5281/zenodo.5101334](https://doi.org/10.5281/zenodo.5101334). URL : <https://storm-irit.github.io/Radium-Engine/> (cf. p. 43).
- [79] Florian NOLTE, Andrew MELNIK et Helge RITTER. “Stroke-based Rendering : From Heuristics to Deep Learning”. In : *arXiv preprint arXiv:2302.00595* (2022). DOI : [10.48550/arXiv.2302.00595](https://doi.org/10.48550/arXiv.2302.00595) (cf. p. 35).
- [80] Peter O'DONOVAN et Aaron HERTZMANN. “Anipaint : Interactive painterly animation from video”. In : *IEEE transactions on visualization and computer graphics* 18.3 (2011), p. 475-487. DOI : [10.1109/TVCG.2011.51](https://doi.org/10.1109/TVCG.2011.51) (cf. p. 34, 35, 39, 40).
- [81] Bui Tuong PHONG. “Illumination for Computer Generated Pictures”. In : *Commun. ACM* 18.6 (juin 1975), p. 311-317. ISSN : 0001-0782. DOI : [10.1145/360825.360839](https://doi.org/10.1145/360825.360839) (cf. p. 16).
- [82] Lincoln RITTER, Wilmot LI, Brian CURLESS, Maneesh AGRAWALA et David SALESIN. “Painting With Texture”. In : *Symposium on Rendering*. Sous la dir. de Tomas AKENINE-MOELLER et Wolfgang HEIDRICH. The Eurographics Association, 2006. ISBN : 3-905673-35-5. DOI : [10.2312/EGWR/EGSR06/371-376](https://doi.org/10.2312/EGWR/EGSR06/371-376) (cf. p. 11).
- [83] Manuel RUDER, Alexey DOSOVITSKIY et Thomas BROX. “Artistic style transfer for videos”. In : *Pattern Recognition : 38th German Conference, GCPR 2016, Hannover, Germany, September 12-15, 2016, Proceedings* 38. Springer. 2016, p. 26-36. DOI : [10.1007/978-3-319-45886-1_3](https://doi.org/10.1007/978-3-319-45886-1_3) (cf. p. 33).

- [84] Manuel RUDER, Alexey DOSOVITSKIY et Thomas BROX. “Artistic style transfer for videos and spherical images”. In : *International Journal of Computer Vision* 126.11 (2018), p. 1199-1219. DOI : [10.1007/s11263-018-1089-z](https://doi.org/10.1007/s11263-018-1089-z) (cf. p. 33).
- [85] D. RUDOLF, D. MOULD et E. NEUFELD. “Simulating wax crayons”. In : *11th Pacific Conference on Computer Graphics and Applications, 2003. Proceedings.* 2003, p. 163-172. DOI : [10.1109/PCCGA.2003.1238258](https://doi.org/10.1109/PCCGA.2003.1238258) (cf. p. 11).
- [86] Dave RUDOLF, David MOULD et Eric NEUFELD. “A Bidirectional Deposition Model of Wax Crayons”. In : *Computer Graphics Forum* 24.1 (2005), p. 27-39. DOI : [10.1111/j.1467-8659.2005.00826.x](https://doi.org/10.1111/j.1467-8659.2005.00826.x) (cf. p. 10).
- [87] Dzhelil RUFAT et al. *Triangle*. 2020. URL : <https://rufat.be/triangle/> (cf. p. 51).
- [88] Stefano SAVIAN, Mehdi ELAHI et Tammam TILLO. “Optical flow estimation with deep learning, a survey on recent advances”. In : *Deep biometrics* (2020), p. 257-287. DOI : [10.1007/978-3-030-32583-1_12](https://doi.org/10.1007/978-3-030-32583-1_12) (cf. p. 28).
- [89] Johannes SCHMID, Martin Sebastian SENN, Markus GROSS et Robert W. SUMNER. “OverCoat : An Implicit Canvas for 3D Painting”. In : *ACM SIGGRAPH 2011 Papers. SIGGRAPH ’11*. Vancouver, British Columbia, Canada : Association for Computing Machinery, 2011. ISBN : 9781450309431. DOI : [10.1145/1964921.1964923](https://doi.org/10.1145/1964921.1964923) (cf. p. 13, 14, 38).
- [90] Jonathan Richard SHEWCHUK. “Triangle : Engineering a 2D Quality Mesh Generator and Delaunay Triangulator”. In : *Applied Computational Geometry : Towards Geometric Engineering*. Sous la dir. de Ming C. LIN et Dinesh MANOCHA. T. 1148. Lecture Notes in Computer Science. From the First ACM Workshop on Applied Computational Geometry. Springer-Verlag, mai 1996, p. 203-222. DOI : [10.1007/BFb0014497](https://doi.org/10.1007/BFb0014497) (cf. p. 51).
- [91] Michio SHIRAISHI et Yasushi YAMAGUCHI. “An Algorithm for Automatic Painterly Rendering Based on Local Source Image Approximation”. In : *Proceedings of the 1st International Symposium on Non-Photorealistic Animation and Rendering*. NPAR ’00. Annecy, France : Association for Computing Machinery, 2000, p. 53-58. ISBN : 1581132778. DOI : [10.1145/340916.340923](https://doi.org/10.1145/340916.340923) (cf. p. 13).
- [92] Francesco SIDDI, Matthew LUHN et Hjalti HJALMARSSON. “Sprite Fright”. In : *Proceedings of the SIGGRAPH Asia 2022 Computer Animation Festival*. SA ’22. Daegu, Republic of Korea : Association for Computing Machinery, 2023. ISBN : 9781450394642. DOI : [10.1145/3550339.3556000](https://doi.org/10.1145/3550339.3556000) (cf. p. 70).
- [93] Peter-Pike SLOAN, William MARTIN, Amy GOOCH et Bruce GOOCH. *The Lit Sphere : A Model for Capturing NPR Shading from Art*. Juin 2001. DOI : [10.20380/GI2001.17](https://doi.org/10.20380/GI2001.17) (cf. p. 31).
- [94] David SMALL. “Simulating watercolor by modeling diffusion, pigment, and paper fibers”. In : *Image Handling and Reproduction Systems Integration*. Sous la dir. de Walter R. BENDER et Wil PLOUFFE. T. 1460. International Society for Optics et Photonics. SPIE, 1991, p. 140-146. DOI : [10.1117/12.44417](https://doi.org/10.1117/12.44417) (cf. p. 10).

- [95] Olga SORKINE-HORNUNG et Michael RABINOVICH. *Least-squares rigid motion using SVD*. Rapp. tech. ETH Department of Computer Science : Zürich, Switzerland, 2017 (cf. p. 55).
- [96] Mario Costa SOUSA et John W. BUCHANAN. “Computer-Generated Graphite Pencil Rendering of 3D Polygonal Models”. In : *Computer Graphics Forum* 18.3 (1999), p. 195-208. DOI : [10.1111/1467-8659.00340](https://doi.org/10.1111/1467-8659.00340) (cf. p. 10).
- [97] Steve STRASSMANN. “Hairy Brushes”. In : *SIGGRAPH Comput. Graph.* 20.4 (août 1986), p. 225-232. ISSN : 0097-8930. DOI : [10.1145/15886.15911](https://doi.org/10.1145/15886.15911) (cf. p. 8, 9, 11).
- [98] Daniel SÝKORA, John DINGLIANA et Steven COLLINS. “As-Rigid-as-Possible Image Registration for Hand-Drawn Cartoon Animations”. In : *Proceedings of the 7th International Symposium on Non-Photorealistic Animation and Rendering*. NPAR’09. New Orleans, Louisiana : Association for Computing Machinery, 2009, p. 25-33. ISBN : 9781605586045. DOI : [10.1145/1572614.1572619](https://doi.org/10.1145/1572614.1572619) (cf. p. 29, 40).
- [99] THE GIMP DEVELOPMENT TEAM. *GIMP*. Version 2.10.12. 12 juin 2019. URL : <https://www.gimp.org> (cf. p. 18).
- [100] Hideki TODO, Kunihiko KOBAYASHI, Jin KATSURAGI, Haruna SHIMOTAHIRA, Shizuo KAJI et Yonghao YUE. “Stroke Transfer : Example-Based Synthesis of Animatable Stroke Styles”. In : *ACM SIGGRAPH 2022 Conference Proceedings*. SIGGRAPH ’22. Vancouver, BC, Canada : Association for Computing Machinery, 2022. ISBN : 9781450393379. DOI : [10.1145/3528233.3530703](https://doi.org/10.1145/3528233.3530703) (cf. p. 34, 37).
- [101] William VAN HAEVRE, Tom VAN LAERHOVEN, Fabian DI FIORE et Frank VAN REETH. “From dust till drawn”. In : *The Visual Computer* 23.9 (2007), p. 925-934. DOI : [10.1007/s00371-007-0144-5](https://doi.org/10.1007/s00371-007-0144-5) (cf. p. 10).
- [102] T. VAN LAERHOVEN, J. LIESENBORGS et F. VAN REETH. “Real-time watercolor painting on a distributed paper model”. In : *Proceedings Computer Graphics International*, 2004. 2004, p. 640-643. DOI : [10.1109/CGI.2004.1309281](https://doi.org/10.1109/CGI.2004.1309281) (cf. p. 10).
- [103] Tom VAN LAERHOVEN et Frank VAN REETH. “Brush up your painting skills”. In : *The Visual Computer* 23.9 (2007), p. 763-771. DOI : [10.1007/s00371-007-0158-z](https://doi.org/10.1007/s00371-007-0158-z) (cf. p. 8).
- [104] David VANDERHAEGHE, Pascal BARLA, Joelle THOLLOT et Francois X. SILLION. “Dynamic Point Distribution for Stroke-based Rendering”. In : *Rendering Techniques*. Sous la dir. de Jan KAUTZ et Sumanta PATTANAIK. The Eurographics Association, 2007. ISBN : 978-3-905673-52-4. DOI : [10.2312/EGWR/EGSR07/139-146](https://doi.org/10.2312/EGWR/EGSR07/139-146) (cf. p. 37, 38).
- [105] David VANDERHAEGHE et John COLLOMOSSE. “Stroke Based Painterly Rendering”. In : *Image and Video-Based Artistic Stylisation*. Sous la dir. de Paul ROSIN et John COLLOMOSSE. T. 42. Computational Imaging and Vision. Springer, London, 2012, p. 3-21. DOI : [10.1007/978-1-4471-4519-6_1](https://doi.org/10.1007/978-1-4471-4519-6_1) (cf. p. 12).

- [106] Ashish VASWANI, Noam SHAZEER, Niki PARMAR, Jakob USZKOREIT, Llion JONES, Aidan N GOMEZ, Łukasz KAISER et Illia POLOSUKHIN. “Attention is All you Need”. In : *Advances in Neural Information Processing Systems*. Sous la dir. d’I. GUYON, U. Von LUXBURG, S. BENGIO, H. WALLACH, R. FERGUS, S. VISHWANATHAN et R. GARNETT. T. 30. Curran Associates, Inc., 2017. DOI : [10.48550/arXiv.1706.03762](https://doi.org/10.48550/arXiv.1706.03762) (cf. p. 33).
- [107] Jue WANG, Steven M. DRUCKER, Maneesh AGRAWALA et Michael F. COHEN. “The Cartoon Animation Filter”. In : *ACM Trans. Graph.* 25.3 (juill. 2006), p. 1169-1173. ISSN : 0730-0301. DOI : [10.1145/1141911.1142010](https://doi.org/10.1145/1141911.1142010) (cf. p. 30).
- [108] Tinghuai WANG, John P. COLLOMOSSE, Andrew HUNTER et Darryl GREIG. “Learnable Stroke Models for Example-based Portrait Painting”. In : *British Machine Vision Conference, BMVC 2013, Bristol, UK, September 9-13, 2013*. Sous la dir. de Tilo BURGHARDT, Dima DAMEN, Walterio W. MAYOL-CUEVAS et Majid MIRMEHDI. BMVA Press, 2013. DOI : [10.5244/C.27.36](https://doi.org/10.5244/C.27.36) (cf. p. 15).
- [109] Brian WHITED, Gioacchino NORIS, Maryann SIMMONS, Robert W. SUMNER, Markus GROSS et Jarek ROSSIGNAC. “BetweenIT : An Interactive Tool for Tight Inbetweening”. In : *Computer Graphics Forum* 29.2 (2010), p. 605-614. DOI : [10.1111/j.1467-8659.2009.01630.x](https://doi.org/10.1111/j.1467-8659.2009.01630.x) (cf. p. 29, 36).
- [110] John WILLATS. *Art and representation : New principles in the analysis of pictures*. Princeton University Press, 1997 (cf. p. 7).
- [111] John WILLATS et Frédo DURAND. “Defining pictorial style : Lessons from linguistics and computer graphics”. In : *Axiomathes* 15.3 (2005), p. 319-351. DOI : [10.1007/s10516-004-5449-7](https://doi.org/10.1007/s10516-004-5449-7) (cf. p. 7).
- [112] Richard WILLIAMS. *The Animator’s Survival Kit–Revised Edition : A Manual of Methods, Principles and Formulas for Classical, Computer, Games, Stop Motion and Internet Animators*. Faber & Faber, Inc., 2009. ISBN : 0571238343 (cf. p. 20, 22, 23, 30).
- [113] Holger WINNEMÖLLER, Jan Eric KYPRIANIDIS et Sven C. OLSEN. “XDoG : An eXtended difference-of-Gaussians compendium including advanced image stylization”. In : *Computers & Graphics* 36.6 (2012). 2011 Joint Symposium on Computational Aesthetics (CAe), Non-Photorealistic Animation and Rendering (NPAR), and Sketch-Based Interfaces and Modeling (SBIM), p. 740-753. ISSN : 0097-8493. DOI : [10.1016/j.cag.2012.03.004](https://doi.org/10.1016/j.cag.2012.03.004) (cf. p. 14).
- [114] Helena T.F WONG et Horace H.S IP. “Virtual brush : a model-based synthesis of Chinese calligraphy”. In : *Computers & Graphics* 24.1 (2000), p. 99-113. ISSN : 0097-8493. DOI : [10.1016/S0097-8493\(99\)00141-7](https://doi.org/10.1016/S0097-8493(99)00141-7) (cf. p. 9).
- [115] Ming XIE. “Feature matching and affine transformation for 2D cell animation”. In : *The Visual Computer* 11 (1995), p. 419-428. DOI : [10.1007/BF02464332](https://doi.org/10.1007/BF02464332) (cf. p. 29).
- [116] Songhua XU, Haisheng TAN, Xiantao JIAO, Francis C.M. LAU et Yunhe PAN. “A Generic Pigment Model for Digital Painting”. In : *Computer Graphics Forum* 26.3 (2007), p. 609-618. DOI : [10.1111/j.1467-8659.2007.01084.x](https://doi.org/10.1111/j.1467-8659.2007.01084.x) (cf. p. 10, 11).

- [117] Songhua XU, Min TANG, Francis C. M. LAU et Yunhe PAN. “Virtual Hairy Brush for Painterly Rendering”. In : *Graph. Models* 66.5 (sept. 2004), p. 263-302. ISSN : 1524-0703. DOI : [10.1016/j.gmod.2004.05.006](https://doi.org/10.1016/j.gmod.2004.05.006) (cf. p. 9).
- [118] Lijie YANG, Tianchen XU, Jixiang DU, Hongbo ZHANG et Enhua WU. “Brushwork master : Chinese ink painting synthesis for animating brushwork process”. In : *Computer Animation and Virtual Worlds* 31.4-5 (2020), e1949. DOI : [10.1002/cav.1949](https://doi.org/10.1002/cav.1949) (cf. p. 38).
- [119] Wenwu YANG, Hock-Soon SEAH, Quan CHEN, Hong-Ze LIEW et Daniel SÝKORA. “FTP-SC : Fuzzy Topology Preserving Stroke Correspondence”. In : *Computer Graphics Forum* 37.8 (2018), p. 125-135. DOI : [10.1111/cgf.13518](https://doi.org/10.1111/cgf.13518) (cf. p. 36).
- [120] Mingliang ZHAI, Xuezhi XIANG, Ning LV et Xiangdong KONG. “Optical flow and scene flow estimation : A survey”. In : *Pattern Recognition* 114 (2021), p. 107861. ISSN : 0031-3203. DOI : [10.1016/j.patcog.2021.107861](https://doi.org/10.1016/j.patcog.2021.107861) (cf. p. 28, 50, 78).
- [121] Mingtian ZHAO et Song-Chun ZHU. “Portrait Painting Using Active Templates”. In : *Proceedings of the ACM SIGGRAPH/Eurographics Symposium on Non-Photorealistic Animation and Rendering*. NPAR '11. Vancouver, British Columbia, Canada : Association for Computing Machinery, 2011, p. 117-124. ISBN : 9781450309073. DOI : [10.1145/2024676.2024696](https://doi.org/10.1145/2024676.2024696) (cf. p. 15).
- [122] Zhengxia ZOU, Tianyang SHI, Shuang QIU, Yi YUAN et Zhenwei SHI. “Stylized Neural Painting”. In : *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Juin 2021, p. 15689-15698. DOI : [10.48550/arXiv.2011.08114](https://doi.org/10.48550/arXiv.2011.08114) (cf. p. 35).

Titre : Génération d'images intermédiaires pour la création d'animations 2D stylisées à base de marques

Mots clés : informatique graphique, rendu expressif, méthodes basées sur des exemples, animation peinture, rendu à base de marques

Résumé : Dans le cadre de ma thèse, je m'intéresse à la problématique de la création d'animations 2D traditionnelles, où toutes les images sont réalisées à la main. Je cherche à explorer comment l'ordinateur peut aider les artistes à produire de manière plus efficace sans pour autant empiéter sur le spectre de la création artistique. Pour répondre à ce problème, mes travaux se placent dans le domaine des méthodes automatiques, où l'animateur travaille de manière itérative avec l'ordinateur.

Je propose une méthode qui, à partir de deux images clés et d'une série de champs vectoriels 2D décrivant le mouvement dans l'espace image de l'animation, génère des images intermédiaires et reproduit le style donné en exemple. Ma méthode se situe à l'intersection de deux techniques manuelles d'animation : l'animation pose à pose et l'animation sous la caméra, et assure un contrôle fort en permettant d'éditer n'importe quelle image générée de la même manière que celles données en exemple.

Mes travaux englobent plusieurs domaines, notamment l'analyse du mouvement, le contrôle de courbe 2D, le rendu à base de marques et la simulation de la peinture.

Title: Inbetweens generation for 2D stylized stroke-based animation

Key words: computer graphics, expressive rendering, example-based methods, painterly animation, stroke-based rendering

Abstract: As part of my thesis, I am interested in the issue of creating traditional 2D animations, where all the images are handcrafted. Specifically, I explore how computers can assist artists in producing animations efficiently without reducing the artistic creative process.

To address this problem, my work falls within the scope of automatic methods, where the animator collaborates iteratively with the computer. I propose a method that takes two keyframe images and a series of 2D vector fields describing the motion in image space of the animation, and generates intermediate images while preserving the given style as an example.

My method combines two manual animation techniques: pose-to-pose and frame-by-frame animation, providing strong control by allowing any generated image to be edited in the same way as the example images provided.

My research covers several domains: motion analysis, 2D curve control, mark-based rendering, and paint simulation.