

Université de Montréal

Rééclairage par génération de cartes de normales
en composition vidéo numérique

par

Nicolas Bergeron

Département d'informatique et de recherche opérationnelle
Faculté des arts et des sciences

Mémoire présenté à la faculté des études supérieures
en vue de l'obtention du grade de
Maître ès sciences (M.Sc.)
en informatique

avril 2006

© Nicolas Bergeron, 2006

Université de Montréal
Faculté des études supérieures

Ce mémoire de maîtrise intitulé

Rééclairage par génération de cartes de normales
en composition vidéo numérique

présenté par
Nicolas Bergeron

a été évalué par un jury composé des personnes suivantes :

Yann-Gaël Guéhéneuc
président-rapporteur

Pierre Poulin
directeur de recherche

Victor Ostromoukhov
membre du jury

Résumé

Dans les productions cinématographiques, les effets d'éclairage sont une grande préoccupation. En composition vidéo numérique, lorsque des séquences sont combinées pour créer des scènes inédites, les différences d'éclairage entre ces séquences sont problématiques et trahissent l'effet de réalisme voulu. Traditionnellement, des outils corrigent les couleurs afin de remettre les séquences dans un même espace lumineux. Basés sur une mise en correspondance des couleurs, ces outils sont limités par le manque d'informations à propos de la géométrie et des fonctions de réflectance de la scène, ce qui donne rarement des résultats convaincants lorsque les éclairages des séquences d'origine sont incompatibles. Récemment, l'extraction de certaines informations manquantes a été rendue possible grâce à des systèmes complexes de caméras et de projecteurs lors du tournage, permettant de simuler un éclairage arbitraire par rendu. Puisque les coûts associés à ces technologies sont présentement trop élevés pour l'industrie, nous proposons un outil intermédiaire qui cadre bien dans le contexte traditionnel de post-production. Notre outil de génération de cartes de normales permet de retrouver automatiquement et en temps réel une géométrie courbe "plausible" des objets d'avant-plan filmés devant l'écran bleu, en exploitant la silhouette du matte et le dégradé d'illumination sur ces objets. Afin d'améliorer le résultat final, une interface usager permet d'ajuster des paramètres simples : une pondération entre les deux stratégies de reconstruction ainsi que des transformations sur les séquences originales et les cartes de normales. Après avoir placé des lumières virtuelles dans un espace tridimensionnel, les cartes de normales sont directement utilisées par le matériel graphique pour refaire le rendu. Dans une composition, notre technique améliore l'effet d'immersion entre les séquences.

Mots clefs :

Composition vidéo numérique, rééclairage vidéo, carte de normales, reconstruction "plausible", reconstruction par dégradé d'illumination, reconstruction à partir d'un matte.

Abstract

Lighting effects are a great concern in cinematographic productions. In digital video compositing, when many sequences are combined to create new scenes, lighting discrepancies are problematic and betray the desired realistic effect of a composite. Traditionally, color correction tools are used to match the lighting space of sequences. Based on color matching, these tools are limited by the lack of information about the objects geometry and reflectance functions, achieving unconvincing results when the original sequences are incompatible. Recently, extracting some of the missing information became possible by using complex camera and projector systems during the filming, allowing to render afterwards the sequence with arbitrary lighting. However, the cost of this technology is currently too high for the industry and we propose an intermediate tool which is consistent with the current post-production pipeline. Our semi-automatic normal map generation tool enables the reconstruction of a “plausible” curved geometry of the foreground objects while using bluescreen footage. Shading and silhouettes of the matte are used for our real-time reconstruction. To improve the final result, a user interface allows to adjust simple parameters : a weighting factor between the two reconstruction schemes and transformations on the original sequences and normal maps. Thereafter, once virtual lights are set in a 3D space, normal maps are directly used by the graphics hardware to render with the new illumination. In a composite, our technique improves the immersion effect between the sequences.

Keywords :

Digital video compositing, video relighting, normal mapping, “plausible” reconstruction, shape from shading, shape from silhouette.

Table des matières

Remerciements	xii
Introduction	1
Composition vidéo numérique	2
Contributions	3
1 Post-production et composition vidéo	5
1.1 Histoire de la composition optique d’images	6
1.2 “Pipeline” actuel de post-production	9
1.2.1 Transfert au numérique	9
1.2.2 Imagerie générée par ordinateur	10
1.3 Composition d’images	12
1.3.1 Dégrainage et dépoussiérage	13
1.3.2 Opérateurs de composition	13
1.3.3 <i>Matting</i>	14
1.3.4 Correction de couleur	16
1.4 Matriçage vidéo	17
2 Rééclairage par correction de couleur	18
2.1 Représentation de la couleur	18
2.1.1 Longueur d’onde	19
2.1.2 Température	19
2.1.3 Espaces de couleur	20
2.2 Correction de couleur	22
2.2.1 Luminance et contraste	22
2.2.2 Ajustement des couleurs	26

2.2.3	Ajout d'ombres	28
2.2.4	Autres effets	28
3	Rééclairage par rendu	30
3.1	Rendu d'images	31
3.2	Sondes d'illumination	32
3.2.1	Les yeux comme sonde d'illumination	33
3.3	Modélisation d'objets réels	34
3.3.1	Modélisation à partir d'images	35
3.3.2	Systèmes interactifs	36
3.3.3	Numériseurs 3D	37
3.3.4	Lumière structurée	37
3.3.5	Combinaison de techniques	37
4	Au-delà de la correction de couleur	40
4.1	Un environnement tridimensionnel	41
4.2	Génération de cartes de normales	42
4.2.1	Surface à partir du matte	43
4.2.2	Surface par dégradé d'illumination	50
4.2.3	Éditer les cartes de normales	51
4.3	Rééclairage	51
4.4	Intégration dans un contexte de composition vidéo numérique	53
5	Résultats	54
5.1	Extraction de cartes de normales	55
5.1.1	Reconstruction automatique	55
5.1.2	Paramètres de reconstruction à partir des silhouettes	55
5.1.3	Paramètres de reconstruction à partir du dégradé d'illumination	56
5.1.4	Combinaison	57
5.2	Rééclairage virtuel	58
5.3	Problèmes de robustesse	59
5.3.1	Reconstruction à partir du matte	60
5.3.2	Reconstruction à partir du dégradé d'illumination	63

5.4 Analyse de la performance	63
Conclusion	66
Travaux futurs	67
Améliorations	67
Extensions	68
Bibliographie	70

Table des figures

1	Composition vidéo numérique tirée de [WP03]	2
1.1	Technique de l'écran bleu en composition numérique tirée de [SJTS04] .	6
1.2	Imagerie générée par ordinateur dans une composition tirée de [NM03] .	11
1.3	Logiciel de composition vidéo numérique (<i>Discreet Flame/Inferno</i>) [ME06]	12
2.1	Diagramme de chromaticité de la CIE tiré de [Owe99]	20
2.2	Égalisation de l'histogramme tirée de [Mig05]	24
2.3	L'addition et la multiplication par une constante	24
2.4	Valeurs de référence des blancs et des noirs entre deux images	25
2.5	Amélioration du contraste et correction gamma	26
2.6	Technique des fausses ombres	28
3.1	Sondes d'illumination tirée de [Deb06]	32
3.2	Les yeux comme sonde d'illumination [NN04]	34
3.3	Modélisation et rendu à partir de séquences vidéo	35
3.4	Images successives prise par le système de <i>performance relighting</i> dans l'espace d'un douzième de seconde, tiré de [WGT ⁺ 05]	38
3.5	Applications du système de <i>performance relighting</i> , tiré de [WGT ⁺ 05] : (a-d) photographies originales et informations retrouvées par le système, (e-h) photographies rééclairées par une carte de radiance en modifiant le ratio diffus / spéculaire, (i-l) autres effets pouvant être réalisés par le système	39
4.1	Système de modélisation <i>Teddy</i> tiré de [IMT99]. Gauche : forme dessinée. Centre et droite : forme 3D affichée avec un style "cartoon"	44

4.2	Reconstruction par la méthode de Oh <i>et al.</i> [OCDD01] par rapport à la nôtre	45
4.3	Illustration en 2D des techniques de génération de hauteurs à partir du matte	46
4.4	Optimisations par <i>quad-tree</i> et par calcul à basse résolution	49
4.5	Illustration de notre système de génération de cartes de normales	52
5.1	Reconstruction automatique : (a) par silhouettes, (b) par dégradé d'illumination, (c) combinaison 50% : 50%	56
5.2	Reconstruction à partir des silhouettes : (a) mise à l'échelle (5), (b) dilatation du matte de 5 pixels, (c) dilatation du matte de 25 pixels	57
5.3	Reconstruction à partir du dégradé d'illumination : (a) mise à l'échelle (0.5), (b) mise à l'échelle (0.05), (c) filtrage gaussien de rayon 5 sur (a), (d) filtrage gaussien de rayon 25 sur (a)	58
5.4	Combinaison des cartes de normales (ratio à <i>partir des silhouettes</i> : à <i>partir du dégradé d'illumination</i>) : (a) 25% : 75%, (b) 50% : 50%, (c) 75% : 25%	59
5.5	Rééclairage avec une source de lumière virtuelle positionnée selon quatre directions différentes	60
5.6	Composition exploitant notre outil : (a) composition avec rééclairage, (b) comparaison avec et sans éclairage	61
5.7	Incohérence temporelle due à l'occlusion : (a, b) sans occlusion, (c, d) occlusion	62
5.8	Différents problèmes de robustesse de notre outil : (a) problèmes près des bordures, (b) problèmes d'ombres (rouge) et de transitions de couleur (bleu), (c et d) images originales	65

Liste des tableaux

1.1	Formats d'images lors du transfert au numérique [Bri99, Wri01]	11
2.1	Correction de couleur avec un vert constant	27
5.1	Temps d'exécution (en secondes) en fonction de la résolution de l'image	64

Remerciements

Je suis fier de remettre ce mémoire, mettant ainsi un terme à un projet qui a occupé une place importante dans ma vie. Je n'aurais pas pu y parvenir sans Pierre Poulin, mon directeur de recherche, qui m'a toujours convaincu de persévérer dans ce projet, même dans les moments les plus difficiles. J'admire son dévouement pour les gens, sa passion contagieuse pour l'infographie et son sens de l'humour.

Merci aux gens de chez *Autodesk Media & Entertainment* qui m'ont offert un stage et un projet exceptionnel pour ma première expérience dans l'industrie. Je me considère privilégié d'avoir été encadré par des gens aussi qualifiés que Jean-François Dumais et Martin Hélie. Merci aussi à Pierre Bouchard qui a appuyé le projet.

Il ne faut pas oublier tous les membres du *LIGUM* avec lesquels j'ai partagé ces trois dernières années. Je voudrais remercier particulièrement Philippe Beaudoin pour ses idées par rapport au projet et François Duranleau pour sa patience et son expertise. Pour les discussions enflammées, les tournois de jeux vidéo, les présentations le midi, les parties de walleyball et les participations aux conférences, je conserve de merveilleux souvenirs de mon séjour au lab.

Merci aux enseignants du *DIRO* de l'Université de Montréal pour les connaissances qu'ils ont su me transmettre et merci au *Conseil de Recherche en Sciences Naturelles et Génie* pour son soutien financier.

Les encouragements et le soutien constant de ma famille et de mes proches m'ont vraiment aidé à rester sur Terre durant les dernières années. Merci à mes parents Josée et Michel, à ma petite soeur Jacinthe, à mes grands-parents Dolorès et Léon, et à ma belle-mère Claire. Je dédie ce mémoire à Valentine, la femme de ma vie.

Introduction

Le développement effréné de l’informatique dans les dernières décennies a drastiquement changé notre mode de vie. L’infographie a fait exploser le domaine du divertissement en introduisant les jeux vidéo et les images de synthèse. Le monde du cinéma n’a pas échappé aux changements. La pellicule utilisée traditionnellement pour l’acquisition, la post-production et la distribution de films laisse graduellement sa place aux outils numériques pour représenter le contenu vidéo.

Même si le processus de production vidéo demeure stable depuis des années, le “pipeline” de post-production d’effets visuels est en constante évolution. De nouveaux outils émergent constamment des avancées technologiques dans les domaines de traitement d’images, de vision par ordinateur, d’animation physique et de synthèse d’images. La convergence de ces domaines en édition vidéo représente actuellement un défi pour les concepteurs d’outils puisque les algorithmes de chaque domaine requièrent des données différentes. Traditionnellement en composition vidéo numérique, seule la couleur des pixels des images était utilisée. Maintenant, les nouveaux algorithmes exigent des données additionnelles.

L’estimation du flux optique, la simulation de fluides et l’insertion d’objets synthétiques dans les photos réelles sont des techniques en plein développement qui permettent de créer des effets spéciaux dans les films. Pour les utiliser, la calibration de la caméra¹, la connaissance de la géométrie de la scène, de sa réflectance et/ou la position des sources de lumière doivent être fournies ou extraites à partir d’images.

Certaines productions cinématographiques à grand déploiement ont innové dans le domaine en introduisant des outils non traditionnels tels que des tableaux de caméras calibrées entre elles pour extraire automatiquement la géométrie de la scène, ainsi que des sondes d’illumination pour extraire une carte d’environnement de la scène. Cepen-

¹le calcul de sa matrice de projection

dant, les producteurs cherchent généralement à réduire la préparation et les interventions lors des tournages déjà coûteux et sont plutôt conservateurs quant aux nouvelles technologies. Ainsi, lorsque les séquences sont acquises sans ces technologies, un travail supplémentaire est laissé à la post-production pour réaliser les mêmes effets.

Composition vidéo numérique

Les outils mentionnés ci-haut sont utilisés pour la composition vidéo numérique (voir le chapitre 1) qui consiste à combiner des séquences d'images provenant de sources différentes. La figure 1 illustre la combinaison de séquences réelles et synthétiques. Le plus grand défi dans ce domaine est de convaincre le public que ce montage de séquences d'images provient d'une seule prise de caméra. Déjà une grande préoccupation en production, les effets d'éclairage viennent souvent trahir le réalisme d'une composition. En effet, en plus de la couleur de l'éclairage, les indices sur la direction des lumières telles que les ombres et les reflets spéculaires sont particulièrement problématiques.



FIG. 1 – Composition vidéo numérique tirée de [WP03]

Pour uniformiser l'éclairage des différents éléments combinés d'une composition, des algorithmes de correction de couleur sont couramment utilisés (voir le chapitre 2). La couleur des images est transformée grâce à une carte de correspondances. Pour établir ces correspondances, une bonne compréhension de la perception de la couleur est primordiale. Aussi, les différentes représentations de la couleur peuvent aider à ajuster le contraste, la luminance, la teinte et la saturation entre l'image d'avant et d'arrière-plan. L'ajout d'ombres et de brouillard améliorent aussi le réalisme. Cependant, les opérations

de correction de couleur, réalisées globalement sur toute l'image, sont limitées. Puisque les problèmes de directionnalité de la lumière ne sont pas traités, certaines séquences sont difficiles à uniformiser.

Des techniques plus modernes commencent à voir le jour pour rééclairer les images et faire le rendu d'objets synthétiques en apportant des modifications à la production (voir le chapitre 3). Le but est de simuler un nouvel éclairage en exploitant la géométrie et les fonctions de réflectance des surfaces extraites de la scène. Pour extraire ces informations, on a recours à plusieurs caméras, des cartes d'environnement extraites à partir de sondes d'illumination et des projections de lumière dans la scène. En combinant ces techniques, il est possible de refaire un rendu photo-réaliste de la scène d'un nouveau point de vue sous un éclairage arbitraire, en y ajoutant des objets synthétiques.

Contributions

Au moment de la post-production, lorsque les séquences proviennent d'un tournage traditionnel et que de nouveaux effets d'éclairage au-delà de la correction de couleur doivent être réalisés, l'artiste doit modéliser les objets de la scène en mouvement, un processus fastidieux. Pour accélérer et faciliter cette tâche, nous avons développé un outil interactif de reconstruction des objets, conçu spécialement pour la composition vidéo numérique, pour des objets courbes filmés devant un écran bleu.

La reconstruction est réalisée en exploitant deux informations présentes sur les séquences d'écran bleu : la forme des silhouettes à partir du matte² et le dégradé d'illumination sur les objets. La reconstruction 3D à partir d'une image étant un problème mal posé³, notre outil ne permet pas de modéliser précisément les objets, mais aide plutôt l'artiste qui connaît la scène à lui inférer une forme "plausible". La reconstruction à partir du matte permet de déduire une forme lisse à l'intérieur des silhouettes de l'objet. La reconstruction à partir du dégradé d'illumination est plus flexible et produit plus de variations à l'intérieur des surfaces.

La géométrie des objets est représentée par une carte de normales animée dans le temps. Des paramètres intuitifs permettent à l'artiste d'éditer et de raffiner le résultat obtenu automatiquement grâce à notre outil. Une pondération et un facteur d'échelle

²carte de transparence

³les données sont insuffisantes pour retrouver la solution exacte

sont déterminés pour ajuster l'importance de la reconstruction à partir des silhouettes par rapport à celle à partir du dégradé d'illumination. Des paramètres de filtrage et des opérations sur le matte sont aussi disponibles.

Une fois les cartes de normales générées, les objets de la composition sont disposés dans un environnement tridimensionnel avant que des lumières virtuelles ne soient ajoutées. Une fois appliquées sur les objets, les cartes de normales, combinées à l'éclairage virtuel, produisent un effet de profondeur à la scène.

L'outil a été implémenté dans le produit *Discreet Flame/Inferno* [ME06], un logiciel commercial de composition vidéo numérique et d'effets visuels. Des optimisations à nos algorithmes procurent une meilleure interactivité à l'artiste qui peut visualiser les effets des paramètres simultanément sur la carte de normales et sur l'image rééclairée en temps réel (voir le chapitre 4).

Des résultats de composition obtenus à partir de scènes où l'éclairage est incompatible (voir le chapitre 5) démontrent bien la flexibilité et l'interactivité de notre outil. La faible puissance de calcul requise permet de créer des effets de rééclairage à faible coût qui vont au-delà de la correction de couleur. Le tout facilite l'intégration de séquences où les effets d'éclairages sont problématiques.

Chapitre 1

Post-production et composition vidéo

Les effets spéciaux existent depuis les débuts du cinéma. Ces techniques appliquées pendant ou après le tournage ont pour effet de modifier la scène originale, de permettre la création de scènes inexistantes ou impossibles, de réduire les coûts et de générer des ambiances. Nous pouvons classer ces effets en plusieurs catégories : maquettes et modèles réduits, illusions d'optique, trucages au tournage et imagerie générée par ordinateur (animation 2D, 3D et physique).

L'imagerie générée par ordinateur est à la base des effets spéciaux que l'on retrouve dans les films d'aujourd'hui. Le fait de modéliser et d'animer de façon réaliste des créatures imaginaires dans des scènes inédites est tout à fait remarquable, mais la difficulté d'intégrer ces objets synthétiques dans les séquences vidéo réelles l'est encore plus. En fait, les effets spéciaux résultent du travail d'artistes oeuvrant dans une panoplie de domaines. De la pellicule de la caméra jusqu'à l'écran de cinéma, chaque image passe à travers le "pipeline" de post-production d'effets visuels.

La composition vidéo numérique se définit comme la création de séquences vidéo, en combinant des images provenant de diverses sources telles que des séquences numériques du monde réel, des films numérisés, des images de synthèse 3D, des animations 2D, des photos numériques et du texte (définition traduite de [Wik06]).

Le matte, une carte de transparence aussi appelée canal α , permet d'isoler les objets dans une image. Une composition consiste en la superposition de couches d'images semi-transparentes sur un arrière-plan opaque. Le défi est de donner l'illusion que ce collage,

à la manière d'un photomontage, provient d'une seule prise de caméra. L'équation de composition (voir l'équation 1.1) proposée par Porter et Duff [PD84] définit la technique où l'image C est générée à partir des images A et B et de leur matte respectif A_a et B_a , où les valeurs de A_a et B_a sont compris dans l'intervalle $[0, 1]$.

$$C = (A_a \times A) + (B_a \times B) \quad (1.1)$$

La technique de l'écran bleu¹, consistant à filmer des objets et des acteurs devant un écran de couleur uniforme (généralement bleu ou vert), facilite l'extraction des mattes et est couramment utilisée aujourd'hui. La figure 1.1 illustre le procédé. Dans les grandes productions, les films sont généralement tournés sur de la pellicule de 35mm, avant d'être numérisés pour l'édition et la composition vidéo. Une panoplie d'outils logiciels sont utilisés à cette fin. Dans ce chapitre, nous allons surtout nous concentrer sur les outils de base : soit le *matting* et les opérateurs de composition.

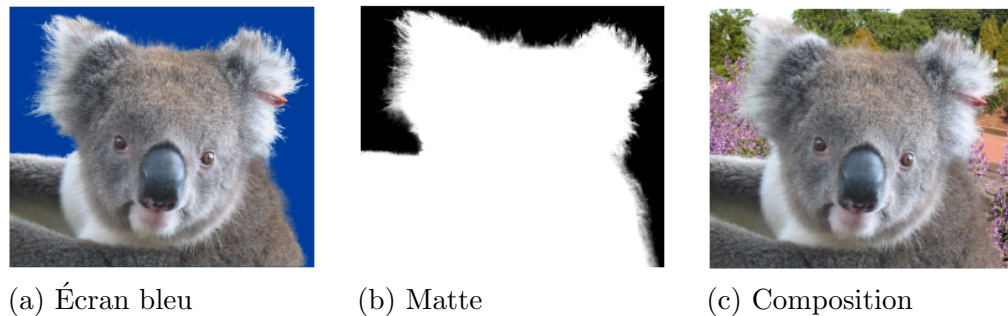


FIG. 1.1 – Technique de l'écran bleu en composition numérique tirée de [SJTS04]

1.1 Histoire de la composition optique d'images

Pour comprendre l'état actuel de la post-production vidéo, il est intéressant de connaître les fondements du domaine. L'histoire de la composition numérique a débuté par l'invention de la caméra, du cinéma et des effets spéciaux. Une grande partie de cette section a été tirée du livre de Rickitt [Ric00] et de sites web [Ana06, Inv06].

Au seizième siècle, Léonard de Vinci a étudié pour la première fois le principe de la chambre noire². Une boîte complètement noire est percée d'un minuscule trou laissant

¹*bluescreen*

²*camera obscura*

entrer la lumière. L'image renversée d'un objet ou d'un paysage est projetée sur un écran blanc. Le problème de l'impression de l'image, qui fait appel à des notions de chimie, a été résolu avec l'invention de la pellicule par William Henry Fox Talbot en 1835 et la photographie est née.

L'inventeur Thomas Edison est à l'origine du cinéma tel qu'on le connaît aujourd'hui grâce à l'invention du kinétographe et du kinétoscope en 1891. Le premier appareil permettait de capturer des vues³ d'une durée comprise entre 30 et 40 secondes au total, à la fréquence de 12 images par tour de manivelle, imprégnées sur une pellicule de 35mm. Le second appareil, soit une boîte dans laquelle les images défilaient les unes à la suite des autres, servait à visionner les vues.

Quelques années plus tard, en 1895, les frères Lumière ont inventé le cinématographe, un appareil permettant de capturer et projeter les vues. La première projection publique a eu lieu à Paris le 25 décembre 1895, où une dizaine de bandes d'une durée d'environ une minute chacune ont été présentées.

Ayant assisté à la première projection, le magicien prestidigitateur Georges Méliès a fait transformé un projecteur en caméra pour réaliser les premiers trucages du cinéma. Le blocage de la caméra lors d'un tournage a créé un effet inattendu, le personnage filmé a passé instantanément d'un endroit à un autre. Appelé "arrêt de caméra", ce trucage qui produit des apparitions, disparitions et substitutions instantanées a été largement utilisé dans les films de Méliès. Pour préserver le rythme en utilisant cet effet, le recours au montage des pellicules en laboratoire a été nécessaire. Ainsi, pour éviter les incongruités temporelles, on filmait un peu plus que nécessaire, puis on raboutait les deux prises à l'endroit idéal.

Celui que l'on surnomme le père des effets spéciaux a aussi inventé la surimpression. La technique nécessite un rembobinage de la pellicule pour une seconde impression. Le trucage permet ainsi d'avoir un même personnage simultanément à plusieurs endroits dans l'image. Les tournages successifs se faisaient généralement devant un décor noir, laissant ainsi la pellicule intacte. La différence d'intensité entre les impressions était cependant le principal problème lié à cette technique.

Inspiré par la surimpression, Edwin Porter en 1903 a apporté une innovation importante en introduisant le *matte*⁴. Initialement utilisée pour ajouter l'arrivée d'un train

³le terme *film* n'existait pas à l'époque

⁴*matte shot*

dans la fenêtre d'une gare, la technique consistait à masquer une partie de la lentille de la caméra (fenêtre de la gare) lors de la première impression, puis à masquer la partie inverse lors de la seconde (arrivée du train). Ainsi, la composition d'images a été rendue possible sans exposer la pellicule plusieurs fois au même endroit.

Un peu plus tard, en 1907, Norman O. Dawn a eu l'idée de peindre des objets de la scène sur une feuille de verre⁵. En plaçant la feuille entre la caméra et les personnages, on avait l'impression que les objets faisaient partie du décor. La technique a plus tard été utilisée pour peindre les mattes et combiner plusieurs images.

Max Fleisher, en 1914, a inventé le rotoscope qui permet de peindre les mattes à la main. Cette technique permet de combiner des images en mouvement grâce à des mattes dynamiques. La qualité obtenue est variable et dépend évidemment de l'artiste.

En 1916, Frank Williams a inventé un procédé pour extraire le matte d'un personnage en mouvement⁶ automatiquement. L'astuce est de placer le personnage devant un écran de couleur uniforme, et d'utiliser une pellicule à fort contraste pour extraire le matte. Il était cependant difficile d'obtenir un contraste suffisant pour obtenir un matte de qualité. Un peu plus tard, en 1927, C. Dodge Dunning a raffiné la technique en utilisant une pellicule qui répond seulement à la couleur orange. L'arrière-plan devait premièrement être filmé, puis la pellicule devait être développée, blanchie et teinte de couleur orange. Cette pellicule positive orange et blanche était alors placée dans une caméra spéciale et déroulait en même temps que la pellicule du tournage. Les objets de l'avant-plan, placés devant un écran bleu étaient éclairés par une lumière orange qui passait à travers les deux pellicules. Une fois remise en noir et blanc, l'illusion était parfaite.

Une autre technique a vu le jour en 1927 quand Eugene Schufftan a eu l'idée d'utiliser un miroir pour combiner plusieurs images. Le procédé Schufftan consiste à placer un miroir à 45 degrés devant la caméra, et de gratter la peinture réfléchissante du miroir à certains endroits. L'action se déroule à l'endroit où le miroir est transparent, et l'autre partie de l'image (généralement des éléments du décor) est réfléchi par le miroir dans la caméra.

En 1928, Linwood G. Dunn a fabriqué une imprimante optique. Ce type d'appareil permettait l'écriture de titres et de génériques, des effets de transition entre plu-

⁵*glass shot*

⁶*traveling shot*

sieurs images et la composition d'images. L'évolution de la technologie a permis le développement de l'imprimante optique à quatre têtes qui sert à combiner quatre pellicules. Ainsi, à partir de deux images, puis d'un matte mâle et femelle, il est possible d'imprimer l'image composée sur une nouvelle pellicule.

En 1933, Willis O'Brien a créé la technique de projection arrière qui consiste à filmer les acteurs devant un écran sur lequel est projeté l'arrière-plan. Cette technique a largement été utilisée dans l'histoire du cinéma, notamment dans les scènes de voiture. La faiblesse de la technique est la différence d'intensité lumineuse entre l'avant et l'arrière-plan.

La technologie de l'écran bleu a véritablement pris son envol vers la fin des années 1950 quand Petro Vlahos a construit un système utilisant des lampes à vapeur de sodium pour éclairer l'écran. Les personnages sont quant à eux éclairés par une lumière blanche. À l'aide d'un prisme placé derrière la lentille qui sépare les deux types de lumière, le matte et l'image d'avant-plan sont simultanément imprimés sur pellicule.

Le procédé ZOptic, inventé par Zoran Peresic en 1977 a été l'une des dernières innovations avant le début de l'ère numérique et consiste en un système de projection arrière où les lentilles optiques du projecteur et de la caméra sont synchronisées. Ainsi, lorsque le plan de la caméra change, le projecteur est automatiquement ajusté.

1.2 “Pipeline” actuel de post-production

Actuellement, tous les projets cinématographiques doivent passer par l'étape de post-production. Cette étape, aussi importante que celle du tournage, ne consiste pas seulement à remettre les scènes en ordre chronologique, mais aussi à transformer les images de façon à les rendre plus attrayantes. Ainsi, les images sont numérisées et corrigées, de l'imagerie générée par ordinateur est ajoutée et, finalement, le tout est converti dans le format voulu. Tout ce qui concerne l'audio est traité séparément de la vidéo et est ajouté quand la post-production d'effets visuels est complétée.

1.2.1 Transfert au numérique

La première étape de post-production consiste à transférer les images de la pellicule à un format numérique. À cette fin, des numériseurs à haute précision sont utilisés pour préserver le plus fidèlement possible les couleurs captées sur la pellicule. Le processus

consiste à faire dérouler le film devant un tableau linéaire de capteurs optiques (*CCD*⁷) qui transforment la couleur du film en une valeur numérique. Sous forme numérique, une image est représentée comme étant un tableau de valeurs où une cellule appelée *pixel* est généralement décomposée en plusieurs canaux. La dimension du tableau est définie par la résolution de l'image.

Les formats d'images les plus courants sont *NTSC*⁸, *PAL*⁹ et *HDTV*¹⁰ pour la télévision, et *Academy*, *Full Aperture*, *Cinamscope*, *Vista Vision*, *70mm* et *Imax* pour le cinéma. Dans les débuts du transfert au numérique, on a établi qu'un transfert sans perte pour une pellicule de 35mm pouvait être réalisé à une résolution de 4K (4096) pixels de largeur. Actuellement, une résolution de 2K (2048) pixels de largeur est jugée suffisante, ce qui réduit considérablement le coût de stockage et de traitement. La qualité est cependant supérieure lorsque la pellicule est numérisée en 4K et reconvertie à 2K. Généralement, une résolution de 2K est utilisée pour les projets en 35mm et une de 4K pour les projets en 70mm. Le tableau 1.1 présente les résolutions standards pour les différents formats d'images ainsi que leur rapport d'aspect¹¹. Ces informations permettent de déduire que les pixels ne sont pas toujours carrés.

La résolution de la couleur, soit la quantité de mémoire allouée pour une couleur, a un effet notoire sur la qualité de l'image. La représentation la plus courante de la couleur est celle qu'utilisent les moniteurs : trois canaux *RGB*¹² de 8 bits par canal. D'autres formats peuvent aussi être utilisés (voir la section 2.1.3). Lors du transfert, pour gagner de la qualité, les couleurs sont généralement extraites à haute résolution (12 ou 16 bits par canal), pour ensuite être converties à plus basse résolution. Les images à haut diapason dynamique (*HDR*¹³) commencent à apparaître sur le marché et leur encodage peut atteindre 32 bits par canal.

1.2.2 Imagerie générée par ordinateur

Plusieurs productions intègrent de l'imagerie générée par ordinateur dans les images réelles. Cette discipline très vaste fait appel à plusieurs types d'expertises tels que des

⁷ *Charged Coupled Device*

⁸ *National Television Systems Committee*

⁹ *Phase Alternating Line*

¹⁰ *High-Definition Television*

¹¹ ratio largeur/hauteur

¹² *Red, Green, Blue*

¹³ *High Dynamic Range*

Format (Télévision)	Résolution	Rapport d'aspect
NTSC	720×486	1.33
PAL	720×576	1.33
HD 720-line	1280×720	1.85
HD 1080-line	1920×1080	1.85

Format (Cinéma)	Demie résolution	Pleine résolution	Rapport d'aspect
Academy (35mm)	1828×1332	3656×2664	1.37
Full Aperture (35mm)	2048×1556	4096×3112	1.33
Cinemascope (35mm)	1828×1556	3656×3112	2.35
Vista Vision (35mm)	3078×2048	6144×4096	1.5
70mm	2048×920	4096×1840	2.2
Imax (70mm)	2808×2048	5616×4096	1.37

TAB. 1.1 – Formats d'images lors du transfert au numérique [Bri99, Wri01]

dessinateurs, modeleurs, animateurs, spécialistes de capture de mouvement, artistes du rendu, scripteurs, etc.

Avec les avancées technologiques en animation, en rendu et en simulation physique, il est maintenant moins coûteux de simuler par ordinateur des créatures fantastiques et certains phénomènes naturels que de les créer avec des matériaux réels. Par exemple, les liquides, fluides et explosions générés par ordinateur peuvent être introduits dans les films de façon réaliste. La figure 1.2 illustre la technique.

Les logiciels les plus communs et complets pour générer de l'imagerie sont *Alias Maya* [Ali06], *Softimage XSI* [Sof06] et *Discreet 3D Studio MAX* [ME06].

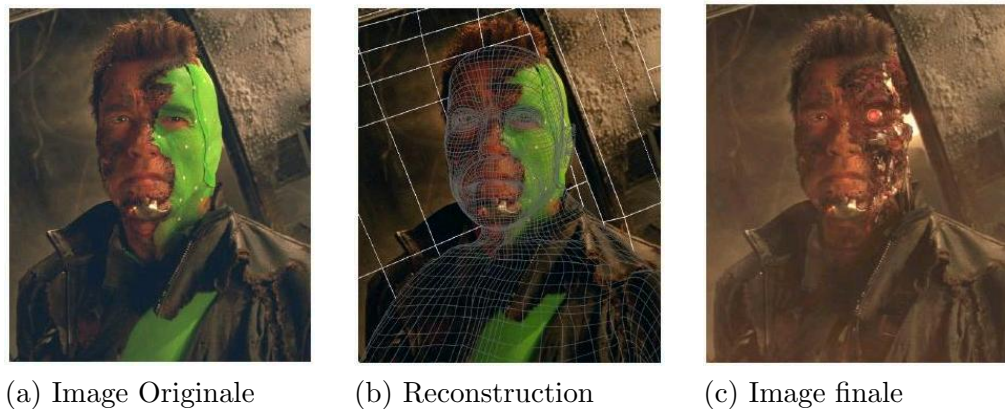


FIG. 1.2 – Imagerie générée par ordinateur dans une composition tirée de [NM03]

1.3 Composition d'images

La composition d'images numériques est un trucage qui consiste à combiner des images numériques de plusieurs sources. Pour y parvenir, dans les logiciels disponibles, chaque outil prend en entrée une (ou plusieurs) image(s) et génère une nouvelle image.

Certains logiciels commerciaux, tel que *Discreet Flame/Inferno* (voir la figure 1.3) [ME06], offrent un ensemble complet d'outils de composition d'images, pouvant être disposés dans un graphe de flux¹⁴. Ainsi, chaque noeud dans le graphe est un outil, et les arcs orientés servent à relier les outils entre eux. On peut alors suivre le chemin à partir des images sources jusqu'aux images finales.

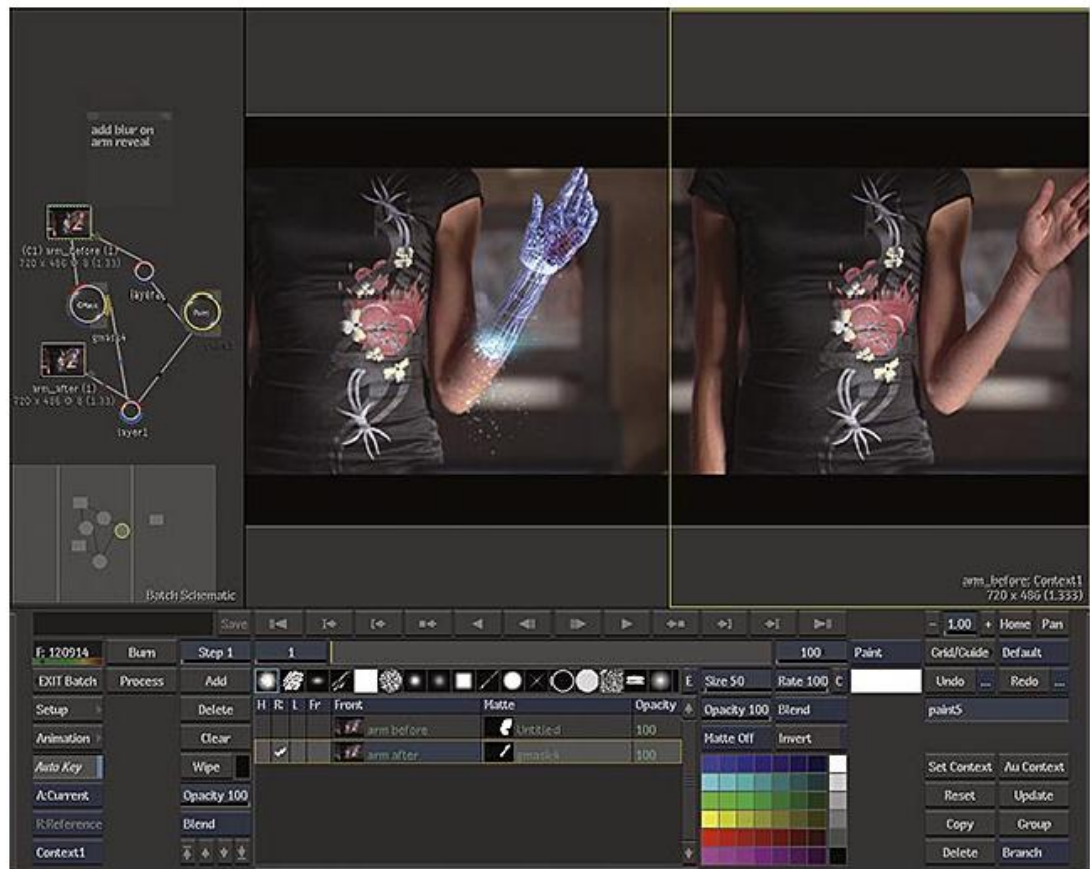


FIG. 1.3 – Logiciel de composition vidéo numérique (*Discreet Flame/Inferno*) [ME06]

Les principaux outils sont des filtres d'images (dégraineurs, dépoussiéreurs, correcteurs de couleur, générateurs de flous), des générateurs de mattes et des opérateurs de composition.

¹⁴flowgraph

1.3.1 Dégrainage et dépoussiérage

Une fois les images numérisées, il faut corriger les défauts de l'image qui proviennent du film et du transfert au numérique. Le grain du film ou des poussières sur la pellicule créent des artéfacts. Les outils de dégrainage et dépoussiérage tentent d'exploiter les caractéristiques de ces artéfacts pour les enlever sans affecter les détails de l'image.

CinePaint [Cin06] est un logiciel libre utilisé professionnellement pour enlever ces types d'artéfacts. Il permet en plus d'appliquer plusieurs algorithmes de base en traitement d'images.

1.3.2 Opérateurs de composition

Les opérateurs de composition, proposés par Porter et Duff [PD84], génèrent une image (C) en combinant deux images (A et B) et leur matte respectif (A_a et B_a).

Au-dessus

L'opérateur de base en composition est le "au-dessus"¹⁵. Il consiste à découper la silhouette de l'image A puis à la coller au-dessus de l'image B . Mathématiquement, l'opération est décrite par l'équation 1.2.

$$C = A_a \times A + (1 - A_a) \times B \quad (1.2)$$

Addition et soustraction

Les opérateurs d'addition et soustraction ont pour effet d'additionner ou soustraire tous les pixels dans l'image A avec les pixels correspondants dans l'image B . Mathématiquement, les opérations sont décrites par les équations 1.3 et 1.4.

$$C = A + B \quad (1.3)$$

$$C = A - B \quad (1.4)$$

Il est important de noter qu'à cause de la discrétisation des pixels en valeurs entières et des valeurs limites numériques, le résultat de ces opérations peut dépasser ces limites.

¹⁵ *over*

Dans ces cas, la majorité des implémentations assigne la valeur à la limite. Aussi, certains opérateurs de différence donnent la valeur absolue entre les deux images.

Dissolution

L'opérateur de dissolution¹⁶ est une somme pondérée et normalisée de deux images. La pondération (p) de chaque image est déterminée par l'utilisateur. Mathématiquement, l'opération est décrite par l'équation 1.5.

$$C = p \times A + (1 - p) \times B \quad (1.5)$$

Dans les films où l'on retrouve des fantômes, c'est souvent cette opération qui est utilisée pour mélanger le fantôme transparent avec l'arrière-plan. On l'utilise aussi pour réaliser des fondus.

Intérieur et extérieur

Certains opérateurs nécessitent seulement un masque comme deuxième image. Les opérateurs d'intérieur et d'extérieur, aussi appelés d'union et d'intersection, ont pour résultat de conserver la région de l'image qui est à l'intérieur (équation 1.6) ou à l'extérieur (équation 1.7) de ce masque.

$$C = A \times B_a \quad (1.6)$$

$$C = A \times (1 - B_a) \quad (1.7)$$

Ces opérateurs ont plusieurs utilités, notamment le raffinement des masques en utilisant des masques de déchets (voir la section 1.3.3).

1.3.3 *Matting*

En composition vidéo, le *matting* consiste à extraire une séquence de masques à partir d'une séquence vidéo. Cette étape est cruciale puisqu'un masque de qualité est nécessaire à une bonne composition.

Des travaux réalisés par Smith et Blinn [SB96] ont démontré que le problème de *matting* est mal posé. En ajoutant une contrainte au problème, soit deux arrière-plans

¹⁶*mix*

de couleurs différentes, il est possible d'isoler automatiquement l'avant-plan de l'arrière-plan avec un système de triangulation.

En pratique, l'utilisation d'un seul écran bleu est la seule solution viable pour obtenir des mattes de qualité. La majorité des outils commerciaux de *matting* sont basés sur trois techniques : le *matting* par luminance, par chrominance et par différence¹⁷.

***Matting* par luminance**

La luminance peut être définie comme l'intensité d'une couleur. L'équation 1.8 est souvent utilisée pour calculer la luminance [Wri01]. Le poids de chaque canal est proportionnel à la sensibilité de l'oeil pour la couleur associée.

$$L = 0.29R + 0.59G + 0.12B \quad (1.8)$$

Le *matting* par luminance consiste à établir une plage de luminance qui correspond uniquement à l'arrière-plan. Pour y parvenir, l'utilisateur étudie l'histogramme (voir la section 2.2.1) de l'image puis détermine à l'aide d'une courbe, une valeur de transparence pour chaque valeur de luminance.

***Matting* par chrominance**

Le *matting* par chrominance utilise la couleur de l'arrière-plan pour déterminer sa transparence. Généralement, l'image *RGB* est convertie dans l'espace *HSV*¹⁸ (voir la section 2.1 pour plus de détails). Ensuite, l'utilisateur détermine la couleur de l'arrière-plan, puis des seuils de tolérance pour la couleur, la saturation et la teinte. À partir de ces critères, la transparence est déterminée pour chaque pixel.

***Matting* par différence**

Le *matting* par différence peut être utilisé lorsque l'arrière-plan de la scène sans l'objet est connu. L'opérateur de différence est alors utilisé entre l'image et l'image d'arrière-plan. L'utilisateur détermine alors une correspondance entre la différence et la transparence de chaque pixel en traçant une courbe sur l'histogramme. En général,

¹⁷*luma-key matting, chroma-key matting et difference matting*

¹⁸*Hue, Saturation, Value*

cette technique est très imprécise et est surtout utilisée pour extraire des mattes de déchets¹⁹ servant à raffiner les mattes bruitées.

Raffinement des mattes

La qualité d'une composition dépend en grande partie de la qualité des mattes. Les algorithmes automatiques de *matting* donnent rarement un résultat parfait. Une deuxième étape de raffinement, ou de nettoyage, consiste à réparer les mattes aux endroits où ces algorithmes ont failli à la tâche.

Une technique consiste à utiliser des mattes de déchets qui consistent généralement en des surfaces grossières dessinées par l'utilisateur. L'intérieur ou l'extérieur de ces surfaces sera nécessairement translucide ou opaque. Par exemple, lorsqu'on utilise un écran bleu et que l'on filme un acteur avec les yeux bleus, la partie du matte devant les yeux est souvent transparente. Les artistes dessinent alors une surface devant les yeux pour que cette région devienne opaque. Les opérateurs de composition d'intérieur et d'extérieur sont très utiles en combinaison avec les mattes de déchets.

Les opérateurs morphologiques [GW01] permettent de dilater ou d'éroder les mattes. Puisque les pixels près des bordures contiennent ou réfléchissent souvent une portion de l'arrière-plan, des problèmes apparaissent. Ces opérateurs sont très utilisés pour isoler ces phénomènes. En faisant la différence entre un matte dilaté puis un matte érodé, nous obtenons le matte des silhouettes. Les artistes travaillent parfois directement sur ce matte pour atténuer les problèmes près des bordures.

Des filtres sont aussi utilisés pour supprimer le bruit dans les images, causé entre autres par la caméra, le grain de la pellicule et le transfert au numérique.

1.3.4 Correction de couleur

L'éclairage est l'un des aspects les plus problématiques en composition d'images, puisque des scènes prises dans différents contextes d'éclairage doivent être combinées. Une attention particulière doit être portée au tournage pour s'assurer que l'ensemble des séquences à combiner aient des éclairages compatibles.

Différents outils de correction de couleur et d'intensité sont fournis par la majorité des logiciels de composition vidéo et d'effets visuels. Ils permettent de mieux uniformiser

¹⁹ *garbage matte* ou *garbage mask*

les séquences. Les algorithmes derrière ces outils sont expliqués en détails au chapitre 2.

1.4 Matriage vidéo

Une fois le film monté en format numérique, il faut le retransformer dans les formats souhaités, c'est-à-dire créer les bandes maîtresses dans un procédé qu'on appelle le mariage²⁰. Pour chacun des formats, les images doivent être modifiées pour que les couleurs apparaissent le plus fidèlement possible sur un téléviseur ou un écran de cinéma à celles sur le moniteur de l'artiste. Certains formats pour la télévision doivent être entrelacés, et un grain artificiel est parfois ajouté (opération inverse du dégrainage) pour donner une impression similaire aux films du cinéma.

Pour les projections au cinéma, les images doivent être enregistrées sur pellicule. À cette fin, il existe deux types d'enregistreurs : ceux à tubes à rayon cathodiques (*CRT*²¹) et ceux au laser. Les deux appareils impriment les images, un canal à la fois. Les appareils au laser sont plus dispendieux, mais plus rapides, plus précis et peuvent être utilisés sur de la pellicule à grain plus petit, nécessitant un plus long temps d'exposition.

²⁰mastering

²¹*Cathode Ray Tube*

Chapitre 2

Rééclairage par correction de couleur

L'un des problèmes les plus difficiles à résoudre en composition vidéo numérique est le camouflage des différences d'éclairage entre les éléments d'avant et d'arrière-plan. Par exemple, lorsque sont utilisées des images provenant d'un studio, de scènes filmées en plein air et de scènes générées par ordinateur, il est essentiel de corriger les couleurs des images de façon à ce que le montage apparaisse réaliste. À ce stade, corriger l'éclairage correspond à remettre les images dans un même espace lumineux. De plus, une correction spécifique au grain doit être préalablement réalisée (voir la section 1.3.1).

Le travail des coloristes, les artistes chargés de corriger la couleur des images, ne repose pas seulement sur une vision artistique mais aussi sur plusieurs notions théoriques à propos de la couleur, de la lumière et de son interaction avec les matériaux. Les différents modèles pour représenter la couleur et le comportement de la lumière dans l'environnement sont discutés ici avant d'étudier les différentes techniques traditionnelles de correction de couleur.

2.1 Représentation de la couleur

Le concept de couleur est intimement lié à notre perception visuelle de la lumière. En effet, la rétine de l'oeil humain est tapissée de cellules qui réagissent à la lumière. Les bâtonnets sont excités par l'intensité lumineuse, tandis que les cônes sont sensibles à la couleur.

La représentation de la couleur diffère d'un domaine à l'autre. Les physiciens préconisent un modèle ondulatoire, tandis que les cinéastes emploient le concept du rayonnement émis par un corps noir en fonction de la température. Les infographistes préfèrent quant à eux séparer la couleur en une coordonnée numérique dans un espace de couleur multi-dimensionnel.

2.1.1 Longueur d'onde

En physique optique, la lumière est décrite par son comportement corpusculaire et ondulatoire. Les photons, soient les particules de lumière, voyagent dans l'espace à partir des sources qui les émettent jusqu'aux surfaces qui les réfléchissent et les absorbent.

La lumière peut être analysée comme la superposition d'ondes sinusoïdales, chacune ayant une longueur différente. Ainsi, le terme couleur est employé lorsque la longueur d'onde se situe à l'intérieur du spectre visible, soit approximativement entre 380 nm et 740 nm. La lumière blanche renferme l'ensemble des couleurs du spectre visible.

Les trois types de cônes présents sur la rétine sont excités par différentes longueurs d'onde. Les cônes de type *S* sont sensibles à la couleur bleue (420 nm), les cônes de type *M* sont sensibles au vert (534 nm) et les cônes de type *L* sont sensibles au rouge (562 nm). Les cônes de type *M* sont présents en plus grande concentration sur la rétine et c'est pour cette raison que l'œil est plus sensible aux variations de vert. Les bâtonnets quant à eux sont indifférents à la longueur d'onde.

2.1.2 Température

La couleur de la lumière émise par une lampe ou un moniteur est souvent décrite en termes de température de couleur. Cette température n'est pas liée à la chaleur générée par les sources de lumière mais plutôt au concept de radiation d'un corps noir. La température d'une couleur est un terme qui désigne la couleur émise par un corps noir lorsqu'on le chauffe. Mesuré en degrés Kelvin (K), ce concept est important en vidéo et en photographie puisque chaque type de pellicule réagit mieux à une certaine plage de températures. Généralement, la pellicule utilisée à l'intérieur est conçue pour l'éclairage des lampes incandescentes de studio dont la température se situe à 5400 K et la pellicule utilisée à l'extérieur est conçue pour l'éclairage d'un ciel bleu à 9000 K.

2.1.3 Espaces de couleur

La perception de la couleur est généralement qualifiée en termes de teinte, de saturation, de luminance et de brillance. La teinte est ce qui distingue les couleurs entre elles, par exemple rouge ou bleu. La saturation représente la pureté de la couleur, ou la distance entre la couleur et le gris. Les couleurs pastelées sont beaucoup moins saturées qu'un bleu royal. La luminance est une mesure de l'intensité perçue d'un objet réfléchissant la lumière, tandis que la brillance est une mesure de l'intensité émise par un objet.

En 1931, la CIE¹ a développé le diagramme de chromaticité (voir la figure 2.1) qui met dans un même référentiel la longueur d'onde (en périphérie en μm), la température de couleur (sur la courbe "Black Body Curve" en K), la teinte (en langage naturel) et la saturation (distance du noyau). Le diagramme de chromaticité n'est pas une palette de couleurs complète puisque la chromaticité est indépendante de la luminance.

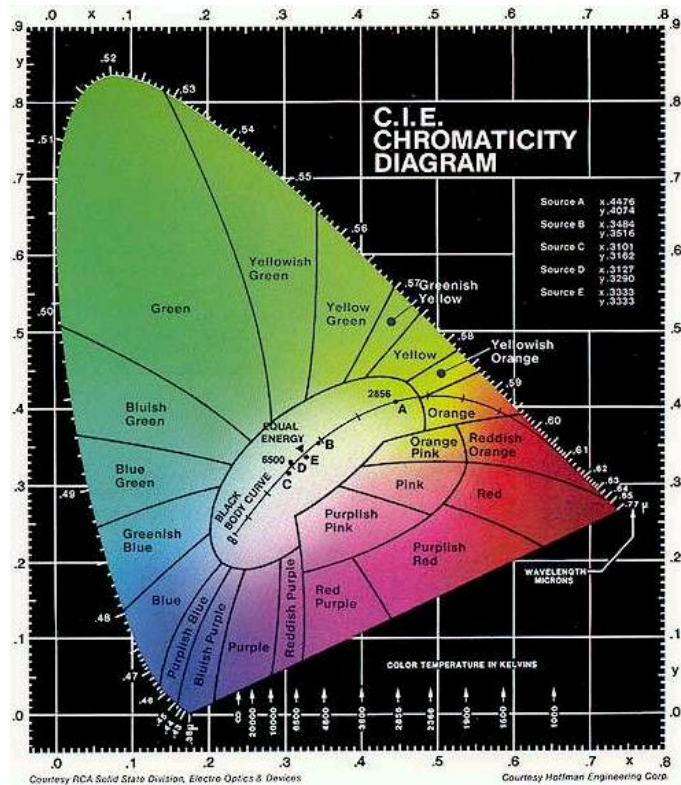


FIG. 2.1 – Diagramme de chromaticité de la CIE tiré de [Owe99]

La représentation de la couleur utilisée aujourd'hui par l'ordinateur consiste à quan-

¹Commission Internationale de l'Éclairage

tifier ces notions en séparant la couleur sur plusieurs canaux, chacun contenant une information spécifique de la couleur. Cette représentation est appelée un espace de couleur, où chaque canal est un axe et chaque couleur correspond à une coordonnée dans l'espace. Parmi les espaces de couleur développés, les plus populaires en composition vidéo numérique sont les espaces RGB^2 et HSV^3 .

Le format le plus répandu est l'espace tridimensionnel RGB qui consiste à reproduire les couleurs par une combinaison de trois couleurs primaires (rouge, vert, bleu) dans un espace cartésien. La pellicule couleur, où trois couches imprégnables par des longueurs d'onde différentes sont superposées, l'écran à rayons cathodiques qui émet les couleurs lorsque des électrons heurtent trois types de phosphores et l'oeil qui perçoit la couleur avec trois types de cônes, utilisent tous cette représentation. Il est plus simple pour les numériseurs d'encoder directement l'information de la pellicule dans ce format. Aussi, il est généralement plus efficace de traiter les images dans ce format, puisqu'on évite ainsi la conversion en RGB pour l'affichage.

La distance entre deux couleurs dans l'espace RGB n'est pas un indice absolu de la différence de couleur perçue par l'oeil. Des espaces de couleur perceptuellement uniformes tels que $CIE L^*u^*v^*$ et $CIE L^*a^*b^*$ sont plus fidèles au système visuel et permettent entre autres de meilleures interpolations de couleur.

Le modèle HSV est une meilleure représentation pour l'utilisateur car les axes correspondent à des éléments plus intuitifs de la couleur, par opposition au modèle RGB qui est avantageux au niveau du matériel. La teinte, la saturation et la luminance sont des concepts bien connus des artistes. Puisque l'éclairage affecte surtout la luminance des objets, ce modèle est utilisé par plusieurs outils de composition, tels que les outils de *matting* par chrominance (voir la section 1.3.3) et de correction de couleur.

Il est possible, par de simples transformations, de convertir les couleurs d'un espace à l'autre. De plus, il existe une panoplie d'autres espaces de couleur, chacun ayant des caractéristiques particulières. Une bonne revue de ces différents modèles est présentée au chapitre 13 du livre de Foley *et al.* [FvDFH96].

²*Red, Green, Blue*

³*Hue, Saturation, Value*

2.2 Correction de couleur

La correction optique de la couleur en photographie consiste à placer un filtre devant la caméra pour simuler un changement de la température de couleur de la lumière. En post-production numérique, on peut simuler ce genre d'effet en établissant une fonction de correspondance entre les couleurs. Ainsi, une correspondance, souvent implémentée à l'aide d'un tableau⁴, est établie entre chaque couleur possible.

La correction de couleur est utilisée en composition vidéo numérique pour créer des ambiances. Le domaine de la psychologie de la couleur a comme objectif de comprendre les émotions provoquées par la combinaison des couleurs. Les coloristes, experts en correction de couleur, exploitent ces concepts dans les films et les publicités. Par exemple, le noir est reconnu comme un symbole d'autorité et de puissance, le bleu représente la tranquillité et le rouge a comme propriété d'augmenter le rythme cardiaque et respiratoire. Les règles à propos des effets et des harmonies des couleurs sont démystifiées dans certains livres [Pfe72, Saw99].

En composition vidéo numérique, un défi consiste à combiner deux images de sources différentes et de créer l'illusion qu'elles proviennent d'un même espace lumineux. En plus de l'intuition artistique nécessaire, cette étape demande une bonne compréhension de la lumière et de la couleur afin de faire correspondre les couleurs. La section 6.3 du livre de Wright [Wri01] explique bien tout le processus en utilisant des techniques de base de traitement d'images.

En premier lieu, en utilisant une version monochrome de la composition, la luminance et le contraste sont corrigés en ajustant le blanc, le noir et les tons de gris intermédiaires. Ensuite, les aspects sur la couleur sont corrigés en appliquant des conversions et transformations à travers les espaces de couleur. Finalement, il est possible d'ajouter d'autres effets comme des ombres ou du brouillard.

2.2.1 Luminance et contraste

La première étape pour remettre l'image d'avant-plan dans le même espace lumineux que l'image d'arrière-plan est de faire correspondre la luminance et le contraste de ces deux images. Souvent implémentée dans l'espace *HSV*, la correspondance entre les valeurs extrêmes doit être établie, avant de faire correspondre les tons de gris in-

⁴*LUT (Look-up Table)*

termédiaires. Ces opérations ont lieu sur le canal de valeur (V). L'histogramme est un outil utile pour analyser et catégoriser chacun de ces ensembles de pixels.

Puisque ces opérations ne tiennent pas compte de la couleur, l'intensité sera utilisée comme valeur des pixels. Pour rester indépendant de la précision des pixels, l'intensité sera exprimée dans la plage comprise entre 0 et 1.

L'histogramme

Un histogramme permet de visualiser la distribution des pixels selon une caractéristique particulière (souvent l'intensité d'un canal). On le représente normalement par un diagramme où les bandes correspondent au nombre d'occurrences d'une caractéristique en fonction de sa valeur. Cet outil permet surtout de compiler des statistiques sur les images et d'isoler un sous-ensemble des pixels.

En appliquant des transformations basées sur l'histogramme, il est possible de modifier la luminance et de rehausser le contraste. Par exemple, une translation horizontale a pour effet d'augmenter ou de réduire la luminance. Aussi, l'égalisation de l'histogramme (voir la figure 2.2), qui consiste à répartir uniformément l'histogramme sur l'ensemble des valeurs de pixels, permet d'obtenir le contraste maximal sans perdre d'information. La coupure d'histogramme est une opération non réversible qui consiste à définir deux extrêmes et répartir les valeurs intermédiaires sur l'ensemble de l'histogramme.

Ajustement de la luminance

L'ajustement de la luminance a pour but de faire correspondre les valeurs extrêmes de luminance, soient les blancs et les noirs. Il est possible de réaliser cette étape en additionnant ou en multipliant la luminance par une constante, ce qui revient à déplacer l'histogramme horizontalement ou à lui appliquer une mise à l'échelle. L'artiste qui utilise cette méthode pour faire correspondre le blanc et le noir entre les images doit avoir une très bonne intuition et être conscient que ces opérations ne sont pas nécessairement réversibles en raison de la saturation de valeurs (c'est le cas des valeurs entre 0.8 et 1.0 sur la figure 2.3).

Une technique plus simple pour faire correspondre les blancs et les noirs consiste à utiliser deux valeurs de référence pour chaque couleur et établir la courbe de correspondance en traçant une droite passant par ces valeurs. La figure 2.4 illustre cette technique

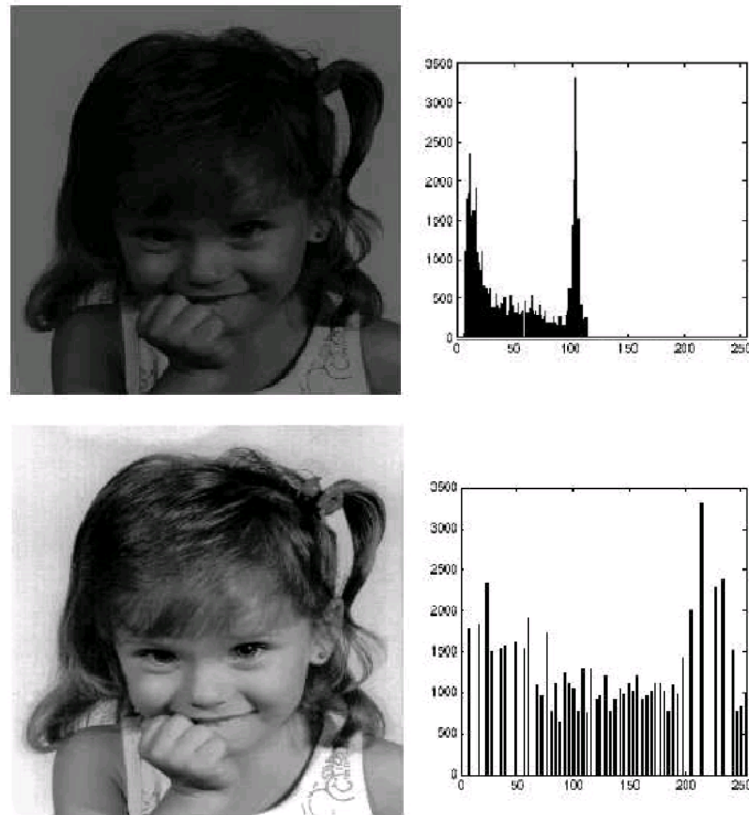


FIG. 2.2 – Égalisation de l'histogramme tirée de [Mig05]

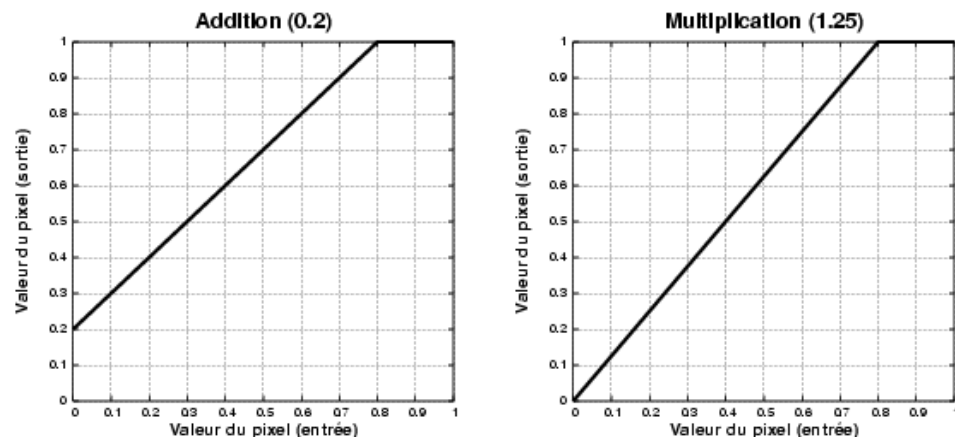
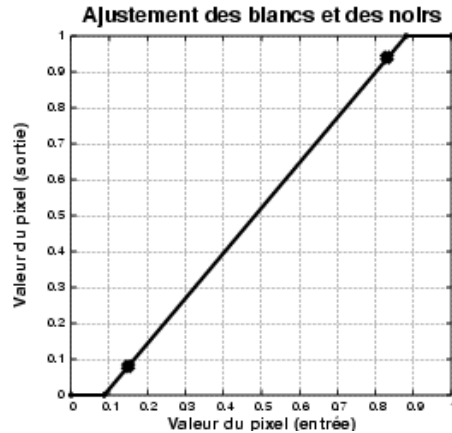


FIG. 2.3 – L'addition et la multiplication par une constante

avec deux valeurs de référence de blanc et de noir.



	Avant-plan	Arrière-plan
Noir	0.15	0.08
Blanc	0.83	0.94

FIG. 2.4 – Valeurs de référence des blancs et des noirs entre deux images

Ajustement du contraste

Une fois les blancs et les noirs ajustés, il faut travailler sur les tons de gris intermédiaires. L'ajustement du contraste et la correction gamma sont normalement les dernières étapes avant de traiter la couleur.

Il existe plusieurs fonctions de correspondance pour corriger le contraste. L'amélioration du contraste sert à augmenter la plage d'intensités entre les valeurs les plus claires et les plus foncées. Comme à la figure 2.4, une technique rudimentaire pour ajuster le contraste consiste à soustraire une constante à la luminance, puis de multiplier le résultat par une autre valeur. Une courbe en forme de S telle qu'illustrée à la figure 2.5 donne généralement de meilleurs résultats.

La correction gamma (γ), définie par l'équation 2.1 où I est la valeur d'entrée et O la valeur de sortie, est probablement la technique la plus utilisée pour ajuster les tons de gris intermédiaires. Comme on peut le constater sur la figure 2.5, la correction gamma préserve l'ensemble de la plage de valeurs. Les tons de gris intermédiaires sont les plus affectés par l'opération. Plus la valeur est au centre de l'histogramme, plus elle sera affectée.

Pour préserver l'ajustement des blancs et des noirs, il est utile d'égaliser l'histogramme avant l'amélioration de contraste ou la correction gamma, puis de remettre l'histogramme dans l'état initial une fois l'opération effectuée.

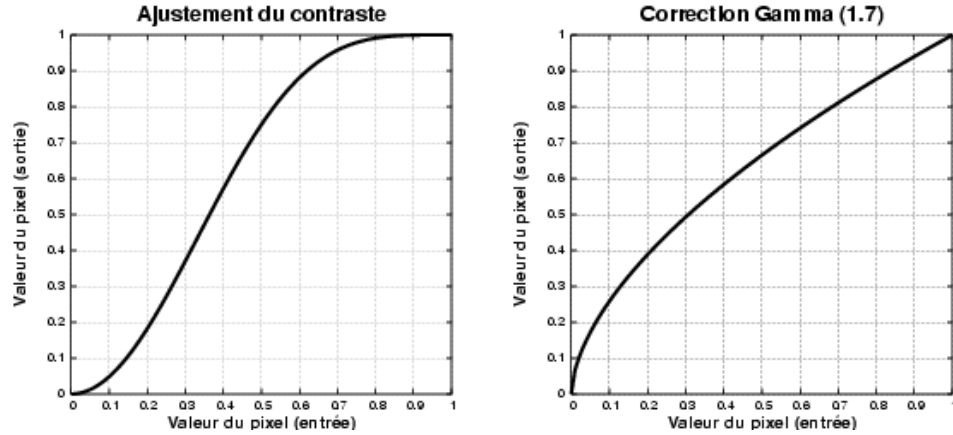


FIG. 2.5 – Amélioration du contraste et correction gamma

$$O = I^{1/\gamma} \quad (2.1)$$

2.2.2 Ajustement des couleurs

Les courbes représentées par les figures 2.3, 2.4 et 2.5 établissent la fonction de correspondance des intensités, entre l'image d'entrée (axe des x) et l'image de sortie (axe des y). Cette technique de correspondance peut aussi être utilisée sur des images couleurs, en traçant une courbe pour chaque canal de couleur. Dans les logiciels commerciaux, ces courbes peuvent être éditées manuellement par l'utilisateur qui déplace des points de contrôle de façon à assurer toute la flexibilité voulue.

La conversion à travers les espaces de couleur s'avère utile en correction de couleur. Par exemple, lorsque la teinte doit être modifiée, mieux vaut travailler dans l'espace *HSV* ou un espace similaire. La majorité du temps, l'ajustement des couleurs se fait dans l'espace *RGB* une fois l'ajustement de la luminance faite.

Puisque la correction de couleur est réalisée globalement sur toute l'image, il est difficile d'obtenir une image parfaite, surtout lorsque l'image d'avant et d'arrière-plan proviennent de sources différentes. Il faut procéder par priorité de façon à faire les choix qui vont engendrer les images les plus réalistes. Par exemple, si l'on veut insérer un personnage filmé à l'extérieur dans un environnement en studio où d'autres personnages sont présents, faire correspondre la couleur de la peau est primordial et la couleur des vêtements devient secondaire. Le jugement du coloriste est important puisque les erreurs d'éclairage de certains types de surfaces sont plus perceptibles que d'autres.

Correction avec un vert constant

La plus grande proportion de la luminance perçue par l'oeil provient du vert, soit 70% selon Wright [Wri01]. Pour conserver l'ajustement de la luminance et du contraste réalisé précédemment, il faut minimiser les modifications à la courbe du canal vert. La technique du vert constant est souvent utilisée à cette fin. Ainsi, en travaillant dans l'espace *RGB*, la contribution d'une couleur doit être augmentée en respectant les règles données au tableau 2.1. Par exemple, pour ajouter du vert, accroître la courbe du vert augmenterait la luminance de l'image. Ainsi, mieux vaut réduire les courbes du rouge et du bleu.

Pour augmenter le	Il faut
rouge	augmenter le rouge
vert	réduire le rouge et le bleu
bleu	augmenter le bleu
cyan	réduire le rouge
magenta	augmenter le rouge et le bleu
jaune	réduire le bleu

TAB. 2.1 – Correction de couleur avec un vert constant

Correction des régions dans l'ombre et des reflets spéculaires

La couleur des régions dans l'ombre est déterminée par l'illumination globale de la scène. Ainsi, les ombres dans une scène extérieure auront souvent une teinte bleutée puisque la source d'illumination indirecte, le ciel, est bleue. En studio, lorsqu'un écran vert est utilisé, la couleur de ces régions est plutôt verdâtre. Il en est de même pour les reflets spéculaires. Ceux-ci dépendent grandement de la température de couleur de la source d'illumination directe.

Pour une composition réaliste, il est fondamental de corriger la couleur des régions dans l'ombre et des reflets spéculaires. Pour y parvenir, la correction de couleur doit être locale, donc seulement dans les régions ciblées. Il est généralement aisé d'extraire ces régions avec le *matting* par luminance (voir la section 1.3.3) puisque ces régions correspondent souvent aux parties extrêmes de l'histogramme. Dans le cas des reflets spéculaires, il faut modifier la saturation et la teinte des régions saturées. L'espace de couleur *HSV* est préférable pour réaliser cette étape.

2.2.3 Ajout d'ombres

Les ombres forment un indice important sur la direction des sources de lumière et elles sont essentielles au réalisme de certaines compositions. En composition d'images, lorsqu'un matte a été extrait d'une image, il est difficile de projeter des ombres réalistes en utilisant des algorithmes d'infographie tels que les volumes d'ombre et les cartes d'ombre puisque la géométrie est inconnue. Les ombres générées par ces images seraient l'équivalent de celles produites par une affiche des objets de l'image.

La méthode des fausses ombres est un trucage populaire en composition vidéo numérique pour imiter une ombre [Wri01]. En appliquant de simples transformations sur le matte (réflexion, rotation, cisaillement, mise à l'échelle, perspective), il est possible d'approximer une ombre projetée sur un sol planaire en appliquant ce matte transformé sous l'objet. La figure 2.6 illustre bien le procédé.

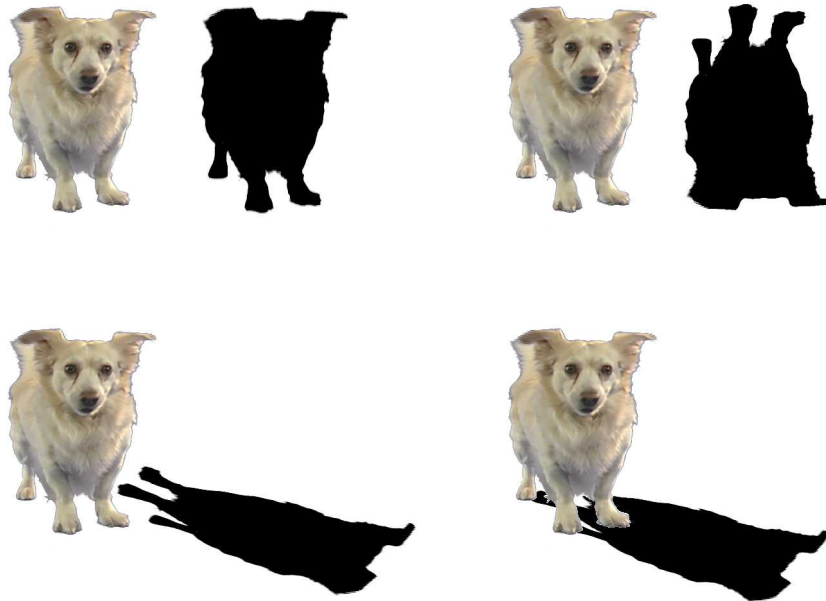


FIG. 2.6 – Technique des fausses ombres

2.2.4 Autres effets

D'autres aspects permettent d'améliorer le réalisme lors de la composition. Il est possible d'ajouter un effet de profondeur de champ en filtrant les images selon la profondeur et du brouillard en appliquant une image de brouillard semi-transparente par-dessus

l'image. L'ajout de défauts typiques à l'image tels qu'un grain similaire à l'arrière-plan peut aussi aider à l'intégration.

Les outils traditionnels de correction de couleur décrits dans ce chapitre ont certainement leur place dans le “pipeline” de post-production. Lorsque les images sont incompatibles, que la direction des lumières est perceptible et différente d'une image à l'autre, il est extrêmement difficile d'uniformiser l'éclairage en utilisant ces techniques. C'est pourquoi, dans ces situations, un bon contrôle de l'éclairage au tournage est crucial.

D'autres techniques plus avancées, plus près de la réalité, ont vu le jour dans les dernières années. Ces techniques requièrent des changements lors du tournage et sont décrites au chapitre suivant.

Chapitre 3

Rééclairage par rendu

Plusieurs aspects de la production de films peuvent être réalisés après le tournage, tels que l'édition et l'ajout d'effets sonores et d'effets visuels. Cependant, l'éclairage demeure un aspect important lors du tournage. La difficulté d'intégrer des scènes dont les éclairages sont incompatibles est considérable due au fonctionnement des techniques de correction de couleur (voir le chapitre 2). Ainsi, les images d'arrière-plan en composition vidéo doivent être choisies avant même de procéder au tournage. L'éclairage en studio doit correspondre le plus fidèlement possible à celui de l'arrière-plan. Généralement, il est difficile de réaliser l'opération correctement même si la majorité du temps en studio est consacrée au positionnement des sources de lumière [WGT⁺05].

Les cinéastes avant-gardistes cherchent constamment à intégrer de nouveaux effets spéciaux dans leurs films. La communauté scientifique en infographie, en traitement d'images et en vision par ordinateur s'implique dans le développement de nouvelles façons de générer des images, de les modifier et d'en extraire de l'information. Depuis quelques années, il est possible de faire des rendus photo-réalistes d'environnements synthétiques. Présentement, le défi est d'extraire d'images réelles la géométrie, la texture, l'illumination et les propriétés de réflectance des surfaces, de façon à leur appliquer les technologies de rendu dans de nouvelles conditions d'éclairage. Ainsi, l'étape coûteuse d'éclairage peut être reléguée à la post-production.

3.1 Rendu d'images

Le rendu consiste à générer une image à partir de la modélisation d'une scène selon un point de vue donné. La valeur de chaque pixel de l'image est définie par l'équation de rendu élaborée par Kajiya [Kaj86], qui évalue la quantité de lumière transférée entre deux points (de x' à x) en tenant compte de la lumière émise et réfléchie par tous les autres points (x'') de la scène :

$$I(x, x') = g(x, x') \left[\varepsilon(x, x') + \int_S \rho(x, x', x'') I(x', x'') dx'' \right] \quad (3.1)$$

où $g(x, x')$ représente le facteur de géométrie ($1/r^2$ s'il n'y a pas d'occlusion entre les deux points, 0 si l'occlusion est opaque, et où r est la distance entre les deux points), $\varepsilon(x, x')$ représente la lumière émise de x' vers x , $\rho(x, x', x'')$ représente la *brdf*¹ à x' , soit la proportion de lumière provenant de x'' réfléchie vers x , et $I(x', x'')$ représente l'intensité lumineuse transférée entre x'' et x' .

L'évaluation analytique de l'équation de rendu donnerait la solution exacte à l'illumination complète de la scène. Cependant, la complexité de ce calcul est trop grande pour les ordinateurs d'aujourd'hui. C'est pourquoi des approches par échantillonnage ont été développées pour approximer cette équation. Pour obtenir des résultats photo-réalistes, la technique du *lancé de rayons* est la plus populaire et consiste à échantillonner la visibilité, la *brdf* et les sources de lumière par l'intersection de demi-droites avec les primitives de la scène. Pour obtenir des images en temps réel en utilisant le matériel graphique, la technique du *z-buffer* est utilisée et consiste à interpoler l'illumination entre les sommets des primitives qui seront projetées sur le plan de l'image.

La *brdf* est la représentation générale des propriétés de réflectance des surfaces. Il s'agit d'une fonction qui détermine la proportion de lumière transférée d'une direction à une autre, lorsque la lumière est réfléchie à la surface d'un objet. Le modèle le plus simple, souvent utilisé ou assumé dans les algorithmes de vision par ordinateur, est le modèle lambertien (*brdf* constante). Ce modèle suppose que la réflexion est complètement diffuse, soit uniformément réfléchie en intensité dans toutes les directions. D'autres modèles simplifiés tiennent aussi compte de la réflexion spéculaire² isotropique³ (*brdf* 3D) et sont par conséquent plus réalistes. Les modèles de Phong [Pho75] et de

¹ *bidirectional reflectance distribution function* ou distribution du coefficient de réflexion bidirectionnel

² réflexion directionnelle qui crée des reflets brillants sur les surfaces

³ indépendante de l'orientation tangentielle de la surface

Blinn [Bli77] sont depuis longtemps supportés par le matériel graphique. Les modèles complets de *brdf* tiennent aussi compte de l'anisotropie et sont plus lourds à calculer et à représenter (*brdf* 4D).

Dans l'industrie cinématographique, en excluant les films d'animation, le rendu est surtout utilisé pour intégrer des objets synthétiques dans l'image. La modélisation des sources de lumière dans le but d'imiter celles de la scène est un processus complexe pour les artistes. L'utilisation de sondes d'illumination est un ajout au niveau du tournage qui permet de simplifier cette étape. Avec les avancées de la technologie, nous espérons voir dans un avenir rapproché l'utilisation des techniques de rendu pour refaire l'illumination complète des objets de la scène en post-production. Présentement, pour y parvenir, des changements considérables et coûteux à la production doivent être réalisés, de façon à extraire la géométrie et la fonction de réflectance des surfaces de la scène.

3.2 Sondes d'illumination

Dans l'équation de rendu (voir l'équation 3.1), les sources de lumière occupent une importance certaine. Il est possible de faire le rendu photo-réaliste d'objets synthétiques en utilisant une carte de radiance (ou d'environnement) qui représente l'ensemble de l'énergie lumineuse qui éclaire la scène à un point précis. Au tournage, une façon simple d'extraire la carte de radiance est d'utiliser une sonde d'illumination qui consiste en une sphère miroir réfléchissant la scène (voir la figure 3.1). En plaçant la sonde d'illumination à l'endroit où l'objet synthétique sera inséré, le rendu de cet objet pourra être réalisé.

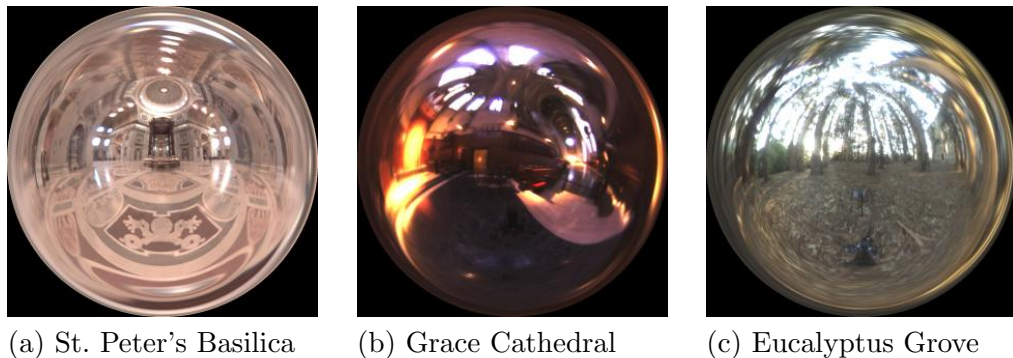


FIG. 3.1 – Sondes d'illumination tirée de [Deb06]

Pour un rendu réaliste, il est préférable d'avoir une représentation des cartes d'envi-

ronnement à haut diapason dynamique (HDR^4). Pour y parvenir, si une caméra HDR n'est pas disponible ou suffisamment précise, il est possible de photographier la sonde avec différents temps d'exposition connus, puis de générer la carte de radiance HDR avec diverses techniques [DM97, FLW02]. Pour obtenir des résultats satisfaisants, il est important de bien positionner la sonde et de conserver la caméra à distance de façon à limiter l'espace occupé par la caméra sur la sonde elle-même.

Une fois la carte de radiance acquise, le rendu des objets synthétiques peut être réalisé. Un ensemble de sources de lumière ponctuelles peut être déduit par un échantillonnage de la carte de radiance. La technique détaillée par Debevec [Deb98] ajoute un modèle des objets environnants la scène de façon à mieux simuler l'illumination globale et les ombres projetées par l'objet synthétique. Plusieurs techniques peuvent être utilisées pour choisir les points à utiliser sur la carte de radiance. La technique de Ostromoukhov *et al.* [ODJ04] permet de déterminer rapidement une bonne distribution de points selon une fonction d'importance générée à partir de la carte de radiance. D'autres techniques [CJAMJ05, BGH05] tiennent aussi compte de la *brdf* de l'objet dans l'échantillonnage. L'objectif de ces techniques par échantillonnage est de réduire le temps de rendu sans affecter la qualité, en lançant les rayons d'une manière stratégique.

3.2.1 Les yeux comme sonde d'illumination

Les cinéastes faisant face à des budgets limités hésitent à modifier leurs techniques de tournage. L'utilisation de sondes d'illumination étant une technique nouvelle, elle demeure encore peu utilisée. En post-production, il est possible d'extraire une carte de radiance même lorsqu'aucune sonde d'illumination n'a été utilisée au tournage. Une idée judicieuse élaborée par Nishino et Nayar [NN04] consiste à utiliser les yeux des acteurs comme sonde d'illumination. En effet, la cornée de l'oeil peut être approximée par un ellipsoïde de taille fixe qui réfléchit la lumière de la scène. Une fois le modèle de cornée calibré avec l'image, celui-ci est utilisé pour échantillonner des sources de lumière avant de faire le rendu d'objets synthétiques.

À la figure 3.2, le visage de l'actrice est substitué par un autre visage. Pour arriver à réaliser ce trucage, la fonction de réflectance du visage à insérer a dû être extraite par une méthode expliquée à la section 3.3.5. Par la suite, il a été possible d'utiliser la carte

⁴*High Dynamic Range*

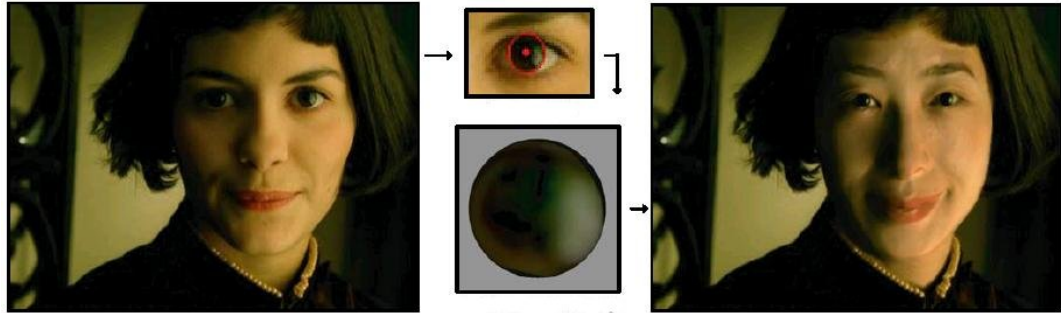


FIG. 3.2 – Les yeux comme sonde d’illumination [NN04]

de radiance extraite à partir de l’oeil pour refaire le rendu.

3.3 Modélisation d’objets réels

La création de scènes synthétiques complexes et réalistes est un processus artistique exigeant et coûteux. La modélisation de géométrie, la création et l’application de textures sur ces modèles, le choix des propriétés de réflectance des surfaces et le positionnement de sources de lumière sont toutes des étapes non triviales. Les groupes de recherche en infographie, en traitement d’images et en vision par ordinateur tentent d’accélérer et d’automatiser ce processus en reconstruisant des objets réels par des procédés optiques (en utilisant des caméras). Les applications de ces technologies sont variées (génération de nouveaux points de vue, interaction de l’imagerie générée par ordinateur, etc.) et ont déjà commencé à donner des résultats en post-production vidéo, notamment pour rééclairer des images.

Certains systèmes interactifs servent à modéliser des scènes réelles à partir d’une ou plusieurs images. D’autres systèmes (généralement plus coûteux) permettent de mieux traiter les séquences vidéo en effectuant la tâche automatiquement. Il existe présentement sur le marché des numériseurs 3D qui permettent d’obtenir automatiquement une image avec un canal de profondeur. D’autres systèmes plus précis utilisent plusieurs caméras ainsi que des projecteurs de lumière structurée. La fonction de réflectance des surfaces peut aussi être retrouvée en utilisant ces techniques. Une revue du fonctionnement de ces techniques permettra de mieux comprendre leur intégration et les difficultés associées dans le contexte de rééclairage en composition vidéo numérique.

3.3.1 Modélisation à partir d'images

L'utilisation de systèmes à plusieurs caméras est une technique de reconstruction passive (sans interaction avec la scène). Le principe de base est de calibrer les caméras entre elles, ce qui permet de déterminer la position du plan de l'image dans l'espace. Pour chaque pixel, une droite est définie par la position du pixel et le centre de projection. La position d'un point 3D correspondra donc à l'intersection des droites associées à ce point. Lorsque ces techniques sont utilisées en studio, les caméras sont souvent calibrées en utilisant des points de calibration sur l'écran bleu.

Le domaine de modélisation et de rendu à partir d'images ou de séquences vidéo⁵ [MPC⁺05] a permis de réaliser des effets spectaculaires dans les films en modélisant les scènes à partir de différents points de vue (voir la figure 3.3). La première étape consiste à segmenter les objets de l'arrière-plan. Une enveloppe visuelle⁶ est ensuite établie en calculant l'intersection des silhouettes des images (mattes) dans l'espace tridimensionnel. Les algorithmes de stéréovision [MP79] et de sculpture d'espace [KS00] permettent de préciser le modèle en réduisant l'enveloppe visuelle.



FIG. 3.3 – Modélisation et rendu à partir de séquences vidéo

La technique de stéréovision [MP79] s'inspire du système visuel humain pour extraire la profondeur dans les images. Elle utilise au moins deux caméras calibrées afin d'extraire la distance entre ces objets et l'image. Dans cette optique, deux problèmes doivent être solutionnés : la correspondance des points d'une image à l'autre et la reconstruction à partir de ces correspondances. La correspondance est souvent basée sur une comparaison des caractéristiques de l'image (couleur, coins, contours, segmentation, etc.) et la profondeur est généralement déterminée par triangulation en exploitant la

⁵image-based and video-based modeling and rendering

⁶visual hull

géométrie épipolaire⁷ [SS02].

La sculpture d'espace [KS00] est une autre technique de reconstruction consistant à exploiter plusieurs points de vue d'un objet dans le but d'éliminer des éléments de volume (*voxels*) disposés dans une grille. À chaque itération, lorsque la projection d'un voxel de surface sur plusieurs images est incohérente, celui-ci est supprimé. Le test de cohérence compare des caractéristiques similaires à celles de stéréovision. Une technique de sculpture d'espace interactive a aussi été développée, et permet de déplacer une seule caméra vidéo autour d'un modèle afin de le raffiner [GPEP04].

3.3.2 Systèmes interactifs

Des systèmes tels que *Façade* [DTM96], *Rekon* [POF98] et *Reality* [DGSR01] permettent de reconstruire des scènes statiques à partir d'un ensemble de photographies. Un artiste résout certaines ambiguïtés en établissant des correspondances entre les images, en détectant les erreurs de reconstruction faites par le système et en ajoutant des contraintes à la scène afin de corriger ces erreurs. Les étapes de reconstruction sont sensiblement les mêmes pour les différents systèmes. Premièrement, les caméras sont calibrées à partir des images. Ensuite, un modèle 3D est généré grâce aux techniques de modélisation à partir d'images. Finalement, les textures de la scène sont retrouvées afin de refaire le rendu d'un point de vue arbitraire. Des extensions à ces systèmes permettent de retrouver les paramètres de réflectance des surfaces en ajoutant des photographies d'un même point de vue sous différentes conditions d'éclairage [LFD⁺99] ou en raffinant les paramètres par des rendus synthétiques [BG01]. Ces systèmes sont efficaces pour modéliser des scènes polygonales tels que des bâtiments et des scènes intérieures.

D'autres systèmes permettent de modéliser une scène, incluant les parties non polygonales, à partir d'une seule image [OCDD01, HEH05]. L'idée est de segmenter l'image en plusieurs couches qui seront disposées dans un espace tridimensionnel. Une interface intuitive permet d'aligner des plans avec les surfaces planaires et la géométrie des éléments courbes sont déterminés par gonflement de silhouettes (voir la section 4.2.1). Des outils tels qu'une brosse de clonage permettent d'éditer les parties cachées de l'image. Il est aussi possible de dissocier la texture de l'illumination à l'aide d'un filtrage bilatéral.

⁷discipline de la géométrie dont les propriétés découlent des correspondances entre des images calibrées

3.3.3 Numériseurs 3D

Les numériseurs 3D au laser permettent de retrouver directement la profondeur d’une scène. Pour l’instant, les appareils commerciaux tels que ceux vendus par *Cyberware* [Cyb06], fonctionnent bien pour les objets de petite taille. Par exemple, il est facile de numériser un visage, un petit objet ou une personne. Pour le projet *Michelangelo* [LPC⁺00], cette technologie a été utilisée pour numériser de grandes statues. Cependant, pour un projet d’aussi grande envergure, beaucoup de temps, de main d’oeuvre ainsi qu’un budget considérable ont été requis.

Dans le contexte de rééclairage, à partir d’une image et de la calibration de la caméra avec la géométrie de la scène (extraite par un numériseur laser), il est possible de refaire le rendu de l’image dans une nouvelle condition d’éclairage [MG97]. Cependant, cette technique n’a pas fait ses preuves pour des images de couleur.

3.3.4 Lumière structurée

La reconstruction par lumière structurée est une solution peu coûteuse pour modéliser automatiquement des objets réels [RCM⁺01]. Grâce à un projecteur calibré de la même façon qu’une caméra, un motif de lumière encodé est projeté dans la scène. La déformation de ces motifs par la géométrie de la scène est capturée par une caméra calibrée. L’analyse de ces déformations permet de retrouver la profondeur des objets de l’image. La technique suppose généralement des objets aux surfaces lambertiennes, ou du moins où il est aisé de distinguer l’illumination en provenance “directe” du projecteur. La technique est par conséquent mal adaptée aux objets spéculaires et aux interréllexions.

3.3.5 Combinaison de techniques

Les différentes versions du *Lightstage* [DHT⁺00, HWT⁺04] combinent plusieurs points de vue et l’utilisation de projections de motifs de lumière multicolores pour capturer la géométrie (représentée par une carte de normales précise) ainsi que les variations temporelles de la réflectance d’un visage humain. Une fois ces informations acquises, il est possible de refaire le rendu d’un visage dans de nouvelles conditions d’éclairage et d’un point de vue arbitraire. En faisant l’extraction de plusieurs configurations de visages, il est possible d’animer ceux-ci en interpolant linéairement la réflectance et la géométrie d’une configuration à l’autre. Une intervention humaine est requise pour

déterminer la correspondance entre les éléments du visage.

La dernière version du *Lightstage* [WGT⁺05] utilise une combinaison de projecteurs de lumière structurée et une caméra, tous calibrés entre eux à hautes fréquences. Cette technologie est destinée au rééclairage en post-production cinématographique. En utilisant des cartes de radiance, il est possible de rééclairer les personnages à l'aide des propriétés de réflectance des surfaces définies par l'utilisateur (voir la figure 3.5). D'autres effets peuvent être réalisés tels que le *matting* automatique, le remplacement des matériaux de surface et l'interpolation d'images par l'analyse du flux optique.

Le coût considérable en équipements, la quantité de mémoire et la puissance de calcul requise pour analyser et traiter la quantité d'images prises au même instant (voir la figure 3.4) sont les facteurs limitants de cette technologie. De plus, une preuve conceptuelle a été réalisée avec des visages humains, mais la flexibilité de la technique pour d'autres objets que les visages de même que pour des scènes plus complexes reste toujours à démontrer.



FIG. 3.4 – Images successives prise par le système de *performance relighting* dans l'espace d'un douzième de seconde, tiré de [WGT⁺05]

Les techniques de rééclairage par rendu en post-production sont prometteuses même si des modifications lors des tournages sont nécessaires. En utilisant ces techniques, les cinéastes pourront bénéficier de d'autres avantages tels que le changement du point de vue en post-production ainsi que des méthodes robustes pour interpoler les images. Ces technologies ont déjà donné des résultats concrets dans l'industrie du film, et des preuves conceptuelles laissent croire que ces technologies vont s'améliorer dans le futur.

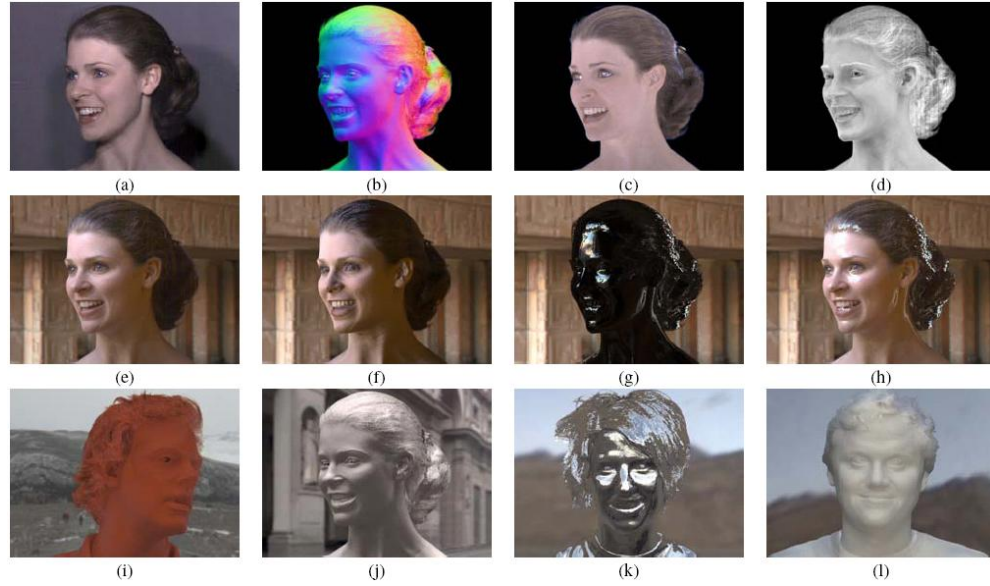


FIG. 3.5 – Applications du système de *performance relighting*, tiré de [WGT⁺05] : (a-d) photographies originales et informations retrouvées par le système, (e-h) photographies rééclairées par une carte de radiance en modifiant le ratio diffus / spéculaire, (i-l) autres effets pouvant être réalisés par le système

Cependant, les avantages des effets réalisés par ces technologies ne compensent pas encore pour l'accroissement des coûts et de la complexité des tournages. Des techniques plus simples pour réaliser des effets similaires en utilisant des séquences filmées de façon traditionnelle seraient bénéfiques à l'industrie. Il en va de même pour tous les tournages ayant déjà eu lieu sans ces technologies. Notre contribution, détaillée au chapitre suivant, tente justement de combler ce vide et permet de réaliser de nouveaux effets d'éclairage, au-delà de la correction de couleur, à partir de séquences traditionnelles.

Chapitre 4

Au-delà de la correction de couleur

Après avoir survolé les moyens utilisés pour modifier l'éclairage en post-production cinématographique, nous pouvons les catégoriser en deux grandes classes : les techniques par correction de couleur (voir le chapitre 2) et les techniques par rendu (voir le chapitre 3). Bien ancrées dans le domaine depuis des décennies, les techniques par correction de couleur, où l'on établit une correspondance entre les couleurs à l'aide d'un tableau, sont rapides, simples d'utilisation et s'appliquent à tous les types d'images. Cependant, pour obtenir des résultats satisfaisant, ces techniques nécessitent des séquences tournées dans des conditions d'éclairages similaires.

En constante évolution, les techniques par rendu exploitent la connaissance de la scène et de ses propriétés pour modifier l'éclairage en post-production, au lieu de fixer l'éclairage lors du tournage. Comme nous l'avons vu au chapitre précédent, extraire des informations supplémentaires telles que les sources de lumière, la géométrie de la scène, la texture et la fonction de réflectance des surfaces peut permettre de réaliser des effets impressionnants. Cependant, une sonde d'illumination, plusieurs caméras, ou encore plusieurs éclairages à un instant donné, ajoutent une complexité de mise en scène et un coût matériel à la fois à la production et à la post-production. Pour l'instant, ces techniques à la fine pointe de la technologie sont donc réservées aux projets cinématographiques à grand déploiement.

Est-il possible d'aller au-delà des techniques par correction de couleur lorsque nous avons une séquence issue d'un tournage traditionnel ? Est-ce qu'un artiste pourrait uti-

liser sa connaissance de la scène pour s’approcher des techniques par rendu tout en conservant les méthodes de production traditionnelles ? Nos travaux répondent au moins partiellement à ces questions en fournissant à l’artiste un système de reconstruction rapide à partir d’une séquence vidéo où les mattes ont été extraits (voir la section 1.3.3). Des effets de rééclairage au-delà de la correction de couleur sont alors rendus possibles en plaçant des sources de lumière virtuelles dans le même environnement que le modèle reconstruit et en manipulant les paramètres de réflectance.

4.1 Un environnement tridimensionnel

Le module *Action* du logiciel de composition vidéo numérique *Discreet Flame/Inferno* [ME06] fournit un environnement de composition quasi-tridimensionnel. Le terme $2D\frac{1}{2}$ est souvent employé pour qualifier ce type d’environnement où les séquences d’images consistent en des panneaux bidimensionnels animés disposés dans un espace tridimensionnel. De façon similaire aux logiciels d’animation, nous pouvons déplacer ces panneaux en fonction du temps pour générer une nouvelle séquence du point de vue d’une caméra virtuelle. De plus, il est possible de déformer la structure planaire des panneaux en éditant une surface bicubique sur laquelle l’image sera appliquée. Modéliser la géométrie du contenu d’une séquence vidéo à l’aide de cet outil reste cependant fastidieux. L’outil a surtout été conçu pour appliquer une déformation grossière à une séquence, sans tenir compte de son contenu.

Il est aussi possible d’ajouter des sources de lumière virtuelles pour rééclairer la scène. Dans ce contexte, le résultat est souvent loin du résultat idéal qui serait de replacer les sources de lumière lors du tournage. Puisque ni la géométrie, ni les fonctions de réflectance des surfaces ne sont connues, le résultat équivaut à placer des sources de lumière autour d’une “affiche” animée de la scène. Il est aussi possible de simuler des ombres planaires, en utilisant la technique des fausses ombres (voir la section 2.2.3) qui consiste à transformer le matte de l’objet afin de créer une ombre.

La majeure partie de nos travaux a été réalisée en collaboration avec l’équipe de recherche et développement en traitement d’images chez *Autodesk Media & Entertainment*. Le premier objectif du projet était de développer un outil interactif pour générer et représenter la géométrie des objets à partir du matte animé. Le second objectif était de pouvoir rééclairer ces objets par des sources de lumière virtuelles en exploitant la

géométrie générée.

Les cartes de normales, introduites par Blinn [Bli78], et les champs de hauteurs, introduits par Cook [Coo84], consistent en des images (ou textures) où chaque pixel (ou texel) correspond respectivement à une normale à la surface et à une hauteur. L'utilisation de l'une ou l'autre de ces structures de données permet de représenter des géométries simples et d'ajouter un effet de profondeur à une image à l'aide du matériel graphique. La conversion d'un champ de hauteurs à une carte de normales est directe.

Pour réaliser nos objectifs, nous avons opté pour le développement d'un outil de génération de cartes de normales à partir d'une séquence vidéo incluant les mattes. Les cartes de normales permettent d'obtenir un effet de profondeur à faible coût lorsque des sources de lumière virtuelles sont utilisées pour modifier l'éclairage.

4.2 Génération de cartes de normales

Comme dans le système *Lightstage* (voir la section 3.3.5), l'extraction automatique et précise des normales des objets de la scène serait idéale. Cependant, retrouver les normales à partir d'une image ou d'une séquence vidéo reste un problème mal posé et donc impossible à résoudre d'une façon unique pour une scène inconnue. Notre objectif est de générer une géométrie "plausible" de la scène. Cete géométrie doit être suffisamment lisse et cohérente dans le temps pour que le changement d'éclairage des images par des sources de lumière virtuelles paraisse réaliste. Notre outil s'ajoute ainsi à la palette d'outils de correction de couleur dans le but de créer un meilleur effet d'immersion lors de la composition vidéo numérique. Il est important de préciser que nous n'adressons pas la reconstruction de surfaces polygonales. Plusieurs systèmes [DTM96, POF98, DGSR01] (voir la section 3.3.2) se sont concentrés sur ces surfaces. Notre approche leur est complémentaire.

La génération de cartes de normales est basée sur deux techniques complémentaires de reconstruction : les silhouettes du matte et le dégradé d'illumination sur les surfaces. La première est insensible à l'éclairage et permet d'inférer une forme globale au contenu de la scène, comme si l'on gonflait les objets à l'intérieur de leurs silhouettes. La surface doit être aussi lisse que possible pour éviter les transitions abruptes d'illumination. Puisque l'oeil est très sensible au contraste, ces transitions empêcheraient d'obtenir un résultat réaliste.

La deuxième technique est plus flexible et utilise l'image comme une carte d'illumination de la scène. Le dégradé d'illumination¹ est analysé de façon à reconstruire une surface de la scène où les objets, incluant des petits détails, sont mieux représentés qu'avec la technique précédente. Cette technique permet d'ailleurs de reconstruire des détails à l'intérieur des silhouettes. Cependant, elle est vulnérable au bruit, aux transitions de couleur, aux reflets spéculaires et aux régions dans l'ombre. Nous verrons que les deux techniques peuvent être combinées d'une manière avantageuse.

4.2.1 Surface à partir du matte

Dans notre contexte, nous déterminons la forme d'un objet dont les silhouettes sont connues, en supposant que l'objet est lisse et bombé vers l'extérieur (*i.e.* vers la caméra). Pour obtenir ce résultat, nous segmentons préalablement l'image en deux régions (voir la section 1.3.3). Un seuillage sur l'intensité du matte est nécessaire pour le transformer en matte binaire. Nous déterminons une hauteur pour chaque pixel (*i.e.* un déplacement perpendiculaire au plan de support du matte) puis nous dérivons une carte de normales du champ de hauteurs en utilisant le produit vectoriel des vecteurs formés par les hauteurs voisines d'un pixel donné.

Travaux antérieurs

Notre technique de gonflement de silhouettes est inspirée du système de modélisation rapide et intuitif *Teddy* [IMT99]. Ce système génère une surface arrondie “plausible” à l'intérieur d'une courbe grossière (silhouette) dessinée par l'utilisateur (voir la figure 4.1). Une triangulation est formée à partir d'une sélection des points de la courbe et de son axe chordal². En expérimentant le système de modélisation, nous réalisons rapidement qu'il est impossible de modéliser des surfaces précises puisque les silhouettes de la surface reconstruite ne correspondent pas exactement à la courbe tracée. Dans notre cas, nous aimerions préserver tous les détails des silhouettes obtenues en rendant un matte binaire. De plus, le système *Teddy* ne supporte pas les cas où des silhouettes sont imbriquées les unes dans les autres en binarisant le matte (*e.g.* un beigne vu de haut), une situation fréquente dans notre contexte.

¹*shading*

²skeleton

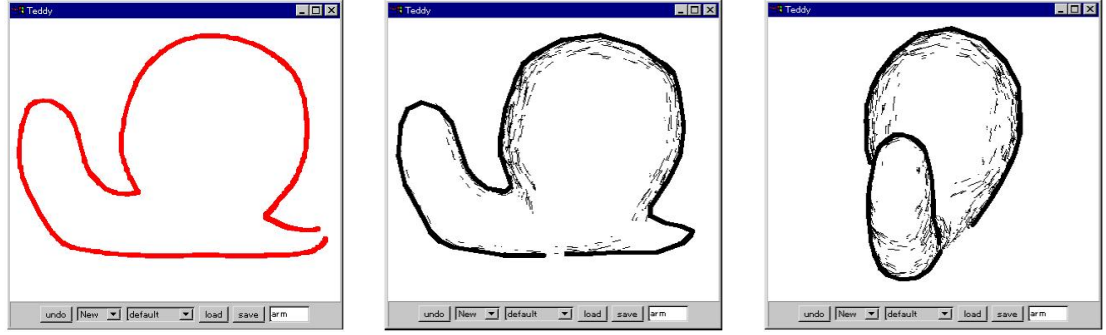


FIG. 4.1 – Système de modélisation *Teddy* tiré de [IMT99]. Gauche : forme dessinée. Centre et droite : forme 3D affichée avec un style “cartoon”

L’approche de Williams [Wil91] génère un champ de hauteurs lisse en appliquant une succession de filtres gaussiens de tailles décroissantes sur un *matte*. À chaque itération, le *matte* (champ de hauteurs) se “diffuse” vers ses silhouettes, mais le résultat est multiplié par le *matte* pour limiter le champ de hauteurs à la région active³ du *matte*. Malheureusement, lorsque la région active du *matte* est beaucoup plus grande que la taille des filtres gaussiens, la partie centrale de la surface ainsi reconstruite demeure plane (voir la figure 4.3(b)), ce que nous souhaitons éviter dans notre système.

Le système de modélisation interactif de Oh *et al.* [OCDD01] (introduit à la section 3.3.2 et illustré par la figure 4.2) génère lui aussi un champ de hauteurs à partir du *matte*. Grâce à une fonction de correspondance, la hauteur d’un pixel $H(p)$ est déterminée en fonction de sa distance à la silhouette $d_{sil}(p)$. Afin d’estimer rapidement cette distance, l’opérateur d’érosion (voir la section 1.3.3) est appliqué sur le *matte* jusqu’à ce que la région active soit nulle. La distance à la silhouette d’un pixel correspond au niveau d’érosion ayant désactivé ce pixel. Par la suite, la hauteur du pixel $H(p)$ est calculée par l’équation 4.1.

$$H(p) = r\sqrt{1 - d'^2} \quad , \quad d'(p) = 1 - \frac{d_{sil}(p)}{\max(d_{sil})} \quad (4.1)$$

La fonction de correspondance forme une ellipse dont les coefficients dépendent de la distance maximale à la silhouette $\max(d_{sil})$ et d’une valeur r définie par l’usager (voir la figure 4.2(a)). Cette technique de reconstruction produit des discontinuités résultant de l’intersection de deux ellipses, chacune définie par une distance plus courte que $\max(d_{sil})$ à la silhouette la plus proche. Celles-ci apparaissent sous formes de crêtes surtout sur

³ensemble des pixels du *matte* où la transparence est nulle

les parties étroites du matte puisqu'à ces endroits, la fonction d' n'atteint jamais zéro. Ainsi, la pente à la surface (gradient de hauteur) n'est jamais nulle dans ces parties. La figure 4.2(b,c) illustre ces problèmes.

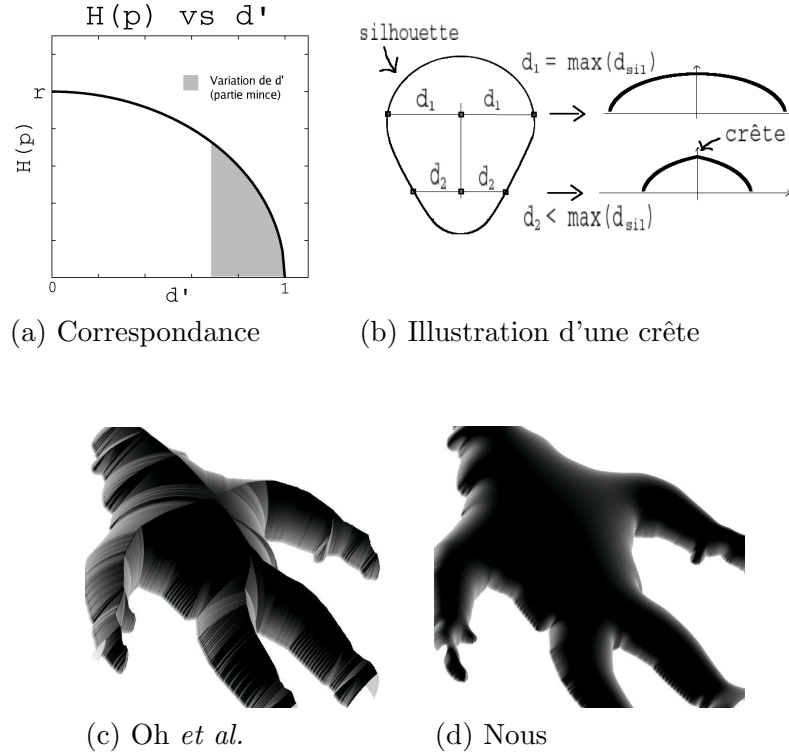


FIG. 4.2 – Reconstruction par la méthode de Oh *et al.* [OCDD01] par rapport à la nôtre

Les faiblesses des travaux précédents nous ont encouragé à développer une nouvelle technique pour générer un champ de hauteurs à partir d'un matte. La figure 4.3 illustre en 2D les résultats typiquement produits avec les différents algorithmes de reconstruction. Initialement, la surface est plane. La technique de Williams [Wil91] produit un effet arrondi près de la silhouette, mais plane vers le centre. La technique de Oh *et al.* [OCDD01] est trop pointue vers le centre lorsqu'il s'agit d'un élément étroit. Notre approche produit une surface mieux bombée près de la silhouette et courbée vers le centre.

Notre approche

Contrairement aux autres techniques de génération de hauteurs à partir d'un matte, notre technique génère une surface sans parties planes et entièrement lisse même

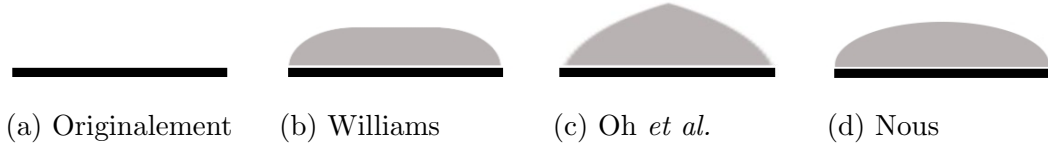


FIG. 4.3 – Illustration en 2D des techniques de génération de hauteurs à partir du matte

dans les parties étroites du matte. Les silhouettes S sont définies comme l'ensemble des pixels de la région inactive du matte touchant à au moins un pixel de la région active. La hauteur d'un pixel $H(p)$ est déterminée en analysant l'ensemble des distances entre ce pixel et les silhouettes. Notre champ de hauteurs est calculé en appliquant l'équation 4.2 sur tous les pixels de la région active du matte.

$$H(p) = \frac{1}{\int_{s \in S} \frac{\partial s}{\text{distance}(p,s)}} \quad (4.2)$$

En analysant l'équation 4.2, nous constatons que l'intégrale de l'inverse des distances aux silhouettes pour un pixel est calculée. Ainsi, plus la distance entre deux de ces points est faible, plus son poids est grand dans l'intégrale. Puisque la hauteur correspond à l'inverse de l'intégrale, plus un pixel est éloigné de la silhouette, plus sa hauteur est élevée. Une fois les hauteurs calculées, la normale à la surface d'un pixel $N(x, y)$ est obtenue en calculant le produit vectoriel entre le gradient de hauteur horizontal et vertical, ce qui se simplifie en l'équation 4.3.

$$N(x, y) = \left(-\frac{\partial H}{\partial x}, -\frac{\partial H}{\partial y}, 1 \right) \quad (4.3)$$

Comparativement à la méthode de Oh *et al.* [OCDD01], comme tous les pixels des silhouettes sont pris en compte, la surface reconstruite est plus lisse et moins sensible lorsque certaines sections se rapprochent les unes des autres. Le calcul est cependant plus intensif puisque tous les pixels de silhouette doivent être traités pour déterminer une hauteur.

La complexité de l'algorithme est dans l'ordre de $O(nm)$, où n et m sont respectivement le nombre de pixels de la région active, et le nombre de pixels sur les silhouettes. Pour des images de petite taille (*e.g.* 360×243), nous pouvons aisément calculer l'équation 4.2 en temps réel. Pour des images de plus grande taille, nous accélérons

le processus avec un *quad-tree* et un calcul à basse résolution pour générer les normales en temps interactif.

Accélération du calcul

Puisque notre fonction pour déterminer la surface est lisse mais coûteuse, il est judicieux d'utiliser l'interpolation pour accélérer le calcul. Un *quad-tree* est un arbre de subdivision qui permet d'accélérer les algorithmes de traitement d'images (voir la figure 4.4(a)). Dans notre cas, puisque notre surface reconstruite est lisse, nous pouvons calculer notre fonction pour déterminer la hauteur de certains pixels, puis utiliser l'interpolation quadratique pour estimer la hauteur des autres pixels. Il faut préciser que le traitement se fait seulement dans les cellules du *quad-tree* contenant une région entièrement active du matte.

Dans les algorithmes 4.1 et 4.2, la fonction *qerp* utilise l'interpolation quadratique pour déterminer la hauteur d'un pixel en le projetant sur la parabole passant par trois points colinéaires (p_1 , p_2 et p_3) où les hauteurs ont été calculées (voir l'équation 4.4). La fonction *interpoler_cellule* utilise la fonction *qerp* à chaque pixel à l'intérieur d'une cellule du *quad-tree* au dernier niveau de subdivision, c'est-à-dire quand l'erreur relative de l'interpolation des coins d'une cellule est faible. L'erreur tolérée *seuil_erreur* est une constante fournie par l'utilisateur.

$$\begin{aligned}
 H(p) = qerp(p, p_1, p_2, p_3) = & \frac{(p - p_2)(p - p_3)}{(p_1 - p_2)(p_1 - p_3)} H(p_1) + \\
 & \frac{(p - p_1)(p - p_3)}{(p_2 - p_1)(p_2 - p_3)} H(p_2) + \\
 & \frac{(p - p_1)(p - p_2)}{(p_3 - p_1)(p_3 - p_2)} H(p_3)
 \end{aligned} \tag{4.4}$$

L'accélération précédente permet de conserver un temps de calcul interactif pour des images de taille moyenne (*e.g.* 720×486). Cependant, dans le cas où nous avons une plus grande image, nous calculons notre champ de hauteurs à l'aide d'un matte à résolution réduite pour conserver l'interactivité. Par la suite, le champ de hauteurs est converti à la résolution initiale avant de calculer les normales. Le coût supplémentaire pour redimensionner les images est nettement inférieur à celui consistant à faire le calcul en pleine résolution. Cette technique doit cependant être utilisée avec prudence, puisqu'il arrive de perdre les détails fins du matte lorsque nous réduisons sa taille (voir la figure

Algorithme : quadtree_interpolation(x_1, x_2, y_1, y_2)

début

$\bar{x} \leftarrow \frac{x_1+x_2}{2}$;

$\bar{y} \leftarrow \frac{y_1+y_2}{2}$;

/* Évaluer les valeurs extrêmes */

pour chaque $x \in \{x_1, \bar{x}, x_2\}$, $y \in \{y_1, \bar{y}, y_2\}$ **faire**

 | évaluer_H((x, y))

fin

/* Estimer l'erreur d'interpolation */

$erreur \leftarrow 0$;

pour chaque $x \in \{\frac{x_1+\bar{x}}{2}, \frac{\bar{x}+x_2}{2}\}$ **faire**

 | $x_{eval} \leftarrow \text{évaluer_H}((x, \bar{y}))$;

 | $x_{interp} \leftarrow \text{qerp}((x, \bar{y}), (x_1, \bar{y}), (\bar{x}, \bar{y}), (x_2, \bar{y}))$;

 | $erreur \leftarrow erreur + 1 - \frac{\min(x_{eval}, x_{interp})}{\max(x_{eval}, x_{interp})}$;

fin

pour chaque $y \in \{\frac{y_1+\bar{y}}{2}, \frac{\bar{y}+y_2}{2}\}$ **faire**

 | $y_{eval} \leftarrow \text{évaluer_H}((\bar{x}, y))$;

 | $y_{interp} \leftarrow \text{qerp}((\bar{x}, y), (\bar{x}, y_1), (\bar{x}, \bar{y}), (\bar{x}, y_2))$;

 | $erreur \leftarrow erreur + 1 - \frac{\min(y_{eval}, y_{interp})}{\max(y_{eval}, y_{interp})}$;

fin

/* Interpoler ou subdiviser */

si $erreur < \text{seuil}_{erreur}$ **alors**

 | interpoler_cellule($x_1, \bar{x}, x_2, y_1, \bar{y}, y_2$) ;

fin

sinon

 | quadtree_interpolation($x_1, \bar{x}, y_1, \bar{y}$) ;

 | quadtree_interpolation($x_1, \bar{x}, \bar{y}, y_2$) ;

 | quadtree_interpolation($\bar{x}, x_2, y_1, \bar{y}$) ;

 | quadtree_interpolation($\bar{x}, x_2, \bar{y}, y_2$) ;

fin

fin

Algorithme 4.1 : Quad-tree pour accélérer la génération du champ de hauteurs

Algorithme : `interpoler_cellule($x_1, \bar{x}, x_2, y_1, \bar{y}, y_2$)`

début

 /* Interpoler verticalement */

pour chaque $y \in \{y_1 < y < y_2\}$ **faire**

pour chaque $x \in \{x_1, \bar{x}, x_2\}$ **faire**

`qerp((x, y), (x, y_1), (x, \bar{y}), (x, y_2)) ;`

fin

fin

 /* Interpoler horizontalement */

pour chaque $y \in \{y_1 < y < y_2\}$ **faire**

pour chaque $x \in \{x_1 < x < x_2\}$ **faire**

`qerp((x, y), (x_1, y), (\bar{x} , y), (x_2, y)) ;`

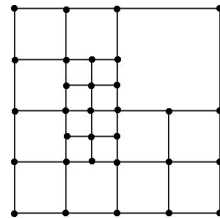
fin

fin

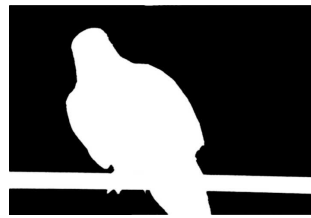
fin

Algorithme 4.2 : Interpolation des cellules du *quad-tree*

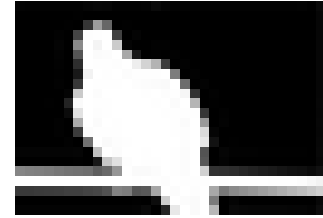
4.4(b,c)). Une analyse de la performance de notre outil est présentée à la section 5.4 pour diverses résolutions d'image.



(a) *Quad-tree*



(b) Pleine résolution



(c) Basse résolution

FIG. 4.4 – Optimisations par *quad-tree* et par calcul à basse résolution

Note sur l'instabilité temporelle

L'outil de génération de normales s'avère très utile dans les scènes où l'occlusion est faible et où le gonflement des silhouettes est une bonne approximation de l'objet filmé. Dans des scènes où l'occlusion est forte, les techniques basées sur le matte pour générer un champ de hauteurs sont plus instables temporellement. Ces problèmes sont surtout causés par l'occlusion qui fait apparaître et disparaître des silhouettes (voir

la section 5.3.1). Une segmentation plus détaillée des parties individuelles de l'objet pourrait atténuer ce phénomène, sans pour autant le résoudre. Nous préférons éviter cette technique de reconstruction lorsque ce problème est trop important.

4.2.2 Surface par dégradé d'illumination

Le dégradé d'illumination sur la surface que l'on souhaite retrouver est une autre propriété que nous pouvons exploiter. L'approche de Horn [Hor90] retrouve une normale à la surface d'un objet en analysant le dégradé d'illumination⁴. Pour l'instant, un sous-ensemble d'hypothèses doit être validé pour assurer la convergence vers la solution. Nous supposons que la surface est éclairée par une seule source de lumière, que la réflectance est uniforme sur la surface (modèle lambertien) et que l'intensité de l'image correspond à la lumière réfléchi (l'image n'est pas saturée, sans ombres ni illumination globale). Puisque l'éclairage dans la réalité est complexe et que les caméras sont imparfaites, ces techniques imprécises arrivent tout de même à retrouver généralement une bonne approximation de la surface dans des conditions normales.

Il y a plusieurs façons de résoudre le problème. Dans notre cas, nous voulons une approximation rapide et grossière de la surface. Fang *et al.* [FH04, FZGH05] ont utilisé une version simpliste d'extraction de surface par dégradé d'illumination à partir du gradient inverse d'une photographie. Avec suffisamment de lissage, leur surface reconstruite est assez précise pour générer et appliquer une texture sur cette surface tout en donnant une bonne approximation des déformations sur la surface.

Notre contexte de réillumination est similaire à celui de Fang *et al.* Utiliser leur technique mène à une surface temporellement continue dans les cas idéaux : source de lumière fixe, aucune ombre, surface lisse lambertienne. Nous pouvons supposer que la normale à l'endroit où la luminance est la plus forte I_{max} pointe vers la lumière et l'endroit où la luminance est la plus faible I_{min} a une normale perpendiculaire à la direction de la lumière. En supposant que la lumière vient de la caméra, nous déterminons la normale au pixel (x, y) par l'équation 4.5. Selon nos expériences, les variations de l'algorithme où la direction précise de la lumière est fournie par l'utilisateur [FH04] ou estimée automatiquement [TV98] ne génèrent pas de meilleurs résultats mais requièrent un calcul beaucoup plus intensif.

⁴*shape from shading*

$$N(x, y) = \left(-\frac{\partial I}{\partial x}, -\frac{\partial I}{\partial y}, \frac{I(x, y)}{I_{max} - I_{min}} \right) \quad (4.5)$$

Le gradient de l'image est très sensible au bruit et aux transitions de couleur. Pour atténuer cette faiblesse, nous utilisons les outils communs de composition vidéo numérique tels que la coupure d'histogramme (voir la section 2.2.1) et l'application de filtres (voir la section 1.3).

4.2.3 Éditer les cartes de normales

Une fois les normales de base générées, il est intéressant de modifier et combiner ces normales, de façon à donner à l'utilisateur un contrôle sur le résultat final. La transformation la plus courante est d'augmenter ou de réduire la dénivellation de la surface reconstruite. Pour ce faire, nous multiplions simplement la hauteur, *i.e.* la composante z de la normale. Lorsque les normales sont bruitées, comme c'est souvent le cas lorsque nous utilisons le dégradé d'illumination, nous filtrons les normales spatialement et temporellement à l'aide d'un filtre gaussien dont la taille est spécifiée par l'utilisateur. Il ne faut pas oublier de renormaliser les normales après chaque opération.

Pour combiner deux normales (n_1 et n_2), nous interpolons linéairement le résultat selon l'importance de chaque composante. Nous fixons la normale finale n_f dans le plan formé par les deux normales initiales. L'angle entre n_1 et n_f est proportionnel à l'importance de n_1 par rapport à n_2 . L'importance des normales peut dépendre de l'angle avec l'axe des z , ou encore être spécifiée directement par l'utilisateur. À la section 5.1, l'ensemble des paramètres d'extraction de cartes de normales sont présentés et illustrés.

La figure 4.5 illustre le résultat des différentes techniques de génération de normales de notre système. L'axe des x utilise le canal bleu de l'image, l'axe des y le canal rouge et l'axe des z , le canal vert.

4.3 Rééclairage

Le système utilisé pour rééclairer les séquences est basé sur le modèle de Phong (voir la section 3.1) qui est supporté par le matériel graphique. Les normales à la surface sont déterminées par la carte de normales et les paramètres de réflectance sont uniformes

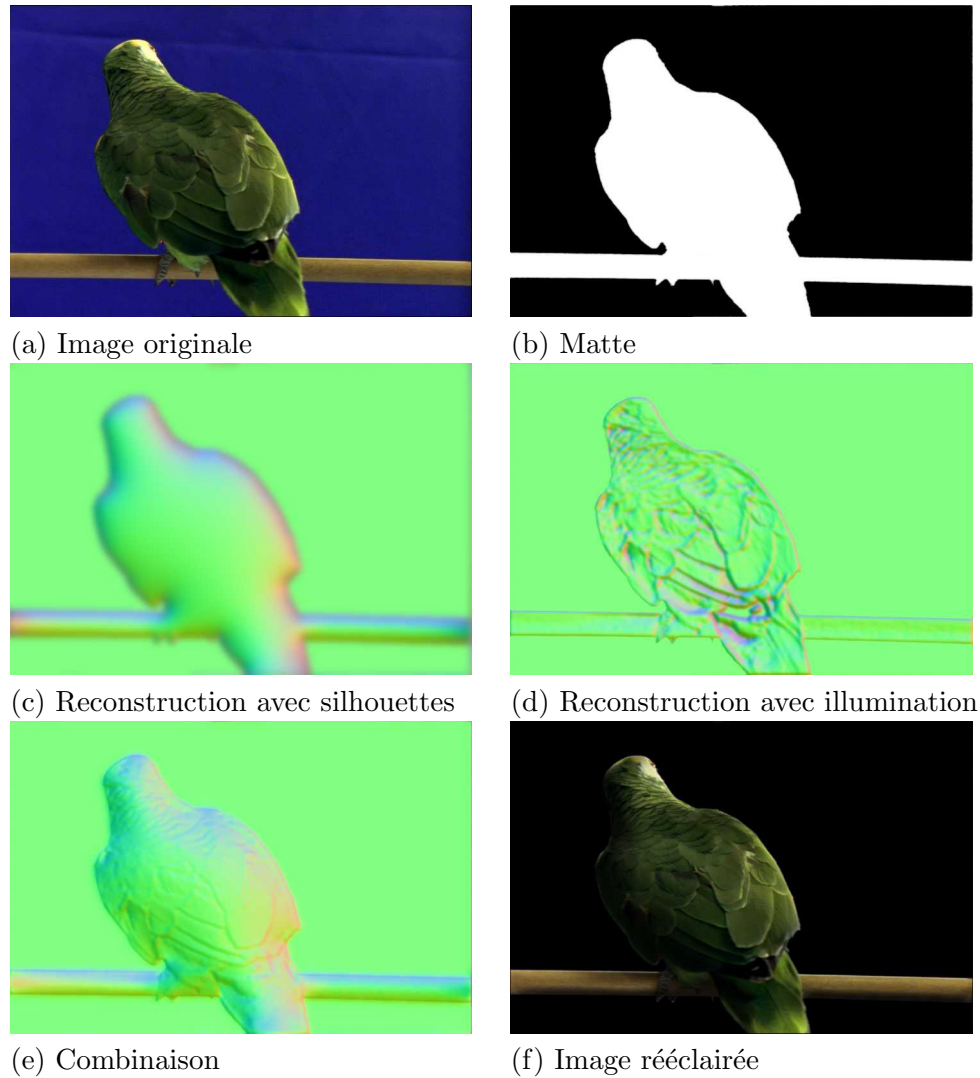


FIG. 4.5 – Illustration de notre système de génération de cartes de normales

sur toute l'image. L'utilisateur a le contrôle sur la position, l'intensité et l'ouverture⁵ des sources de lumière virtuelles ainsi que sur les paramètres de réflectance de la surface pour ajuster l'illumination dans les séquences.

Il est possible d'utiliser la carte de normales pour générer la géométrie de l'objet par un champ de hauteurs, permettant ainsi d'utiliser des techniques de calcul d'ombre réalistes telles que les volumes d'ombres et les cartes d'ombres [Cro77, Wil78]. Nous ne sommes plus limités à utiliser la technique des fausses ombres, uniquement valide sur les plans. Projeter des ombres sur des surfaces arbitraires est désormais possible en utilisant nos cartes de normales, ce qui facilite l'intégration d'objets synthétiques dans les scènes réelles et vice versa.

⁵l'angle de projection d'une source de lumière directionnelle

4.4 Intégration dans un contexte de composition vidéo numérique

Notre technique de génération de normales a été implémentée dans une version future⁶ de *Discreet Flame* [ME06], un système de composition vidéo numérique professionnel. L'outil principal de génération de normales, un noeud dans le module *Batch*, prend en entrée une séquence d'images et de mattes. Le résultat est une séquence de cartes de normales.

L'interface de l'outil permet d'activer la génération de normales par matte et par dégradé d'illumination. Si les deux modes sont activés, les normales doivent être combinées et la pondération des normales est établie par l'utilisateur.

Dans le cas des normales à partir du matte, nous pouvons dilater le matte (voir la section 1.3.3) et mettre à l'échelle la hauteur de la surface. Dans le mode par dégradé d'illumination, nous pouvons préfiltrer et ajuster l'histogramme de l'image, mettre à l'échelle la hauteur de la surface et filtrer la carte de normales résultante.

L'outil de réillumination a été implémenté dans le module *Action* (voir la section 4.1). Il est maintenant possible d'utiliser une carte de normales pour rééclairer une séquence vidéo en déplaçant des sources de lumière virtuelles.

⁶il est impossible d'inclure une copie d'écran de l'outil en action puisqu'il s'agit d'une version non-annoncée

Chapitre 5

Résultats

Dans les chapitres précédents, nous avons exploré le “pipeline” de post-production, les techniques rapides et abordables de correction de couleur et les effets réalistes pouvant être obtenus avec les techniques coûteuses de rééclairage par rendu. L’intérêt est évident pour un outil pouvant créer des effets similaires aux techniques par rendu au coût des techniques par correction de couleur. Dans cette optique, nous avons développé une nouvelle technique de rééclairage basée sur la reconstruction d’une surface courbe “plausible” à partir d’une image et d’un matte. Dans le “pipeline” de post-production, cette technique s’insère entre le *matting* et la correction de couleur.

Le présent chapitre illustre le processus de reconstruction à partir d’images. Après que le système ait généré automatiquement une carte de normales, des paramètres pour chaque stratégie de reconstruction peuvent être ajustés afin d’améliorer le résultat final. Quelques exemples de résultats typiques ainsi que différents problèmes découlant de notre technique sont présentés. De plus, une analyse de performance de notre outil a été réalisée pour différentes résolutions d’images.

Il est important de spécifier que les résultats ont été obtenus par l’auteur de cet outil, un étudiant gradué qui n’est pas un artiste professionnel. Les résultats proviennent de la version préliminaire du logiciel *Discreet Flame* 10 dans laquelle notre outil a été implémenté. Les séquences filmées devant un écran bleu ont été fournies par *Autodesk Media & Entertainment* [ME06].

5.1 Extraction de cartes de normales

L'extraction des normales est l'étape déterminante pour assurer l'intégration réussie d'un avant-plan dans une composition. Pour accomplir cette tâche, l'artiste a le choix de conserver les normales générées automatiquement par le système (voir le chapitre 4) ou d'améliorer le résultat obtenu en modifiant les paramètres de reconstruction. Pour mieux observer l'effet du changement des paramètres, l'image du *perroquet* (voir la figure 4.5) a été rendue avec une source de lumière virtuelle directionnelle située en haut à droite. Puisque l'outil fonctionne en temps réel, il est possible de visualiser simultanément la carte de normales et l'image rééclairée dans le processus de reconstruction. La différence absolue de couleurs avec l'image originale est aussi illustrée à droite sur les figures afin d'apprécier les différences plus subtiles.

5.1.1 Reconstruction automatique

Comme nous l'avons vu au chapitre précédent, deux stratégies de reconstruction sont présentées à l'utilisateur : à partir des silhouettes et à partir du dégradé d'illumination. Sans intervention de l'utilisateur, notre système génère une carte de normales automatiquement. La figure 5.1 illustre le résultat obtenu automatiquement selon l'approche de reconstruction sélectionnée. La reconstruction basée sur les silhouettes (a) présente des effets d'illumination localisés près des silhouettes tandis que la reconstruction basée sur le dégradé d'illumination présente des effets localisés où le gradient d'intensité lumineuse est élevé, c'est-à-dire autour de certaines plumes. En combinant les deux cartes de normales, les deux effets sont générés.

5.1.2 Paramètres de reconstruction à partir des silhouettes

La reconstruction à partir des silhouettes permet de générer une surface lisse et bombée. L'effet sur l'illumination varie en fonction de la direction de la source de lumière virtuelle et de la distance aux silhouettes. L'intensité est accrue du côté de la source de lumière et atténuée du côté opposé. Le paramètre de mise à l'échelle détermine l'ampleur de cet effet. Plus ce paramètre est élevé, plus la surface est bombée et plus l'effet d'illumination est prononcé. Par la dilatation du matte (voir la section 1.3.3), il est possible de réduire l'effet dû au rééclairage près des silhouettes ou encore d'éliminer certains détails des silhouettes qui créent un effet indésirable (*e.g.* incohérence temporelle).

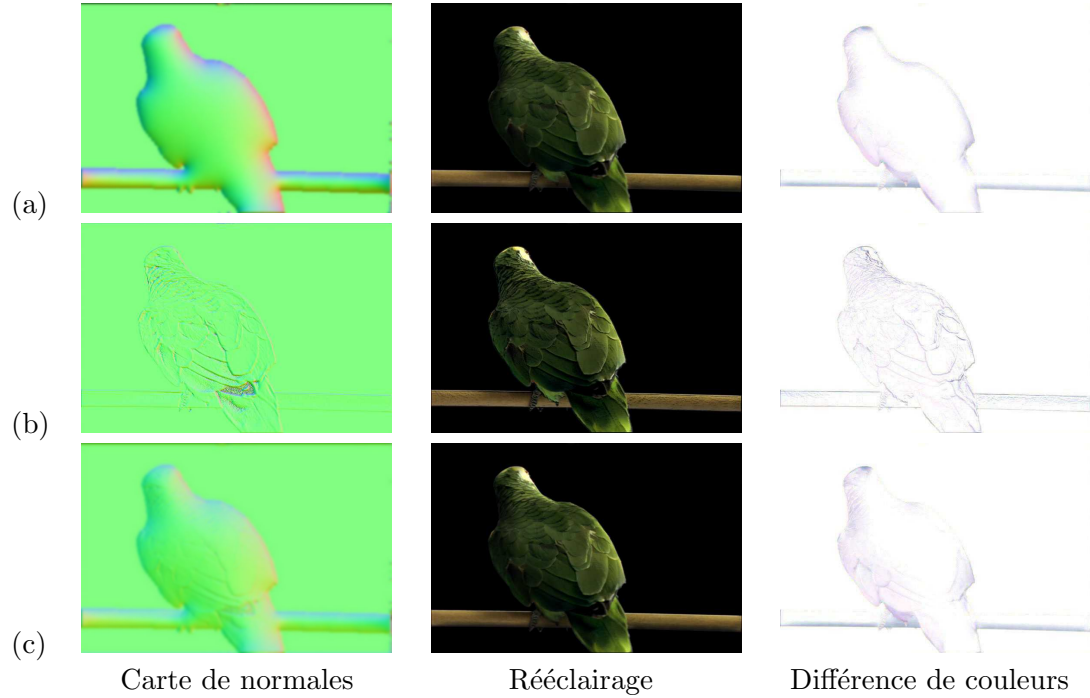


FIG. 5.1 – Reconstruction automatique : (a) par silhouettes, (b) par dégradé d'illumination, (c) combinaison 50% : 50%

L'effet des différents paramètres est illustré à la figure 5.2. En (a), nous remarquons que l'effet d'ombrage est plus important lorsque la mise à l'échelle est grande. L'effet de la dilatation est subtile. Une dilatation élevée (en (c)) ne préserve pas tous les détails du *matte*, ce qui engendre moins de variations d'intensité qu'une dilatation faible (en (b)). Ainsi, on peut observer en (c) une perte de détails près de la patte gauche du *perroquet*. Il est important de noter que le *matte* est modifié seulement pour la génération de normales et non pour la composition.

5.1.3 Paramètres de reconstruction à partir du dégradé d'illumination

La reconstruction à partir du dégradé d'illumination modifie l'éclairage aux endroits où le gradient d'intensité lumineuse est élevé dans l'image. Un facteur de mise à l'échelle est appliqué sur la composante z des normales, ce qui a pour effet de rapprocher ou d'éloigner la normale vers le plan de l'image (xy). Puisque le gradient d'intensité lumineuse varie localement sur l'image, les normales obtenues à ce stade peuvent être bruitées et incohérentes entre elles. Pour réduire ces phénomènes et adoucir l'illumination, nous filtrons les normales en appliquant un filtre gaussien sur le gradient d'illumination.

À la figure 5.3, nous observons des cartes de normales générées avec deux facteurs

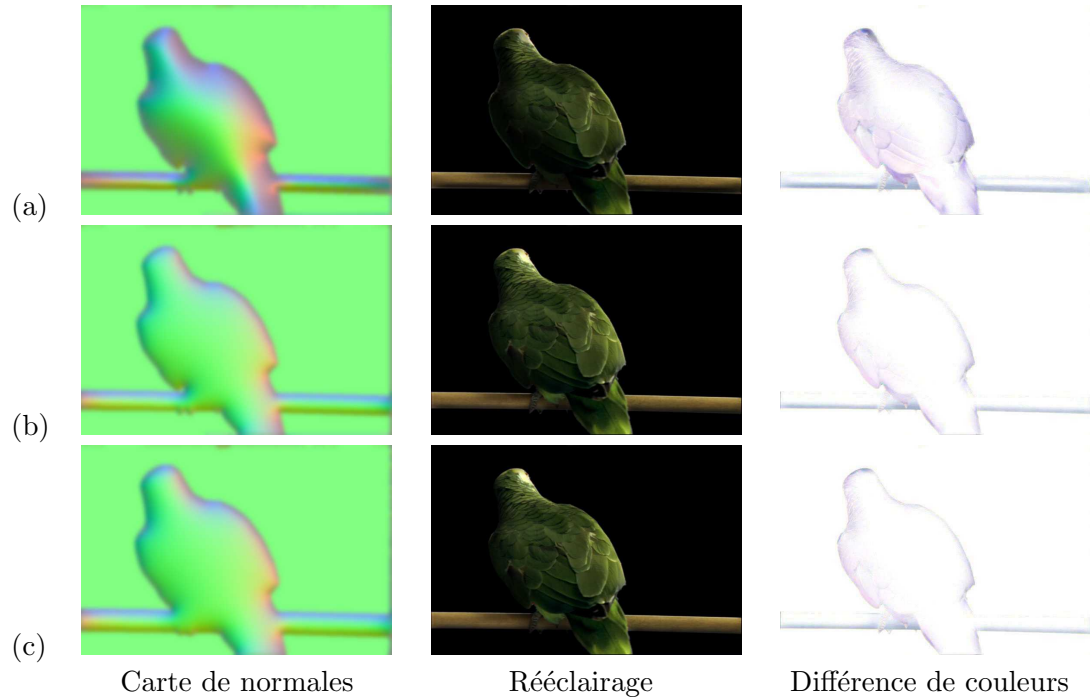


FIG. 5.2 – Reconstruction à partir des silhouettes : (a) mise à l'échelle (5), (b) dilatation du masque de 5 pixels, (c) dilatation du masque de 25 pixels

(0.5 et 0.05) de mise à l'échelle différents ((a) et (b)). Plus le facteur est faible et plus l'image rééclairée est bruitée. Même si l'effet du bruit peut être exploité d'un point de vue artistique (il s'agit d'un type de rehaussement de contours), nous souhaitons réduire fortement ces artefacts de façon à atteindre un niveau de réalisme plus acceptable. Lorsqu'un filtre gaussien (rayon de 5 et 25 pixels) est appliqué ((c) et (d)), les normales sont lissées et l'image rééclairée est plus réaliste.

5.1.4 Combinaison

Une fois les paramètres de reconstruction déterminés pour chaque stratégie, nous pouvons combiner les deux cartes de normales résultantes afin d'obtenir les deux effets d'illumination. Une pondération permet d'équilibrer l'importance de chacune. La figure 5.4 illustre la combinaison de deux cartes de normales avec des pondérations différentes. Nous remarquons que la technique à partir des silhouettes est lisse et que la technique utilisant le dégradé d'illumination permet d'ajouter des détails d'illumination à l'intérieur de l'objet.

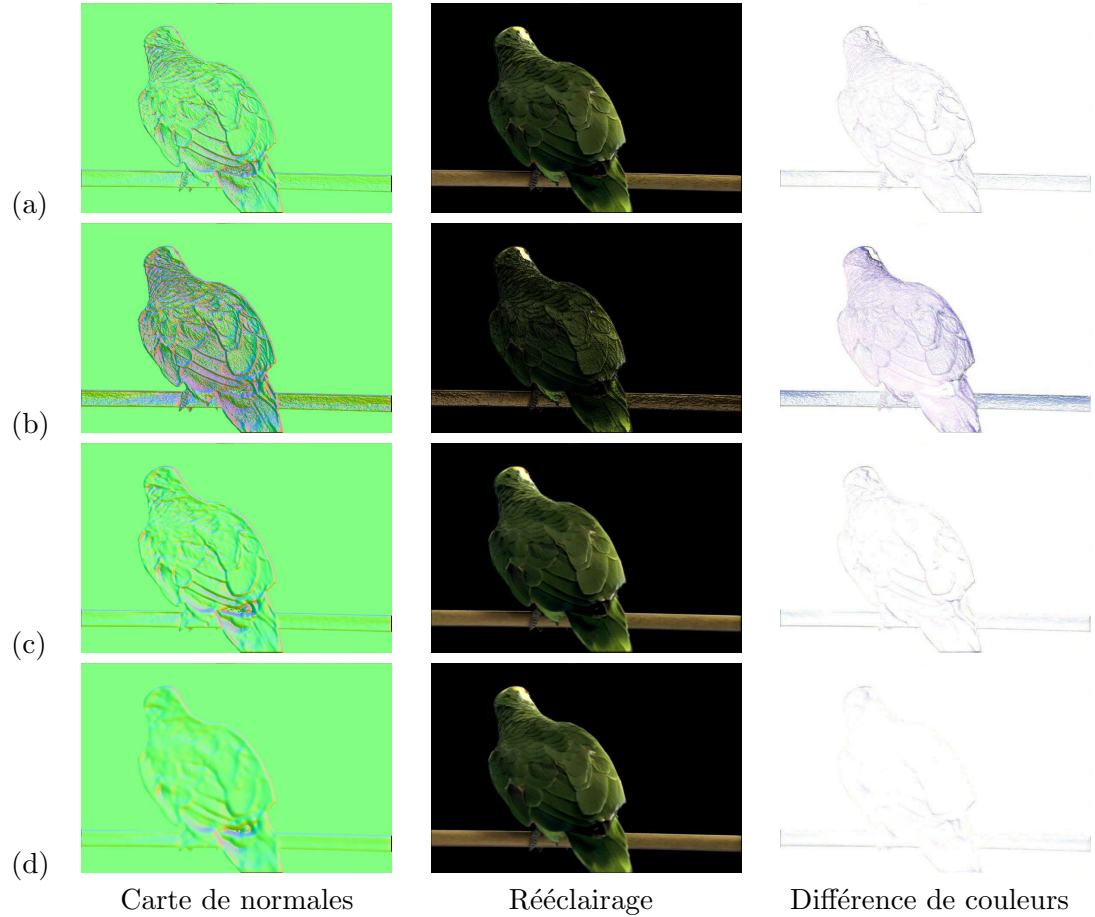


FIG. 5.3 – Reconstruction à partir du dégradé d’illumination : (a) mise à l’échelle (0.5), (b) mise à l’échelle (0.05), (c) filtrage gaussien de rayon 5 sur (a), (d) filtrage gaussien de rayon 25 sur (a)

5.2 Rééclairage virtuel

Une fois la carte de normales reconstruite, le défi consiste à rééclairer l’image de façon réaliste avec des sources de lumière virtuelles. Si la carte de normales est une bonne estimation de la surface, un effet de profondeur devrait apparaître lorsque la source de lumière est déplacée. Parfois, l’effet d’éclairage est trop intense pour être réaliste. Dans ce cas, l’opérateur de dissolution (voir la section 1.3.2) permet de doser le résultat en générant une image intermédiaire entre le résultat et l’image originale.

À la figure 5.5, le *chien* est rééclairé grâce à une carte de normales générée par notre outil et une source de lumière virtuelle positionnée selon quatre directions différentes.

Pour démontrer l’avantage de notre outil dans une composition, nous avons intégré l’image du *chien* sur un arrière-plan où la direction de la lumière est facilement per-

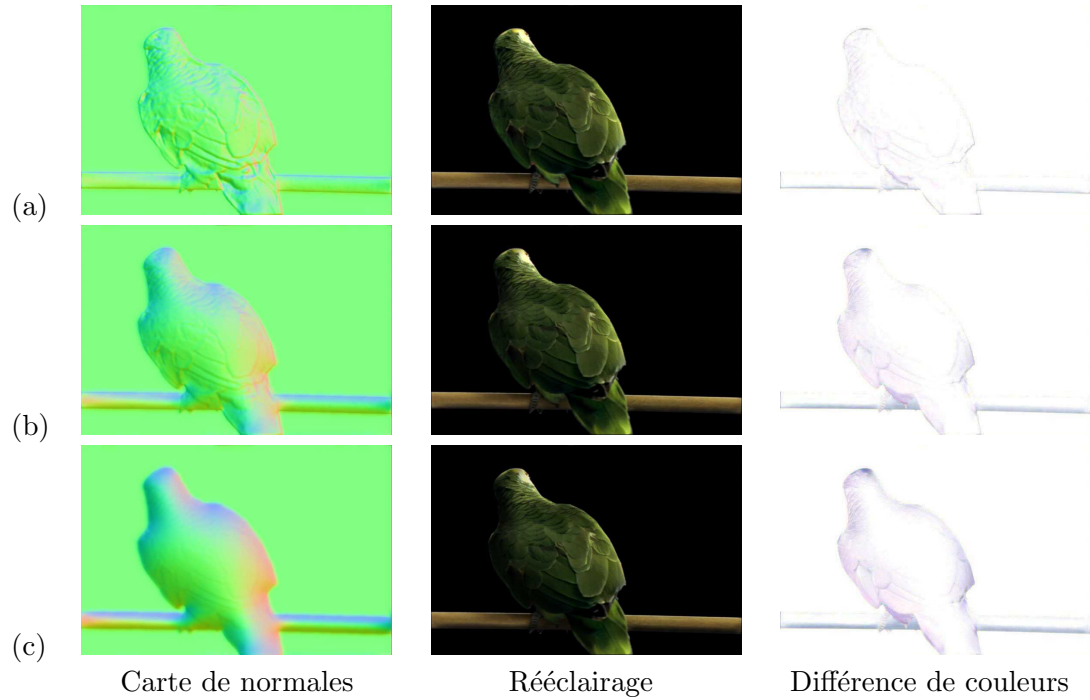


FIG. 5.4 – Combinaison des cartes de normales (ratio à partir des silhouettes : à partir du dégradé d'illumination) : (a) 25% : 75%, (b) 50% : 50%, (c) 75% : 25%

ceptible (voir la figure 5.6). Il s'agit là d'un cas problématique en composition vidéo numérique. Même si le *chien* n'a pas été filmé dans les mêmes conditions d'éclairage que l'arrière-plan, nous simulons des conditions similaires. En ajoutant une fausse ombre (voir la section 2.2.3) et en dosant l'ajout d'éclairage, nous arrivons à un résultat satisfaisant. La séquence vidéo complète est disponible sur notre site web [Ber06].

Pour illustrer l'effet de notre outil, un gros plan du *chien* est présenté avec et sans l'éclairage par une source de lumière virtuelle (voir la figure 5.6). Nous observons que l'ajout de cet effet aide à l'intégration de l'image d'avant-plan.

5.3 Problèmes de robustesse

Notre outil de génération de cartes de normales permet d'ajouter des effets d'éclairage directionnels. Pour tirer profit de l'outil au maximum, il est nécessaire de connaître les faiblesses de chacune des stratégies de reconstruction.

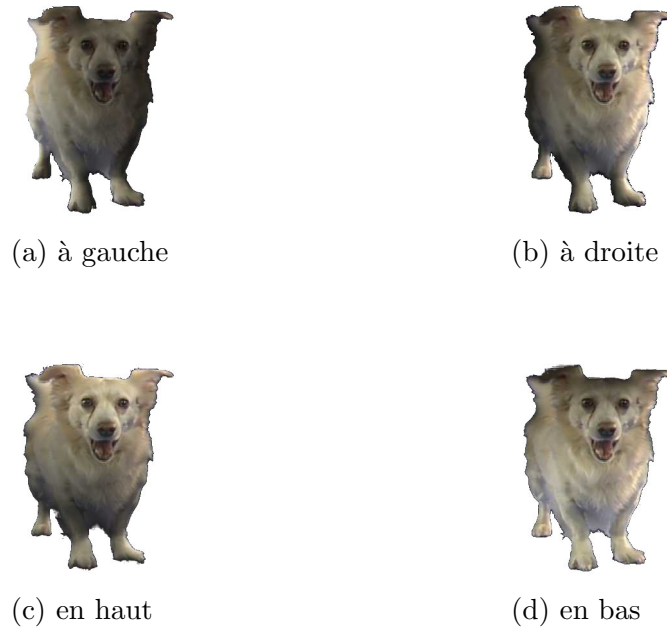


FIG. 5.5 – Rééclairage avec une source de lumière virtuelle positionnée selon quatre directions différentes

5.3.1 Reconstruction à partir du matte

Grâce à notre technique de reconstruction à partir du matte, la normale d'un point est déterminée en analysant les distances entre le point et les silhouettes. Les parties des silhouettes près de ce point ont une influence plus importante sur la normale calculée que les parties éloignées. Le comportement des silhouettes en fonction du temps est problématique, ainsi que les régions près des bordures du matte.

Incohérence temporelle

Le grand problème de l'approche de reconstruction à partir des silhouettes est l'incohérence temporelle des silhouettes causée par les occlusions. Avec notre technique, lorsqu'un objet passe devant un autre, les silhouettes sont fusionnées. Les normales sont alors déterminées à partir de silhouettes complètement différentes d'un matte au suivant, causant ainsi des fluctuations inattendues d'éclairage. La figure 5.7 illustre un rectangle et une sphère obstrués par une sphère. Nous voyons que les normales reconstruites ne sont pas cohérentes avant et après l'occlusion, ce qui change l'illumination ajoutée sur les primitives. Des exemples sur des séquences réelles sont présentés sur le web [Ber06].

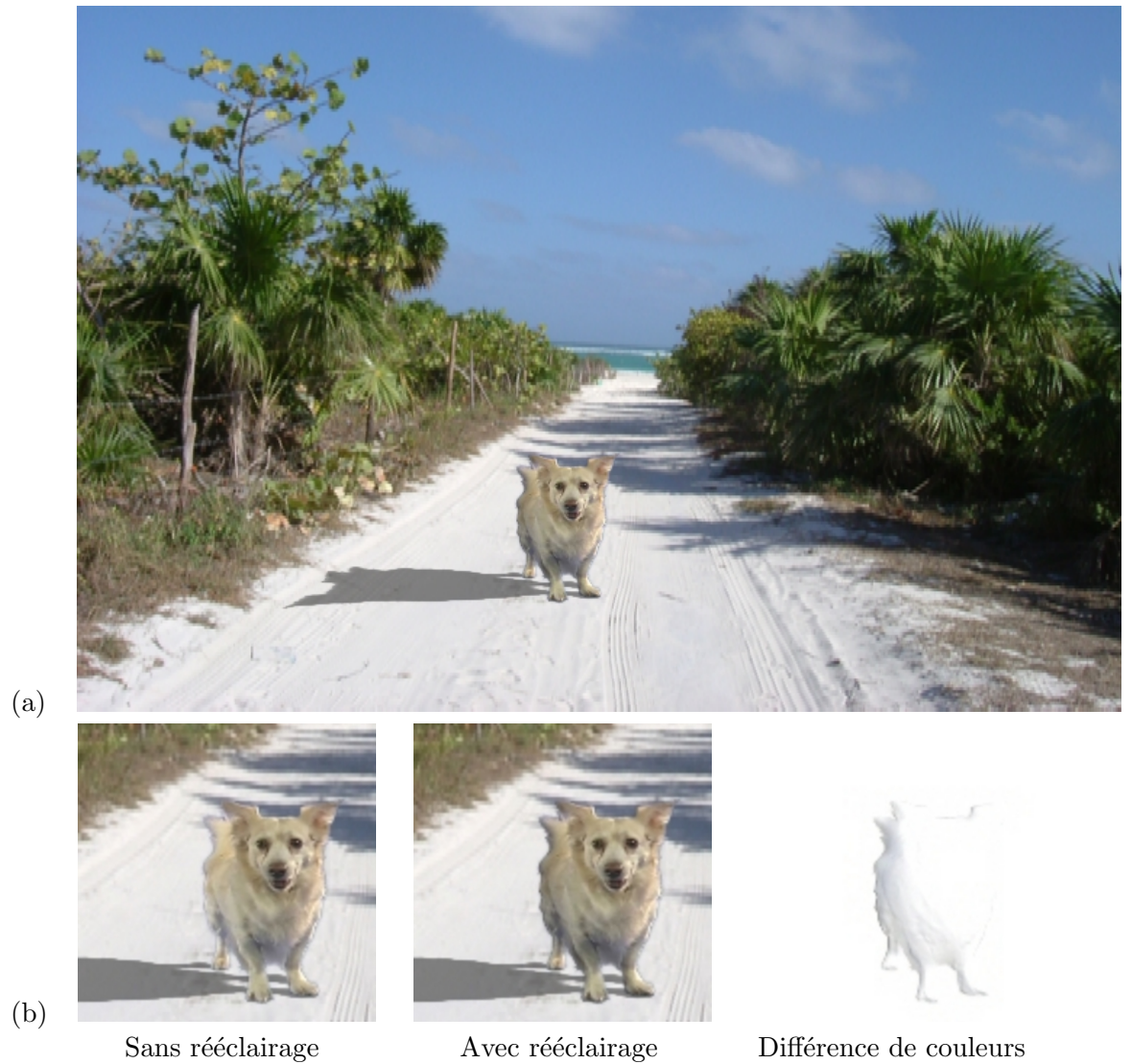


FIG. 5.6 – Composition exploitant notre outil : (a) composition avec rééclairage, (b) comparaison avec et sans éclairage

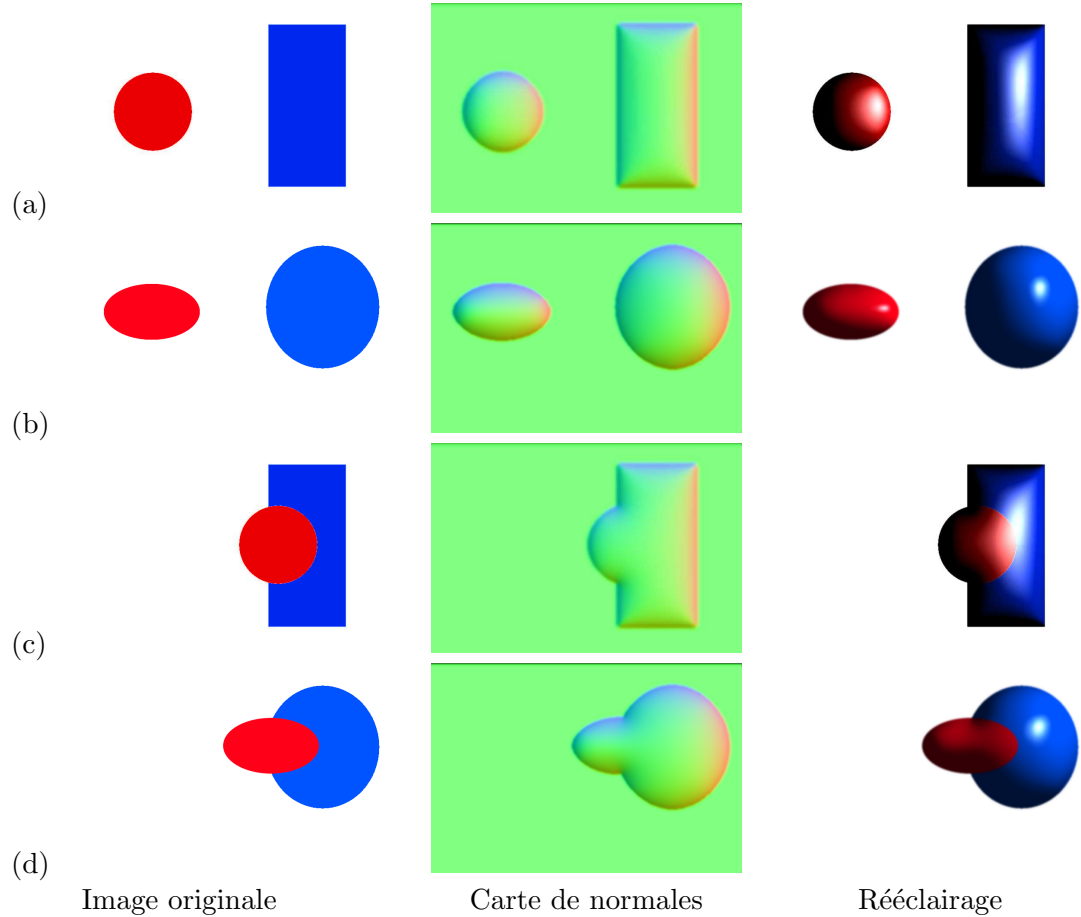


FIG. 5.7 – Incohérence temporelle due à l’occlusion : (a, b) sans occlusion, (c, d) occlusion

Bordures du matte

Lorsqu’un point est situé près de la bordure du matte, un problème de continuité survient puisque les silhouettes complètes de l’objet n’apparaissent pas dans l’image. En effet, une partie de celles-ci repose à l’extérieur du matte. Une bordure artificielle peut être créée par la bordure de l’image en ajoutant une bande noire autour du matte. À la figure 5.8(a), l’effet dû au rééclairage à l’intersection du perchoir et de la bordure de l’image apparaît erroné. Ceci est dû au fait que le prolongement des silhouettes du bâton se trouvant à l’extérieur du matte n’est pas pris en compte dans le calcul des normales. L’erreur est beaucoup moins perceptible sur la queue du *perroquet* puisque cette surface est moins régulière.

5.3.2 Reconstruction à partir du dégradé d'illumination

Grâce à notre technique de reconstruction à partir du dégradé d'illumination, la normale d'un point est déterminée en analysant le gradient d'intensité lumineuse à ce point. L'algorithme suppose que les variations d'intensité lumineuse sont causées par la courbure de la surface. Des problèmes surviennent lorsque d'autres phénomènes causent ces variations. C'est le cas des ombres, des reflets spéculaires et des transitions de couleur.

Ombres et reflets spéculaires

Lorsqu'un objet projette une ombre dans une image, le gradient d'intensité lumineuse de la région dans l'ombre est faible contrairement au gradient très élevé de la région du contour de l'ombre (passage d'une région foncée à claire). Dans le cas des reflets spéculaires, le même phénomène se produit à l'exception du fait que le gradient est inversé sur le contour (passage d'une région claire à foncée). Dans ces situations, puisque le gradient d'intensité lumineuse est affecté par d'autres phénomènes que la courbure de la surface, la normale reconstruite est erronée et des problèmes apparaissent lors du rééclairage. Puisque ces erreurs sont importantes, l'application d'un filtre empire le problème puisque les erreurs sont propagées dans les régions voisines. Les artéfacts d'ombres sont illustrés à la figure 5.8 (b).

Transitions de couleur

Pour calculer le gradient d'intensité lumineuse, l'image couleur est convertie en image monochrome. Les différences d'intensité, résultantes de la conversion de différentes couleurs, créent des discontinuités sur le gradient. Ainsi, aux endroits où des transitions de couleur apparaissent, la normale reconstruite sera biaisée en direction de la couleur la plus sombre. Le phénomène est aussi illustré à la figure 5.8 (b).

5.4 Analyse de la performance

Les résultats générés dans ce chapitre ont été obtenus grâce à notre outil de génération de cartes de normales implémenté dans une version préliminaire du logiciel *Discreet Flame* 10 [ME06]. La station de travail *Linux* utilisée (*Redhat Enterprise 3 [Lin06]*)

possède deux processeurs *AMD Opteron 244 1800 MHz*, 2 Giga-octets de mémoire vive et une carte graphique *nVidia Quadro FX 3000*. L'implémentation exploite le parallélisme pour un nombre arbitraire de processeurs, chacun effectuant la même tâche.

Les temps de génération de cartes de normales sont donnés au tableau 5.1. L'image du *perroquet* (voir la section 5.1) a été redimensionnée à différentes résolutions pour calculer la performance avec la même image de référence. Les temps de chacune des étapes de reconstruction ont été étudiés : la reconstruction à partir des silhouettes du matte, la reconstruction à partir du dégradé d'illumination et la combinaison des deux techniques. Le temps total de reconstruction ne tient pas compte du filtrage ou de toute autre manipulation pouvant être effectuée pour corriger le résultat final. Pour procéder au rééclairage, 0.2 secondes ont été nécessaires pour l'ensemble des résolutions d'images. Du point de vue de l'artiste, le temps requis pour déterminer les meilleurs paramètres de génération de normales est inférieur à une minute par image, selon notre expérience.

Résolution	Reconstruction (Silhouettes)	Reconstruction (Illumination)	Combinaison	Total
360×243	0.66	0.01	0.01	0.68
720×486	0.86	0.03	0.01	0.90
1280×720	1.02	0.12	0.06	1.18
1920×1440	1.29	0.25	0.15	1.69
4760×3988	2.28	1.47	0.90	4.75

TAB. 5.1 – Temps d'exécution (en secondes) en fonction de la résolution de l'image

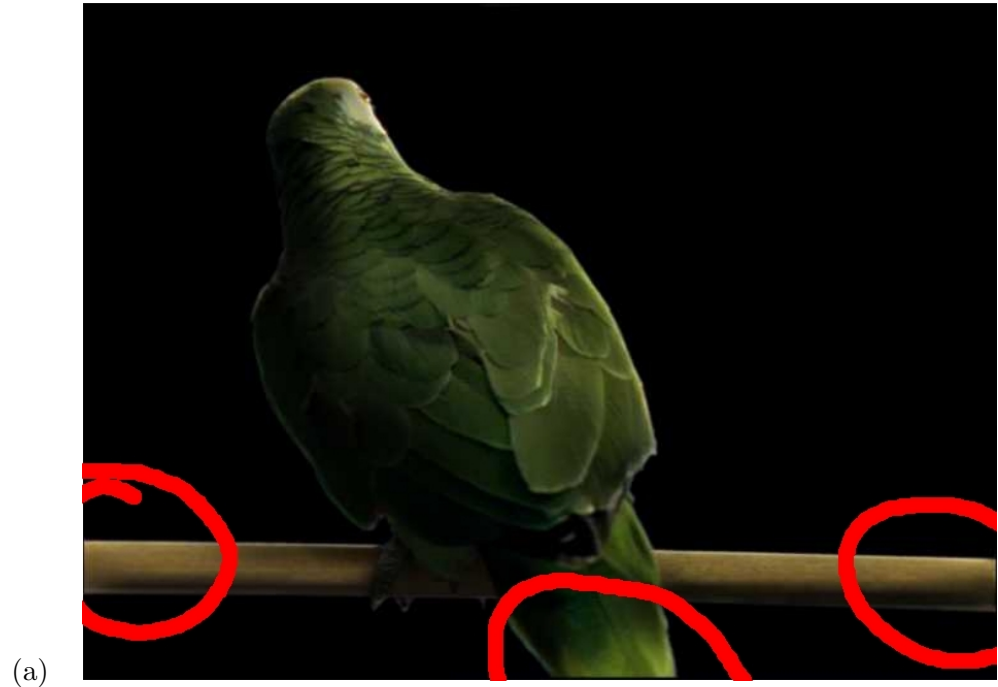


FIG. 5.8 – Différents problèmes de robustesse de notre outil : (a) problèmes près des bordures, (b) problèmes d'ombres (rouge) et de transitions de couleur (bleu), (c et d) images originales

Conclusion

Afin de créer des effets spéciaux ou de réduire les coûts du tournage, la composition vidéo numérique est devenue une technique importante dans l'industrie du cinéma. Ainsi, des séquences sont combinées dans le but de donner au public l'illusion que la séquence résultante provient d'un seul tournage. De nos jours, ce trucage est réalisé en post-production, ce qui n'était pas le cas dans les débuts du cinéma. Grâce aux techniques d'extraction de mattes, il est facile d'intégrer une séquence en filmant les objets d'avant-plan devant un écran bleu.

Actuellement, pour réaliser des compositions de qualité, il est primordial d'avoir un éclairage en studio qui correspond avec celui de l'arrière-plan. À cette fin, l'ajustement de l'éclairage est un processus fastidieux qui occupe la majeure partie du tournage. Une fois le tournage terminé, des outils de correction de couleur sont utilisés pour raffiner l'éclairage entre les séquences. Ainsi, la luminance, le contraste et la teinte sont modifiés en exploitant les différentes représentations des couleurs par l'ordinateur. Les coloristes doivent tenir compte de notre perception des couleurs ainsi que des effets d'ambiance évoqués par celles-ci. Le réalisme est parfois difficile à atteindre puisque ces techniques ne tiennent pas compte du positionnement des sources de lumière ni des propriétés de réflectance des surfaces.

Dans les tournages à grand déploiement, grâce aux techniques par rendu, il est possible de refaire l'éclairage avec des sources de lumière virtuelles arbitraires (souvent des cartes de radiance extraites à partir de sondes d'illumination). Pour y parvenir, des changements considérables doivent être réalisés à l'étape du tournage. La géométrie de la scène est reconstruite avec des systèmes à plusieurs caméras, des numériseurs 3D ou des projections de lumière. Certaines techniques permettent aussi d'extraire les propriétés de réflectance des surfaces à partir d'images, de façon à ce que le rendu synthétique soit encore plus réaliste. L'accroissement considérable des coûts de mise en scène engendrés

par ces techniques explique leur faible utilisation.

Après avoir résumé l'état actuel des techniques de rééclairage en post-production, nous concluons qu'une technique intermédiaire serait bénéfique pour l'industrie de façon à faciliter la transition vers les techniques par rendu. Dans cette optique, nous avons développé un outil de reconstruction interactif à partir d'une séquence vidéo, qui ne requiert aucune modification aux étapes de tournage réalisées devant l'écran bleu. Grâce à la surface courbe "plausible" obtenue, il est possible de rééclairer la séquence d'avant-plan avec des sources de lumière virtuelles. L'implémentation a notamment été réalisée à l'intérieur d'un logiciel de composition vidéo numérique professionnel.

L'idée originale de combiner les techniques de reconstruction par gonflement de silhouettes et par dégradé d'illumination a permis le développement d'un outil de reconstruction rapide, flexible et intuitif qui s'insère facilement dans le "pipeline" de post-production actuel. L'introduction des cartes de normales en composition vidéo numérique aide à réaliser des effets d'éclairage en temps réel, ce qui en fait une représentation bien adaptée pour la géométrie générée par notre système. Le tout a permis de réaliser une preuve de concept que la reconstruction d'une surface "plausible" à partir d'une séquence vidéo permet d'ajouter des effets d'éclairage réalistes. Dans le cas où la différence d'éclairage entre deux séquences est notable, notre outil s'avère un choix judicieux pour ajouter un effet d'éclairage directionnel. L'effet serait impossible à réaliser avec les outils traditionnels de correction de couleur.

Perspectives futures

En s'appuyant sur le talent des artistes plutôt que sur des techniques automatiques moins flexibles, nous croyons que le développement de nouveaux outils puissants permettrait de résoudre d'autres problèmes complexes. En conservant cette philosophie, nous sommes enthousiastes quant aux futures applications de notre outil.

Améliorations

Afin d'améliorer la robustesse de notre outil, il serait bénéfique de corriger certains problèmes causant des artefacts tels que les occlusions, les transitions de couleur, les reflets spéculaires et les régions dans l'ombre.

Le suivi des silhouettes pourrait contrer en partie l'incohérence temporelle causée

par les occlusions. Lorsque cette situation survient, les silhouettes des deux régions où l'intersection a lieu disparaissent. À ce moment, les parties visibles des silhouettes pourraient être déduites par une détection de contours sur l'image et les parties des silhouettes cachées pourraient être extrapolées grâce à une estimation par flux optique [HS80]. Chaque région utiliserait les silhouettes correspondantes dans le calcul de ses normales. L'extrapolation pourrait aussi être utilisée lorsque les silhouettes sortent du cadre de l'image.

L'intégration d'une technique pour dissocier la texture de l'illumination permettrait d'accroître la robustesse de notre système. En effet, les ombres, les reflets spéculaires et les transitions de couleur sont problématiques avec notre algorithme de reconstruction de surfaces par dégradé d'illumination. Ainsi, il serait possible d'enlever certaines ombres sur les images grâce à un filtrage bilatéral [OCDD01]. Par la suite, en convertissant les couleurs dans l'espace de *chromaticité-intensité inverse*, il serait possible d'enrayer plusieurs reflets spéculaires et de séparer l'illumination de la texture [TNI03]. Nous croyons que ces techniques aideraient à retrouver des normales plus précises et à atténuer certains artefacts. Puisqu'elles ne sont démontrées que sur des images statiques, nous ignorons le comportement de ces techniques sur des séquences d'images. Malheureusement, ces techniques feraient grimper le temps de calcul au-dessus d'une minute par image sur des plateformes, ce qui va à l'encontre de la philosophie du système. L'exploitation des nuanceurs de pixels¹ programmables semble très appropriée dans un tel contexte.

Extensions

La génération de géométrie à partir d'une séquence vidéo est un problème complexe et impossible à résoudre de façon générale. Bien que l'ajustement des paramètres de reconstruction de notre outil offre à l'utilisateur une certaine flexibilité, la reconstruction de certains types de surface demeure impossible à réaliser. Pour plus de généralité, le support pour les objets polygonaux serait essentiel. Il existe déjà plusieurs techniques de reconstruction mieux adaptées aux polygones (*Façade* [DTM96], *Rekon* [POF98], *Reality* [DGSR01], etc.). Intégrer ces techniques dans notre outil tout en exploitant les informations supplémentaires dues aux séquences animées des mattes pourrait nous

¹*pixel shaders*

permettre de mieux reconstruire les scènes d'intérieur.

Présentement, la fonction de réflectance est constante sur toute l'image. Cette limitation nous empêche de réaliser des effets complexes de rééclairage. Nous croyons qu'un système similaire au nôtre pourrait être développé afin d'extraire les paramètres de réflectance des surfaces. En se basant sur plusieurs modèles de réflectance et en exploitant certains indices présents dans les images, nous espérons étendre notre système rapide et intuitif de rééclairage en tenant compte de la géométrie et des paramètres de réflectance des surfaces. Des effets d'éclairage plus réalistes en découleront, ainsi que de nouveaux effets artistiques.

Le rééclairage n'est pas le seul domaine en composition vidéo numérique qui peut bénéficier d'une reconstruction "plausible" d'objets réels. Par exemple, un système de particules [Ree83] pourrait tirer avantage de la géométrie grossière pour simuler une variété d'effets tridimensionnels tels que l'inclusion de pluie ou de fumée dans les scènes. Des effets de réfraction pourraient également être réalisés en temps réel à partir de nos cartes de normales [Wym05].

Bibliographie

- [Ali06] Alias. www.alias.com, janvier 2006.
- [Ana06] « Analyse filmique ». analysefilmique.free.fr, janvier 2006.
- [Ber06] Nicolas Bergeron.
www.iro.umontreal.ca/labs/infographie/theses/bergerni, avril 2006.
- [BG01] Samuel Boivin et Andre Gagalowicz. « Image-based rendering of diffuse, specular and glossy surfaces from a single image ». Dans *SIGGRAPH '01 : Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, pages 107–116. ACM Press, 2001.
- [BGH05] David Burke, Abhijeet Ghosh et Wolfgang Heidrich. « Bidirectional importance sampling for direct illumination. ». Dans *Proceedings of the Eurographics Symposium on Rendering Techniques, Konstanz, Germany, June 29 - July 1, 2005*, pages 147–156, 2005.
- [Bli77] James F. Blinn. « Models of light reflection for computer synthesized pictures ». Dans *SIGGRAPH '77 : Proceedings of the 4th annual conference on Computer graphics and interactive techniques*, pages 192–198. ACM Press, 1977.
- [Bli78] James F. Blinn. « Simulation of wrinkled surfaces ». Dans *SIGGRAPH '78 : Proceedings of the 5th annual conference on Computer graphics and interactive techniques*, pages 286–292. ACM Press, 1978.
- [Bri99] Ron Brinkmann. *The art and science of digital compositing*. Morgan Kaufmann Publishers Inc., 1999.
- [Cin06] CinePaint. www.cinepaint.org, janvier 2006.
- [CJAMJ05] Petrik Clarberg, Wojciech Jarosz, Tomas Akenine-Möller et Henrik Wann Jensen. « Wavelet importance sampling : efficiently evaluating products

- of complex functions ». Dans *SIGGRAPH '05 : Proceedings of the 32th annual conference on Computer graphics and interactive techniques*, pages 1166–1175. ACM Press, 2005.
- [Coo84] Robert L. Cook. « Shade trees ». Dans *SIGGRAPH '84 : Proceedings of the 11th annual conference on Computer graphics and interactive techniques*, pages 223–231. ACM Press, 1984.
- [Cro77] Frank C. Crow. « Shadow algorithms for computer graphics ». Dans *SIGGRAPH '77 : Proceedings of the 4th annual conference on Computer graphics and interactive techniques*, pages 242–247. ACM Press, 1977.
- [Cyb06] Cyberware. www.cyberware.com, janvier 2006.
- [Deb98] Paul Debevec. « Rendering synthetic objects into real scenes : bridging traditional and image-based graphics with global illumination and high dynamic range photography ». Dans *SIGGRAPH '98 : Proceedings of the 25th annual conference on Computer graphics and interactive techniques*, pages 189–198. ACM Press, 1998.
- [Deb06] Paul Debevec. www.debevec.org/Probes, janvier 2006.
- [DGSR01] Sebastien Dedieu, Pascal Guitton, Christophe Schlick et Patrick Reuter. « Reality : An interactive reconstruction tool of 3D objects from photographs ». Dans *Vision, Modeling, and Visualization 2001*, pages 195–202, novembre 2001.
- [DHT⁺00] Paul Debevec, Tim Hawkins, Chris Tchou, Haarm-Pieter Duiker, Westley Sarokin et Mark Sagar. « Acquiring the reflectance field of a human face ». Dans *SIGGRAPH '00 : Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, pages 145–156. ACM Press/Addison-Wesley Publishing Co., 2000.
- [DM97] Paul Debevec et Jitendra Malik. « Recovering high dynamic range radiance maps from photographs ». Dans *SIGGRAPH '97 : Proceedings of the 24th annual conference on Computer graphics and interactive techniques*, pages 369–378. ACM Press/Addison-Wesley Publishing Co., 1997.
- [DTM96] Paul Debevec, Camillo Taylor et Jitendra Malik. « Modeling and rendering architecture from photographs : a hybrid geometry- and image-based

- approach ». Dans *SIGGRAPH '96 : Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, pages 11–20. ACM Press, 1996.
- [FH04] Hui Fang et John C. Hart. « Textureshop : texture synthesis as a photograph editing tool ». Dans *SIGGRAPH '04 : Proceedings of the 31th annual conference on Computer graphics and interactive techniques*, pages 354–359. ACM Press, 2004.
- [FLW02] Raanan Fattal, Dani Lischinski et Michael Werman. « Gradient domain high dynamic range compression ». Dans *SIGGRAPH '02 : Proceedings of the 29th annual conference on Computer graphics and interactive techniques*, pages 249–256. ACM Press, 2002.
- [FvDFH96] James D. Foley, Andries van Dam, Steven K. Feiner et John F. Hughes. *Computer graphics (2nd ed. in C) : principles and practice*. Addison-Wesley Longman Publishing Co., Inc., 1996.
- [FZGH05] Hui Fang, Steve Zelinka, Michael Garland et John C. Hart. « Interactive material replacement in photographs ». Dans *GI '05 : Proceedings of the 2005 conference on Graphics interface*, pages 227–232. Canadian Human-Computer Communications Society, 2005.
- [GPEP04] Martin Granger-Piché, Emric Epstein et Pierre Poulin. « Interactive hierarchical space carving with projector-based calibrations ». Dans *Vision, Modeling, and Visualization 2004*, pages 159–166, novembre 2004.
- [GW01] Rafael C. Gonzalez et Richard E. Woods. « digital image processing ». pages 401–441. Addison-Wesley Longman Publishing Co., Inc., 2001.
- [HEH05] Derek Hoiem, Alexei A. Efros et Martial Hebert. « Automatic photo pop-up ». Dans *SIGGRAPH '05 : Proceedings of the 32th annual conference on Computer graphics and interactive techniques*, pages 577–584. ACM Press, 2005.
- [Hor90] Berthold K. P. Horn. « Height and gradient from shading ». volume 5, pages 37–75. Kluwer Academic Publishers, 1990.
- [HS80] Berthold K.P. Horn et Brian G. Schunck. « Determining optical flow ». Rapport technique, 1980.

- [HWT⁺04] Tim Hawkins, Andreas Wenger, Chris Tchou, Andrew Gardner, Fredrik Göransson et Paul E. Debevec. « Animatable facial reflectance fields ». Dans *Proceedings of Eurographics Workshop on Rendering 2004*, pages 309–321, 2004.
- [IMT99] Takeo Igarashi, Satoshi Matsuoka et Hidehiko Tanaka. « Teddy : a sketching interface for 3D freeform design ». Dans *SIGGRAPH '99 : Proceedings of the 26th annual conference on Computer graphics and interactive techniques*, pages 409–416. ACM Press/Addison-Wesley Publishing Co., 1999.
- [Inv06] « Invention and technology ». inventionandtechnology.com, janvier 2006.
- [Kaj86] James T. Kajiya. « The rendering equation ». Dans *SIGGRAPH '86 : Proceedings of the 13th annual conference on Computer graphics and interactive techniques*, pages 143–150. ACM Press, 1986.
- [KS00] Kiriakos N. Kutulakos et Steven M. Seitz. « A theory of shape by space carving ». Dans *International Journal of Computer Vision*, volume 38, pages 199–218. Kluwer Academic Publishers, 2000.
- [LFD⁺99] Celine Loscos, Marie-Claude Frasson, George Drettakis, Bruce Walter, Xavier Granier et Pierre Poulin. « Interactive virtual relighting and remodeling of real scenes ». Dans *Proceedings of Eurographics Workshop on Rendering 99*, juin 1999.
- [Lin06] Red Hat Linux. www.redhat.com, janvier 2006.
- [LPC⁺00] Marc Levoy, Kari Pulli, Brian Curless, Szymon Rusinkiewicz, David Koller, Lucas Pereira, Matt Ginzton, Sean Anderson, James Davis, Jeremy Ginsberg, Jonathan Shade et Duane Fulk. « The digital Michelangelo project : 3D scanning of large statues ». Dans *SIGGRAPH '00 : Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, pages 131–144. ACM Press/Addison-Wesley Publishing Co., 2000.
- [ME06] Autodesk Media et Entertainment. www.discreet.com, janvier 2006.
- [MG97] Stephen R. Marschner et Donald P. Greenberg. « Inverse lighting for photography ». Dans *Proceedings of the Fifth Color Imaging Conference, Society for Imaging Science and Technology*, 1997.

- [Mig05] Max Mignotte. *IFT 6150 : Digital image processing, image enhancement by point operation course*. www.iro.umontreal.ca/~mignotte/ift6150.html, 2005.
- [MP79] D. Marr et T. Poggio. « A computational theory of human stereo vision ». Dans *Proceedings of the Royal Society of London*, volume B 204, pages 301–328, mai 1979.
- [MPC⁺05] Marcus Magnor, Marc Pollefeys, German Cheung, Wojciech Matusik et Christian Theobalt. « Video based rendering ». SIGGRAPH 2005 course notes, 2005.
- [NM03] Marla Newall et David Meny. « 3D performance matching for Terminator 3 ». Dans *Proceedings of the SIGGRAPH '03 conference on Sketches & applications*, 2003.
- [NN04] Ko Nishino et Shree K. Nayar. « Eyes for relighting ». Dans *SIGGRAPH '04 : Proceedings of the 31th annual conference on Computer graphics and interactive techniques*, pages 704–711. ACM Press, 2004.
- [OCDD01] Byong Mok Oh, Max Chen, Julie Dorsey et Frédo Durand. « Image-based modeling and photo editing ». Dans *SIGGRAPH '01 : Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, pages 433–442. ACM Press, 2001.
- [ODJ04] Victor Ostromoukhov, Charles Donohue et Pierre-Marc Jodoin. « Fast hierarchical importance sampling with blue noise properties ». Dans *SIGGRAPH '04 : Proceedings of the 31th annual conference on Computer graphics and interactive techniques*, pages 488–495. ACM Press, 2004.
- [Owe99] G. Scott Owen.
www.siggraph.org/education/materials/HyperGraph/color/color0.htm, juin 1999.
- [PD84] Thomas Porter et Tom Duff. « Compositing digital images ». Dans *SIGGRAPH '84 : Proceedings of the 11th annual conference on Computer graphics and interactive techniques*, pages 253–259. ACM Press, 1984.
- [Pfe72] Henri Pfeiffer. *L'harmonie des couleurs, cours théorique et pratique*. Dunod, 1972.

- [Pho75] Bui Tuong Phong. « Illumination for computer generated pictures ». *Communication of the ACM*, volume 18, numéro 6, pages 311–317, 1975.
- [POF98] Pierre Poulin, Mathieu Ouimet et Marie-Claude Frasson. « Interactively modeling with photogrammetry ». Dans *Proceedings of Eurographics Workshop on Rendering 98*, pages 93–104, juin 1998.
- [RCM⁺01] C. Rocchini, P. Cignoni, C. Montani, P. Pingi et R. Scopigno. « A low cost 3D scanner based on structured light ». *Computer Graphics Forum*, volume 20, numéro 3, 2001.
- [Ree83] W. T. Reeves. « Particle Systems – a technique for modeling a class of fuzzy objects ». Dans *SIGGRAPH '83 : Proceedings of the 10th annual conference on Computer graphics and interactive techniques*, pages 253–259. ACM Press, 1983.
- [Ric00] Richard Rickitt. *Special effects : the history and technique*. Watson-Guptill Publications, 2000.
- [Saw99] Lesa Sawahata. *Color harmony workbook*. Rockport Pub, 1999.
- [SB96] Alvy Ray Smith et James F. Blinn. « Blue screen matting ». Dans *SIGGRAPH '96 : Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, pages 259–268. ACM Press, 1996.
- [SJTS04] Jian Sun, Jiaya Jia, Chi-Keung Tang et Heung-Yeung Shum. « Poisson matting ». Dans *SIGGRAPH '04 : Proceedings of the 31th annual conference on Computer graphics and interactive techniques*, pages 315–321. ACM Press, 2004.
- [Sof06] Softimage. www.softimage.com, janvier 2006.
- [SS02] Daniel Scharstein et Richard Szeliski. « A taxonomy and evaluation of dense two-frame stereo correspondence algorithms ». Dans *International Journal of Computer Vision*, volume 47, pages 7–42. Kluwer Academic Publishers, 2002.
- [TNI03] Robby T. Tan, Ko Nishino et Katsushi Ikeuchi. « Illumination chromaticity estimation using inverse-intensity chromaticity space. ». Dans *2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 673–682, 2003.

- [TV98] Emanuele Trucco et Alessandro Verri. *Introductory techniques for 3D computer vision*. Prentice Hall, 1998.
- [WGT⁺05] Andreas Wenger, Andrew Gardner, Chris Tchou, Jonas Unger, Tim Hawkins et Paul Debevec. « Performance relighting and reflectance transformation with time-multiplexed illumination ». Dans *SIGGRAPH '05 : Proceedings of the 32th annual conference on Computer graphics and interactive techniques*, pages 756–764. ACM Press, 2005.
- [Wik06] Wikipedia. www.wikipedia.org, janvier 2006.
- [Wil78] Lance Williams. « Casting curved shadows on curved surfaces ». Dans *SIGGRAPH '78 : Proceedings of the 5th annual conference on Computer graphics and interactive techniques*, pages 270–274. ACM Press, 1978.
- [Wil91] Lance Williams. « Shading in two dimensions ». Dans *GI '91 : Proceedings of the 1991 conference on Graphics interface*, pages 143–151, 1991.
- [WP03] Guy Williams et Dana Peters. « Making the treebeard's beard for "The lord of the rings" ». Dans *Proceedings of the SIGGRAPH '03 conference on Sketches & applications*, 2003.
- [Wri01] Steve Wright. *Digital compositing for film and video with cdrom*. Butterworth-Heinemann, 2001.
- [Wym05] Chris Wyman. « An approximate image-space approach for interactive refraction ». Dans *SIGGRAPH '05 : Proceedings of the 32th annual conference on Computer graphics and interactive techniques*, pages 1050–1053. ACM Press, 2005.