

Vers plus de contrôle pour le créateur d'images de synthèse

Rémi Cozot

▶ **To cite this version:**

Rémi Cozot. Vers plus de contrôle pour le créateur d'images de synthèse. Synthèse d'image et réalité virtuelle [cs.GR]. Rennes 1, 2014. tel-01141846

HAL Id: tel-01141846
https://inria.hal.science/tel-01141846

Submitted on 14 Apr 2015

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

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

HABILITATION À DIRIGER DES RECHERCHES

présentée devant

L'Université de Rennes 1
Spécialité : informatique

par

Rémi Cozot

Vers plus de contrôle pour le créateur d'images de synthèse

soutenue le 8 septembre 2014 devant le jury composé de :

Pr. Éric Marchand - Université de Rennes 1 - Président
Pr. Touradj Ebrahimi - École Polytechnique Fédérale de Lausanne - Rapporteur
Pr. Céline Loscos - Université de Reims Champagne-Ardenne - Rapporteur
Pr. Pascal Guitton - Université de Bordeaux - Rapporteur
Pr. Patrick Le Callet - Université de Nantes
Pr. Jean-Marc Jézéquel - Université de Rennes 1

Table des matières

1 Avant Propos **5**
1.1 Remerciements 5

2 Introduction **6**
2.1 Contexte 6
2.2 Domaines d'application 8
2.3 Enjeux 9
2.4 Contributions 12
2.4.1 Modélisation automatique de scènes 3D 12
2.4.2 Rendu des milieux participatifs en illumination globale 12
2.4.3 Modèle d'attention visuelle et ses applications 13
2.4.4 Cohérences temporelles dans la réduction des gammes de luminances 13
2.4.5 Estimation du blanc de référence en illumination globale . . . 14
2.5 Structure du document 15

3 De la lumière à l'esthétique de l'image **16**
3.1 Introduction 16
3.2 La lumière : de la luminance à la colorimétrie 17
3.2.1 Luminance et Équation de rendu 17
3.2.2 Photométrie et colorimétrie 18
3.2.3 Perception et apparence 20
3.3 Dessiner avec la lumière : une intention 23
3.3.1 Composition et cadrage 23
3.3.2 Lumière 24
3.3.3 Couleur 26
3.3.4 Netteté et profondeur de champ 28
3.4 Éléments de contrôle 28

4 Modélisation automatique de scènes 3D **32**
4.1 Introduction 32
4.2 Travaux précédants connexes 33
4.3 Environnement de création automatique de monde 3D 34
4.3.1 Vue d'ensemble 35

4.3.2 Calcul du génotype 35
4.3.3 Calcul du phénotype 37
4.4 Résultats 40
4.5 Conclusion et Perspectives 41

5 Rendu des milieux participatifs en illumination globale 45
5.1 Introduction 46
5.2 État de l'art 46
5.3 Architecture de l'algorithme 47
5.4 Implémentation 48
5.4.1 Pré-traitements 49
5.4.2 Émission des photons dirigée selon la visibilité 49
5.4.3 Mise à jour des *luminances* 50
5.4.4 Mise à jour de l'image 50
5.4.5 Mises à jours des rayons 51
5.5 Résultats 51
5.6 Conclusion 52

6 Post-traitements 57
6.1 Attention visuelle et applications 58
6.1.1 Introduction 58
6.1.2 Modèles d'attention visuelle précédents 59
6.1.3 Notre modèle d'attention visuelle 60
6.1.4 Résultat du modèle d'attention visuelle 62
6.1.5 Algorithme de flou de profondeur avec autofocus 63
6.1.6 Conclusion 65
6.2 Adaptation des gammes de luminances pour les séquences d'images . 65
6.2.1 Problématique de la réduction des gammes de luminances pour les vidéos 66
6.2.2 Travaux précédants connexes 68
6.2.3 Algorithmes de maintient de la cohérence temporelle 69
6.2.4 Résultats 73
6.2.5 Conclusion et perspectives 76
6.3 Gestion de la couleur 76
6.3.1 Notion de couleur d'adaptation 76
6.3.2 Estimation de la couleur d'adaptation 77
6.3.3 Notre méthode d'estimation de la couleur d'adaptation centrée sur l'oeil 80
6.3.4 Résultats 84
6.3.5 Conclusion sur l'estimation de la couleur d'adaptation 86

7 Discussion et Perspectives 88
7.1 Discussion 88
7.1.1 Modélisation automatique de scène 3D 88
7.1.2 Rendu des milieux participatifs en illumination globale 89
7.1.3 Attention visuelle dans les environnements 3D interactifs . . 90

7.1.4 Autofocus intelligent pour le flou de profondeur 90
7.1.5 Adaptation des gammes de luminance pour les séquences d'images 90
7.1.6 Détermination de la couleur d'adaptation 91
7.1.7 Analyse 91
7.2 Perspectives 92
7.2.1 Introduction 92
7.2.2 Vue d'ensemble 92
7.2.3 Description et évaluation de l'esthétique 94
7.2.4 Optimisation des paramètres 95
7.2.5 Rendu et des post-traitements 96
7.3 Conclusion 97

Références **98**

List of Figures **110**

List of Tables **114**

Chapitre 1

Avant Propos

Ce document présente un aperçu des travaux de recherche que j'ai menés et encadrés dans les équipes Bunraku puis FRVSense de l'IRISA en tant que maître de conférences à l'Université de Rennes 1. La période couverte dans ce document commence à mon retour de disponibilité pour création d'entreprise en 2005. Mes travaux qui précédent ma période de disponibilité ne sont pas traités dans ce document car ils portent sur des thématiques différentes (animation 3D par modèle physique). Mon retour dans le domaine de la recherche académique a coïncidé avec un changement de thèmes de recherche. Mes travaux actuels portent sur la synthèse d'images par ordinateur.

Remerciements

Je tiens à remercier chaleureusement Pr. Touradj Ebrahimi, Pr. Céline Loscos et Pr. Pascal Guitton d'avoir acceptés d'être rapporteurs de mon Habilitation à Diriger des Recherches. Je remercie Pr. Éric Marchand, Pr. Jean-Marc Jézéquel et Pr. Patrick Le Callet d'avoir accceptés de faire partie de mon jury d'Habilitation à Diriger des Recherches.

Je tiens à remercier très chaleureusement les doctorants que j'ai eu la chance d'encadrer dans leur thèse : Mlle Noémie Esnault, M. Sébastien Hillaire, M. Ronan Boitard, M. Adrien Gruson, M. Billal Merabti, M. Matis Hudon, M. Antonin Gilles et Mlle Hristina Hristova.

Je remercie également infiniment Pr. Kadi Bouatouch, responsable de l'équipe FRVSense de l'Institut de Recherche en Informatique et Systèmes Aléatoires (IRISA) pour m'avoir accueilli dans son équipe, pour m'avoir donné tous les moyens nécessaire à l'accomplissement des recherches présentés dans ce document, pour ses conseils avisés, pour son écoute et pour son amitiée.

Chapitre 2

Introduction

2.1 Contexte

La synthèse d'images dans son acception la plus large correspond à la création d'images numériques au moyen d'algorithmes. Nous nous intéressons ici à la synthèse d'images dite 3D, c'est à dire aux algorithmes qui calculent une image à partir d'une représentation géométrique tridimensionnelle d'une scène, appelée scène 3D. Habituellement ces algorithmes sont classés en deux grandes catégories : synthèse d'images photo-réaliste et synthèse non photo-réaliste. La synthèse d'images photo-réaliste correspond aux logiciels dont le but est de calculer une image dont l'apparence est proche d'une photographie. A l'opposée, la synthèse d'images non photo-réaliste cherche à calculer une image dont l'apparence est celle d'un dessin, d'une peinture, etc. Cette classification ne prend pas en compte l'algorithmique interne qui calcule l'image. À cette classification nous préférons définir la notion de *synthèse d'images physique* (*physically based rendering*) comme étant les algorithmes qui calculent une image d'une scène 3D en effectuant une simulation physique des interactions entre la scène et la lumière. Les travaux présentés ici se situent dans le contexte de la *synthèse d'images physique*.

Dans le cadre de la synthèse d'images physique, les éléments clés sont (Cf. Figure 2.1) :

1. la description de la scène 3D,
2. les caractéristiques de l'algorithme de synthèse d'images,
3. les images brutes de données physiques, résultats de la simulation,
4. les post-traitements convertissant les données physiques brutes en une image affichable,
5. l'image finale compatible avec les dispositifs de restitutions (écrans, imprimantes, etc.).

La description de la scène 3D comprend la définition de la géométrie des surfaces et des volumes, la caractérisation du comportement de la surface ou du volume vis

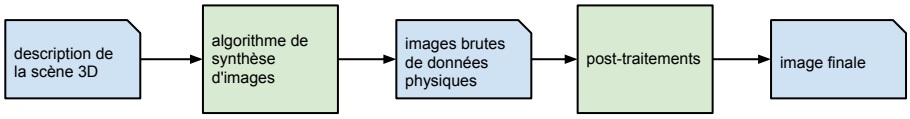


FIGURE 2.1 – Éléments clés de la synthèse d'images physique.

à vis de la lumière, le positionnement et les caractéristiques des sources de lumière et de la caméra. Les objets solides sont habituellement modélisés par la description de leur surface. Pour les objets gazeux, la modélisation se fait classiquement en spécifiant leur densité volumique dans l'espace. La description du matériaux définit leur comportement vis à vis de la lumière. Celui est généralement donné, pour les surfaces, par une fonction appelée *BRDF : Bidirectional Reflectance Distribution Function*. Elle exprime pour un ensemble de longueurs d'ondes, pour chaque direction incidente, le ratio de lumière réfléchie dans une direction donnée.

Le moteur de synthèse d'images physique se caractérise principalement par les interactions physiques entre la lumière et la matière qu'il est capable de simuler. D'un point de vue physique, l'éclairement d'un élément de surface provient, premièrement, de l'éclairement direct arrivant des sources de lumières, deuxièmement, de la réflexion de la lumière sur les autres surfaces de la scène et, troisièmement, de la lumière transmisse au travers des volumes. L'appellation *illumination directe* est utilisée dans le cas où le moteur de rendu ne prend en compte que la lumière venant directement des sources. Quand la réflexion de la lumière sur les surfaces est prise en compte, on parle alors d'*illumination globale*. Les moteurs d'*illumination globale* se différencient essentiellement par le nombre de réflexions de la lumière pris en compte, par le fait qu'ils simulent les réflexions *diffuse* et/ou *spéculaire*, par le fait qu'ils prennent en compte ou non les milieux participatifs (gaz, verre, etc.). Le résultat de la simulation d'un moteur d'*illumination globale* est la *luminance* (flux par unité de surface, par unité d'angle solide, exprimée en $Wm^{-2}sr^{-1}$) d'un élément de surface visible (depuis la caméra) dans la direction de la caméra.

Le logiciel de synthèse d'image physique calcule, pour chaque pixel x_i de l'image, la luminance L réfléchie dans la direction de la caméra ω_{x_i} de l'élément de surface visible s_{x_i}. La luminance est calculée pour un ensemble de longueur d'onde λ_k : $L(s_{x_i}, \omega_{x_i}, \lambda_k)$. Les luminances calculées sont des valeurs réelles positives non bornées. Elles sont qualifiées de valeurs *HDR* (*High Dynamic Range*). Elles ne sont pas directement affichables sur les écrans standards dit *LDR* (*Low Dynamic Range*) qui n'acceptent des valeurs de pixels codées entre 0 et 255 dans des espaces de couleurs RGB. Ces valeurs *LDR* sont habituellement appelées : *Luma*. Dans la suite du document, nous les appelons luminances *LDR*. Seul les écrans *HDR*, encore réservés aux laboratoires de recherche, permettent de restituer directement les images de luminances *HDR*.

Comme les images de luminances *HDR* ne peuvent pas être directement affichées

ou imprimées, il est nécessaire de réduire ou compresser les luminances afin de les ramener dans un espace de couleur compatible avec les dispositifs de restitution classiques. Cette étape est habituellement appelée *tone mapping*. Les couleurs de l'image brute de luminances *HDR* ne correspondent pas aux couleurs perçues par un observateur humain situé dans la scène. En effet les couleurs calculées par le moteur d'*illumination globale* représentent la couleur physique. Le système visuel humain adapte les couleurs afin d'avoir une perception constante des couleurs des objets quelque soient les conditions d'illumination. Cette étape dite d'*adaptation chromatique* peut prise en compte dans les post-traitements de l'image brute afin de restituer des couleurs paraissant naturelles. Dans cette étape de post-traitement il est possible de transformer l'image afin de contrôler son apparence. Par exemple, un flou de profondeur peut être appliqué à l'image afin de prendre en compte la profondeur de champ d'une optique ou de simuler le comportement de la vision humaine. Le contraste, la luminosité, la saturation peuvent être également modifiés.

Le dernier élément est l'image finale habituellement codé dans un espace de couleur RGB afin qu'elle puisse être affichée directement sur un écran ou imprimée. L'image est un tableau 2D de pixels où chaque pixel est codé comme un triplé de valeur entière comprise entre 0 et 255. La valeur n'est pas une valeur physique de luminance mais une valeur relative habituellement appelée LUMA.

Ces différents éléments seront plus détaillés et expliqués dans les différents chapitres du document. Nous préciserons les enjeux qui leur sont liés et nos contributions.

2.2 Domaines d'application

Les domaines d'applications de la synthèse d'images physique sont nombreux. De part sa capacité à simuler physiquement les interactions entre la lumière et la scène 3D, l'une des applications de la synthèse d'images physique est la simulation de l'éclairage dans le domaine de l'architecture, du design ou de la conception. En effet outre la capacité de visualiser un ensemble architectural (ou un objet) avant sa construction, la synthèse d'images physique permet également de calculer les valeurs d'éclairement, de calculer des images montrant fidèlement l'apparence des objets.

Un autre domaine d'application important est celui du cinéma. La synthèse d'images physique peut y être utilisée dans deux contextes : premièrement, pour la création complète de films dits d'animation et, deuxièmement, pour la créations d'effets spéciaux.

Dans le cadre de la production des films d'animation, la synthèse d'images physique offre aux créateurs la capacité de reproduire de nombreuses interactions entre la lumière et la scène 3D comme les multiples réflexions diffuses et spéculaires, l'absorption et la diffusion dans des nuages, des brumes ou des fumées. Il est ainsi

possible de produire des images avec une lumière d'une grande richesse. De même, la capacité de la synthèse d'images physique de restituer l'apparence d'une large gamme de matériaux est un élément important.

Les effets spéciaux consistent principalement à mélanger des images issues de prise de vues réelles et des images de synthèse, par exemple ajouter des objets virtuels dans des images réelles. Il est alors crucial que l'éclairement des objets virtuels corresponde à l'éclairage réel de la prise de vue. La capacité de la synthèse d'images physique de simuler précisément l'éclairage est alors un élément clé.

La synthèse d'images physique est également utilisée dans le domaine des jeux vidéo. Dans ce cadre, la principale contrainte est la capacité de calculer en temps réel les images. La contrainte temps réel est atteinte en simplifiant les interactions physiques prises en compte. La simulation se limite classiquement à un seul rebond diffus.

2.3 Enjeux

Les enjeux vis à vis de la synthèse d'images physique sont nombreux. Celui sur lequel se porte une part importante des travaux est à la réduction des temps de calcul des images. Le temps de calcul est en effet un critère clé quant à l'utilisation effective de la synthèse d'images physique pour la production de films ou pour son utilisation dans le domaine des jeux vidéo. Le temps de calcul est souvent à mettre en relation avec la qualité d'image. En effet dans de nombreuses approches algorithmiques, il est possible de fixer un temps de calcul, le critère devient alors la qualité obtenu dans le temps de calcul donné. La qualité s'évalue habituellement avec une mesure d'écart, PSNR : *Peak Signal to Noise Ratio* par exemple, entre l'image calculée dans un temps donné et une image dite de référence. Cette dernière est obtenue par une méthode de calcul non biaisée (convergeant vers la solution mathématique) après un temps de simulation très long. Dans ce cadre, l'élément discriminant devient à la vitesse de convergence de l'algorithme. Elle mesure la vitesse avec laquelle l'algorithme converge vers la l'image de référence. Une grande vitesse de convergence indique que pour un temps donné l'algorithme donne une image proche de la solution de référence.

Dans le cadre de la production de séquences d'images, la stabilité temporelle du moteur de rendu est également très importante. En effet, si entre deux images successives le rendu diffère alors ces différences sont souvent très visibles et gênantes lorsque que la vidéo est jouée. Par exemple, de très légères fluctuations dans le bruit de l'image apparaissent comme de la neige. Tous les aspects concernant la cohérence temporelle des images calculées sont donc un élément clés dans le cadre de la production de séquences d'images. Ils portent sur le moteur de synthèse d'image en tant que tel, mais ils sont également à prendre en compte dans les autres étapes de la production. Ainsi, l'étape de post-traitement, et en particulier dans la phase de

réduction des gammes de luminance, est également très sensible aux problèmes de cohérences temporelles.

Dans le cadre des effets spéciaux, où le but est de mélanger des prises de images réelles (issues d'une caméra) avec des images « virtuelles » (calculées par synthèse d'images physique), la précision physique est très importante. Comme le but est que le spectateur ne voit pas de différences entre les parties réelles et les parties calculées de l'image finale, il est nécessaire que les parties calculées soient semblables aux parties réelles de l'image. Cette capacité à produire des images semblables aux images réelles provient de différents facteurs : la précision de la modélisation des sources de lumières réelles, la modélisation des matériaux dont le comportement doit être proche du comportement des matériaux réels, la précision de l'algorithme de calcul des images (algorithme non biaisé qui converge vers la solution mathématique), les effets pris en compte dans la simulation (diffusion multiple de la lumière dans les fumées par exemple), le traitement des gammes de luminances qui doit être conforme au comportement de la caméra et la concordance entre les couleurs calculées et celles restituées par la caméra (même adaptation chromatique par exemple). Tous ces éléments peuvent entrainer une apparence différentes entre les images réelles et les images calculées.

Un autre enjeu important partagé par tous les domaines d'application, concerne la capacité offerte au créateur de contrôler simplement l'image (son apparence, son style, son esthétique) qu'il souhaite obtenir. Cela passe par le réglage approprié de tous les paramètres de contrôle disponibles dans tous les éléments de la production des images : modélisation, moteur de rendu et post traitements. Dans l'étape de modélisation, les outils de contrôle sont naturellement ceux permettant de contrôler la création de la scène. Dans le cas de scènes complexes composées de très nombreux objets (modélisation de villes par exemple), la modélisation interactive via une interface utilisateur atteint ses limites. Des méthodes de modélisation automatique peuvent être utilisées. La capacité à contrôler le résultat de la modélisation est alors un enjeu important. Dans la modélisation entre également la modélisation des matériaux. La capacité à modéliser des apparences spécifiques peut être une tâche délicate et souvent hors de porteur d'un créateur n'ayant pas des compétences fortes en programmation. La modélisation inclue également la spécification des sources de lumières. L'élément clé permettant d'obtenir l'éclairage désiré dans l'image finale nécessite la capacité à modéliser finement la géométrie de flux de sortie de la lumière, sa puissance et sa couleur. Le temps nécessaire aux réglages des paramètres peut être grand du fait de la succession des essais (calcul d'une image avec un nouvel ensemble de paramètres) nécessaires pour déterminer les bonnes valeurs des paramètres. Dans ce processus de réglage, il est capital que les changements effectués se répercutent dans un sens qui paraisse naturel dans l'image calculée. La simulation physique permet normalement d'assurer ce comportement. Néanmoins du fait des limites de la simulations, des éventuels biais et des erreurs de la méthode de calcul, une modification de la modélisation peut entrainer un changement non escompté dans l'image finale. Cela montre également l'importance du paramétrage du moteur

de rendu. Par exemple, limiter le nombre de réflexion/de rebond de la lumière pris en compte dans la simulation peut entrainer un comportement qui parait peu réaliste. En dehors du moteur de rendu, les algorithmes de post traitement ont également des paramètres important vis à vis de contrôle de l'image finale. Chaque traitement peut modifier l'apparence de l'image finale allant dans une direction opposée aux changements souhaités par le créateur. Par exemple l'algorithme de réduction des gammes de luminances se comporte généralement comme le réglage automatique de l'exposition d'un appareil photo en essayant de trouver la meilleure exposition pour l'image finale. Le créateur peut, par exemple, augmenter la puissance des sources de lumière de la scène 3D afin d'obtenir une image finale plus lumineuse. Mias s'il laisse l'algorithme réduction des gammes de luminances traiter automatiquement l'image de luminance, alors l'image finale peut être moins lumineuse (qu'avant l'augmentation de la puissances des sources). Cette exemple montre la complexité pour le créateur de maîtriser l'image finale car elle nécessite la compréhension des différents paramètres mis à sa disposition.

Le tableau 2.1 résume les enjeux importants par domaine d'application. Les enjeux sont notés : Temps de calcul, Cohérence tempor. pour cohérence temporelle, Précision et Contrôle. Les domaines d'applications sont notés : simul. pour la simulation d'éclairage, films d'anim. pour la production de films d'animation, SFX pour les effets spéciaux et jeux vidéo.

	simul.	films d'anim.	SFX	jeux vidéos
Temps de calcul	faible	important	important	capital
Cohérence tempor.	faible	capital	capital	capital
Précision	capital	important	capital	faible
Contrôle	important	capital	capital	capital

Table 2.1 – Les enjeux par domaines d'applications

Les travaux présentés ici sont principalement situés dans ce cadre dans les outils pour un meilleur contrôle du créateur sur l'image finale. Ils se situent dans les différentes étapes de la chaîne de production des images : modélisation, moteur de rendu et post-traitements. Nous présentons une méthode de modélisation automatique de scène 3D à partir de bases de données relationnelles. Cet outil permet à un créateur de construire facilement des mondes 3D à partir d'un langage de modélisation qui extrait les informations d'une base de données. Nous proposons également une méthode de rendu efficace qui prend en compte des milieux participatifs en illumination globale. Cette contribution en dehors des innovations algorithmes rend accessible aux créateurs la richesse de la lumière diffusée au seing des milieux participatifs. Cela est indispensable pour obtenir certaines atmosphère dans l'image finale. Dans le domaine des post-traitements, nous présentons premièrement un modèle précis d'attention visuelle qui permet de calculer le point de la scène sur lequel se fixe le regard de l'utilisateur. Un tel modèle d'attention visuelle est utile pour les post-traitements qui nécessite de connaître le point regardé. Deuxièmement, nous

proposons des solutions afin de régler les problèmes de cohérence spatio-temporelle issus de la réduction des gammes de luminances. Enfin, nous proposons une estimation précise et robuste de manière temporelle de la couleur d'adaptation adaptée au contexte de l'illumination globale. Ces contributions sont replacées dans leur contexte et précisées dans la section suivante.

2.4 Contributions

2.4.1 Modélisation automatique de scènes 3D

Dans le cadre de la thèse CIFRE avec France Telecom R&D de Noémie Esnault [1], nous avons proposé un outil de production automatique de mondes 3D à partir de données extraites d'une base de données relationnelle. La problématique de la thèse porte sur la possibilité de construire un monde 3D à partir de données stockées dans une base de données relationnelle (une base de données de films en vidéo à la demande par exemple). Le monde 3D ne peut pas être statique car il doit prendre en compte les ajouts/suppressions d'informations de la base seul la structure relationnelle de la base est considérée comme statique. Dans ce cadre il faut proposer un outil de construction automatique de monde 3D à partir d'extractions d'informations faites à la volée. Nous avons proposé un langage permettant, d'une part, d'extraire et de structurer sous forme de hiérarchie les informations contenues dans la base de données et, d'autre part, de construire à partir de la hiérarchie un monde 3D. Les tests menés avec des infographistes ont démontré la facilité de contrôle, la flexibilité, de l'approche proposée. Cette contribution a donné lieu à différentes publications : une publication à destination de la communauté francophone de l'informatique graphique [2], deux conférences internationales [3, 4] et une publication dans une revue [5] qui correspond à une version détaillée et complétée de la publication acceptée en conférence.

2.4.2 Rendu des milieux participatifs en illumination globale

Le Master 2 Recherche de Charly Collin [6] a porté sur la prise en compte des milieux participatifs (gaz, fumée, brume, verre, etc.) en illumination globale. La prise en compte précise des milieux participatifs permet de restituer des effets complexes de lumière provenant de sa diffusion multiple au seing du milieu. La simulation de la lumière dans le volume contenant le milieu participatif accroît considérablement le coût de calcul et le coût mémoire. Il est donc nécessaire de développer des méthodes spécifiques, rapides et maîtrisant le coût mémoire. Dans ce contexte, nous avons proposé une méthode progressive (succession de simulations) permettant la maîtrise du coût mémoire et exploitant une structure accélératrice originale spécifique à la gestion des volumes. Ces travaux ont été poursuivis dans le cadre des thèses de Charly Collin (sous la direction Sumanta Pattanaik, University of Central Florida, Orlando, USA) et d'Adrien Gruson. Cette contribution à la simulation de la

lumière dans les milieux participatifs a été publiée sous forme de *Talk* SIGGRAPH [7] et dans une revue internationale [8]. Nous avons également proposé une méthode de simulation des milieux participatifs rapide fonctionnant sur carte graphique [9].

2.4.3 Modèle d'attention visuelle et ses applications

Dans le cadre de la thèse CIFRE avec France Telecom R&D de Sébastien Hillaire [10], nous avons travaillé sur la création d'un modèle d'attention visuelle pour les applications 3D interactives. Le modèle d'attention visuelle calcule automatiquement le point observé dans la scène par un utilisateur qui se ballade et effectue des tâches dans un monde 3D. Nous avons donc proposé une méthode exploitant des informations provenant de l'image tel que la détermination des régions saillantes, les données issues des actions de l'utilisateur dans le monde et des données hauts niveaux provenant du scénario d'utilisation du monde 3D. Ce modèle original mixant des informations bas et haut niveaux a été publié dans deux conférences internationales [11, 12] et deux revues internationales [13, 14].

Afin d'améliorer la qualité d'image dans mes mondes 3D interactifs (jeux vidéos, réalité virtuelle, etc.), nous avons proposé d'inclure au moteur d'illumination globale temps réel, un post-traitement permettant de restituer la profondeur de champ d'une caméra ou de système visuel humain. Nous avons donc proposé un algorithme de flou de profondeur temps réel et nous avons évalué son impact sur l'utilisateur. Afin de déterminer quelle zone de l'image doit être nette, nous avons proposé un système d'autofocus intelligent qui utilise les actions de l'utilisateur, le scénario et des données du monde 3D. Cette contribution a donné lieu à trois publications : une conférence nationale [15], une conférence internationale [16] et une revue internationale [17].

2.4.4 Cohérences temporelles dans la réduction des gammes de luminances

Nous avons commencé à travailler sur la gestion des gammes de luminances et des couleurs dès 2007 [18]. Dans le cadre du projet FUI Nevex, du projet européen COST HDRi et de la thèse CIFRE de Ronan Boitard avec Technicolor R&D, nous avons focalisé nos travaux sur la réduction des gammes de luminances des vidéos HDR. Nous avons formalisé les problématiques propres aux vidéos : le scintillement haute fréquence et le manque de cohérence temporelle basse fréquence. Plutôt que de travailler sur la proposition d'un nouvel algorithme de réduction des gammes de luminances HDR, nous avons choisi de développer des algorithmes de contrôle de la cohérence qui fonctionnent avec la majorité des solutions proposées pour réduire les gammes de luminance. Cette approche laisse la liberté aux utilisateurs de choisir la méthode de réduction des gammes de luminances qui donne le meilleur résultat sur les contenus, tout en leur proposant des outils de contrôle de la cohérence. En autre enjeu important propre aux vidéos est la compression du flux vidéo HDR afin

de pouvoir distribuer le contenu sur les réseaux existant. Nous avons montré que le fait d'assurer une bonne cohérence temporelle permet un gain substantiel dans la compression du flux. Ces travaux toujours en cours ont donné lieu à des publications : deux workshops internationaux [19, 20], trois conférences internationales [21, 22, 23] et une revue internationale [24].

2.4.5 Estimation du blanc de référence en illumination globale

En complément de la gestion des gammes de luminances, nous avons travaillé sur la gestion de la couleur, en particulier dans le cadre des images calculées par illumination globale. Les couleurs de images brutes calculées par le moteur d'illumination globale correspondent aux couleurs physiques et diffèrent des couleurs telles qu'elles seraient perçues par un observateur observant la scène. Cette différence provient de l'adaptation chromatique caractéristique du système visuel humain. L'adaptation chromatique du système visuel humain nous permet de percevoir les couleurs comme constante quelque soient les conditions d'illumination. Les modèles d'adaptation chromatique nécessitent l'estimation d'une couleur appelée couleur d'adaptation. Dans le cadre de l'illumination globale, c'est à dire dans le cas où la scène 3D est connue, nous avons proposé une méthode originale d'estimation de la couleur d'adaptation. Notre méthode donne une estimation précise et correcte même dans les cas de conditions d'illumination complexes où les méthodes proposées précédemment ne fonctionnent pas. En plus de son estimation précise, notre méthode est également plus cohérente temporellement et elle peut donc être utilisée pour l'adaptation chromatique des séquences d'images. Cette contribution a donné lieu à une publication [25] dans une revue internationale.

Le tableau 2.2 résume les publications (revues internationales, conférences internationales, *workshops* internationaux et conférences nationales) relatives aux contributions issues des travaux de recherche présentés dans ce document.

	Références	Total
revues int.	[18, 17, 13, 5, 14, 8, 24, 25]	8
Conférences Internationales	[16, 11, 12, 3, 4, 9, 7, 21, 22, 23]	10
Workshop Inter.	[19, 20]	2
Conférences Nationales	[15, 2]	2

Table 2.2 – Résumé des publications

2.5 Structure du document

La suite du document est organisée comme suit. Le chapitre 3 fournit, dans la partie 3.2, les notions physique et mathématiques nécessaires à la compréhension du document. Celles ci portent sur la lumière et les grandeurs de physiques associées, sur les espaces de couleur. La modélisation des interactions entre la lumière avec les objets et sa formalisation sous la forme d'équation de rendu est détaillée. La partie 3.3 présente les éléments classiques qu'un créateur d'images, plus précisément un photographe, a à sa disposition pour obtenir l'image qu'il souhaite. Les principales esthétiques de la lumière sont détaillées. Cette analogie avec la gestion de la lumière en photographie met en évidence le besoins d'outils permettant un contrôle précis de l'esthétique de l'image calculée 3.4.

Les chapitres suivant présentent en détails nos contributions. Le chapitre 4 présente nos travaux sur la modélisation de scène 3D à partir d'informations issues de bases de données relationnelles. Le chapitre 5 détaille nos algorithmes d'illumination globale incluant les milieux participatifs. Le chapitre 6 regroupe nos travaux portant, d'une part, sur la création d'un modèle d'attention visuelle temps réel 6.1 adaptée à la visualisation des mondes 3D, d'autre part, sur la compression des gammes de luminances des séquences d'images HDR 6.2 et sur l'estimation de la couleur d'adaptation pour l'adaptation chromatique des images calculées en illumination globale 6.3.

Le chapitre 7 présente les perspectives ouvertes pour de futures recherches qui prolongent les résultats déjà obtenus et explore de nouvelles problématiques. Nous nous proposons notamment de travailler sur les problématiques portant sur la prise en compte de l'esthétique souhaitée de l'image calculée.

Chapitre 3

De la lumière à l'esthétique de l'image

3.1 Introduction

La synthèse d'images physique est un outil de création d'images comparable à la photographie ou au cinéma du fait qu'ils partagent les deux mêmes éléments clés : premièrement, le rôle de la lumière qui dessine/modèle la scène et deuxièmement, le rôle de la caméra qui capture une image correspondant à un point de vue et à un rendu particulier de la scène mise en lumière. L'étymologie de mot photographie provient du préfixe « photo » qui signifie lumière et du suffixe « graphie » qui signifie « dessiner ». Elle montre bien l'importance de la lumière. Dessiner avec la lumière sous-entend gérer la lumière depuis les sources jusqu'à l'image. Le but est d'obtenir une certaine esthétique de lumière dans l'image. Cette esthétique de la lumière dans l'image passe donc par le contrôle :

- des sources de lumières : leur position, leur puissance, leur couleur,
- des objets et de la composition de la scène : les éléments avec lesquels la lumière interagit,
- de l'appareil qui capture cette lumière et la transforme en image.

Afin de comprendre le processus de création de l'image, il est nécessaire de comprendre, d'une part, la physique de la lumière, ses interactions dans la scène 3D, sa transformation en image, et d'autre part, les familles d'esthétiques de la lumière.

Dans une première partie 3.2, nous présentons brièvement les modèles physiques permettant de décrire la lumière, ses interactions avec les matériaux et le processus de transformation en valeur de pixels des grandeurs physiques lumineuses. Quelques notions clés sur la perception de la lumière et de la couleur sont également présentées. La deuxième partie 3.3 porte sur la description des solutions habituelles permettant de gérer la lumière et son esthétique en photographie.

3.2 La lumière : de la luminance à la colorimétrie

Dans cette partie nous rappelons brièvement les éléments (grandeurs physiques, équation de rendu, espace de couleurs) important pour la compréhension des parties suivantes.

3.2.1 Luminance et Équation de rendu

Quand on regarde un point x sur une surface, on perçoit le flux émis par ce point x dans la direction d'observation $\vec{\omega}$. La grandeur physique correspondante est le flux par unité de surface par unité d'angle solide ($W/(m^2 sr)$ ou cd/m^2) , c'est à dire la luminance spectrale $L(x, \vec{\omega}, \lambda)$.

La luminance $L(x, \vec{\omega}, \lambda)$ provient de la surface elle même (si c'est une source de lumière), terme $L_e(x, \vec{\omega}, \lambda)$ dans l'équation 3.1, et de la réflexion des luminances incidentes en x provenant des autres surfaces de la scène, terme $L_i(x, \vec{\omega_i}, \lambda)$. Seule une partie des luminances incidentes est réfléchie dans la direction de l'observateur, cette partie est donnée par la fonction de réflectance bidirectionnelle (BRDF). On obtient donc l'équation de rendu[26] :

$$L(x, \vec{\omega}, \lambda) = L_e(x, \vec{\omega}, \lambda) + \int_{\Omega} f_r(x, \vec{\omega}, \vec{\omega_i}, \lambda) L_i(x, \vec{\omega_i}, \lambda)(\vec{\omega_i} \cdot \vec{n}) d\omega_i \qquad (3.1)$$

avec x le point observé selon la direction $\vec{\omega}$, $\vec{n}$ la normale en ce point, $L_e(x, \vec{\omega}, \lambda)$ la luminance spectrale émisse en x dans la direction $\vec{\omega}$, f_r la BRDF (fonction de réflectance bidirectionnelle), Ω l'hémisphère centré en x, $L_i(x, \vec{\omega_i}, \lambda)$ la luminance spectrale incidence en x depuis la direction $\vec{\omega_i}$ et λ la longueur d'onde (Cf. figure 3.1).

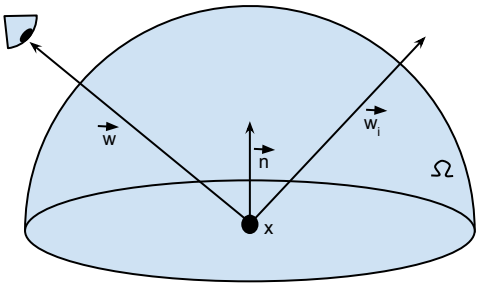


FIGURE 3.1 – Élément géométrique de l'équation 3.1.

En présence de milieu participatif, l'équation devient [27] :

$$\begin{aligned} L(x,\vec{\omega},\lambda) &= \int_0^s Tr(x \leftrightarrow x')\sigma_a(x')L_e(x',\vec{\omega},\lambda)dx' \\ &+ \int_0^s Tr(x \leftrightarrow x')\sigma_s(x')L_i(x',\vec{\omega},\lambda)dx' \\ &+ Tr(x \leftrightarrow x_s)L(x_s,\vec{\omega},\lambda) \end{aligned} \tag{3.2}$$

$L(x,\vec{\omega},\lambda)$ est la luminance arrivant en x selon la direction $-\vec{\omega}$, x_s est le point visible par la demi-droite partant de x dans la direction $-\vec{\omega}$, s est la distance traversée par la demi-droite dans le volume, $L_i(x',\vec{\omega},\lambda)$ est la luminance arrivant en x' et diffusée dans la la direction $\vec{\omega}$, $L_e(x',\vec{\omega},\lambda)$ est la luminance émise en x' dans la direction $\vec{\omega}$, $L(x_s,\vec{\omega},\lambda)$ est la luminance provenant de x_s dans la direction $\vec{\omega}$, $Tr(x \leftrightarrow x')$ est la transmittance entre les points x et x', σ_a est le coefficient d'absorption, et σ_a est le coefficient de diffusion.

La luminance énergétique est l'intégrale des luminances spectrales :

$$L_E(x,\vec{\omega}) = \int_0^{+\infty} L(x,\vec{\omega},\lambda)d\lambda \tag{3.3}$$

3.2.2 Photométrie et colorimétrie

L'œil humain a une sensibilité différente à la luminance spectrale en fonction de la longueur d'onde. Elle est donnée par la fonction d'efficacité lumineuse relative spectrale $V(\lambda)$ (Cf. figure 3.2, normalisée par la CIE).

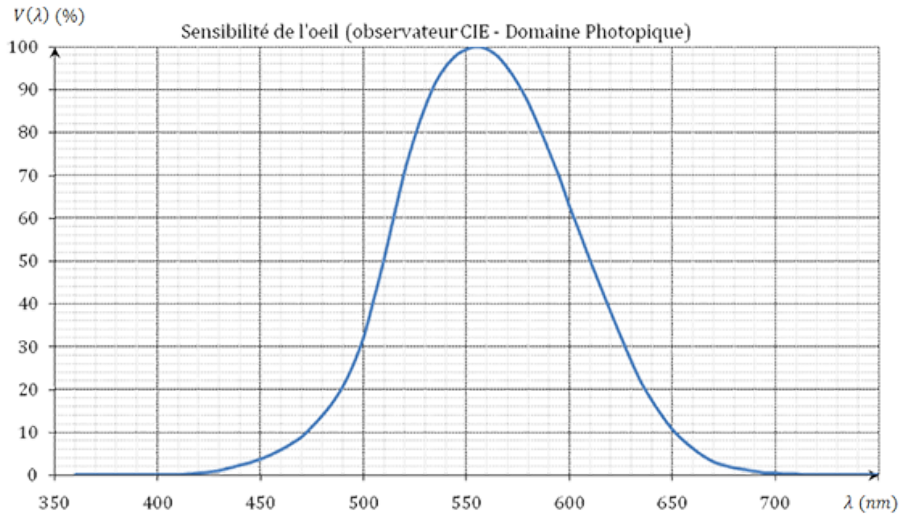


FIGURE 3.2 – Éfficacité lumineuse relative spectrale.

La luminance photométrique est alors définie comme :

$$L_p(x,\vec{\omega}) = \int_{\lambda=380}^{\lambda=780} V(\lambda)L(x,\vec{\omega},\lambda)d\lambda \tag{3.4}$$

La luminance photométrique ne donne que l'indication de puissance par unité de surface par unité d'angle solide. Afin de conserver les informations de couleur, la luminance physique spectrale peut être projetée dans un espace de couleur trichromatique comme l'espace CIE RGB [28, 29] :

$$R = \int_{\lambda=380}^{\lambda=780} \bar{r}(\lambda)L(x,\vec{\omega},\lambda)d\lambda \tag{3.5}$$

$$G = \int_{\lambda=380}^{\lambda=780} \bar{g}(\lambda)L(x,\vec{\omega},\lambda)d\lambda \tag{3.6}$$

$$B = \int_{\lambda=380}^{\lambda=780} \bar{b}(\lambda)L(x,\vec{\omega},\lambda)d\lambda \tag{3.7}$$

avec $\bar{r}$, $\bar{g}$ et $\bar{b}$ les fonctions de bases trichromatiques (Cf. figure 3.3).

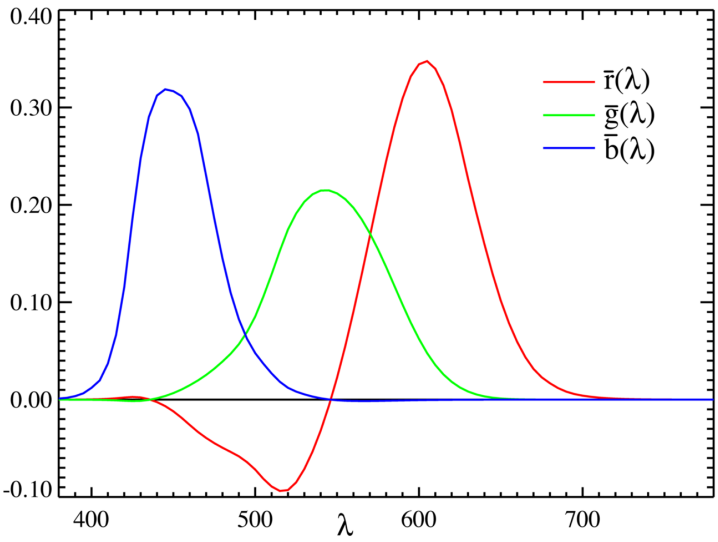


FIGURE 3.3 – Fonctions $\bar{r}$, $\bar{g}$ et $\bar{b}$ de l'espace de couleur CIE RGB.

Les valeurs de luminances (R, G, B) sont des valeurs réelles et peuvent varier de 10^{-3} (ciel étoilé) à 10^5 (Cf. table 3.1). On parle alors de grande gamme de luminance (HDR : *High Dynamic Range*). Le système visuel humain est sensible à cette large gamme de luminance. Ces luminances HDR doivent être réduites entre 0 et 255 et quantifiées en valeur entière pour être affichées sur les écrans standard. Ces luminances appelées *luma* ou luminances LDR (LDR : *Low Dynamic Range*). La

réduction des gammes de luminances (*tone mapping*) a donné lieu à de nombreux travaux [30, 31].

Condition	Luminance (Cd/m^2)
Nuit étoilée	10^{-3}
Lumière de la lune	10^{-1}
Lumière intérieure	10^{2}
Soleil	10^{5}

TABLE 3.1 – Ordre de grandeur des luminances

Chaque opérateur de réduction des gammes de luminances a sa propre stratégie de réduction des gammes de luminances : de la simulation du système visuel humain [32] à une reproduction du développement photographique [33]. Chaque opérateur de réduction donne donc un résultats différents des autres comme le montre la figure 3.4. Les opérateurs de réduction de gamme de luminances sont généralement classés selon deux grandes catégories : opérateurs globaux et opérateurs locaux.

Les opérateurs globaux utilisent une unique fonction f_{tmo} de compression des luminances pour toute l'image :

$$l_{LDR} = f_{tmo}(L_{HDR}) \tag{3.8}$$

avec L_{HDR} la luminance et l_{LDR} la luminance réduite.

Les opérateurs locaux utilisent quant à eux une fonction f_{tmo} qui dépend de la luminance $L_{HDR}(p)$ du pixel p mais également du voisinage du pixel $V(p)$:

$$l_{LDR}(p) = f_{tmo}(L_{HDR}(p), V(p)) \tag{3.9}$$

dès lors à une valeur de luminance HDR L_{HDR} correspond plusieurs valeurs de luminance LDR l_{LDR} en fonction du voisinage du pixel.

3.2.3 Perception et apparence

Les espaces de couleurs permettent d'encoder les luminances HDR ou LDR. Chaque espace de couleur a ses propres propriétés : proche du matériel pour CIE RGB, couvrant l'ensemble des couleurs visibles CIE XYZ, perceptuellement uniforme CIE Lab. Mais ils ne décrivent pas la perception ou l'apparence d'un stimulus coloré par un observateur. C'est le rôle des modèles d'apparence de couleurs [29]. Ils sont construits à partir d'observations menées dans des conditions contrôlées. Ils prennent en compte (Cf. figure 3.5) le stimulus coloré (d'une taille correspondant à 2 ° d'angle), l'arrière-plan (*background*, d'une taille de 8 ° d'angle) et le champ environnant (*surround*).



FIGURE 3.4 – Résultats de la réduction des gammes de luminances avec des opérateur différents (de gauche à droite et de bas en haut : [34], [35], [36], [37], [32] et [33]).

L'apparence de la couleur est décrite selon des valeurs absolues et des valeurs relatives : *Brightness*, *Lightness*, *Chroma*, *Saturation*, *Colorfulness* et *Hue*. Les modèles d'apparence sont capables de reproduire certains comportement du système visuel humain et notamment l'adaptation chromatique. C'est l'une des caractéristiques clés de la perception humaine des couleurs. Nous percevons comme constante la couleur des objets quelque soient les conditions d'éclairement et en particulier quelque soit la couleur des sources de lumières [38]. La figure 3.6 montre la différence entre les couleurs physiques et les couleurs perçues par un observateur.

J. Von Kries [39] fut le premier à proposer un modèle pour l'adaptation chromatique. En 1905, il fait l'hypothèse que chaque type de photorécepteur (L, M, S) s'adapte linéairement indépendamment les uns des autres. Des améliorations ont depuis été apportées au modèle de Von Kries, en particulier pour prendre en compte le comportement non-linéaire de l'adaptation chromatique [40, 41, 42, 43] et le phénomène d'adaptation incomplète dans les faibles valeurs de luminances [44, 45, 46]. L'adaptation de Von Kries s'écrit :

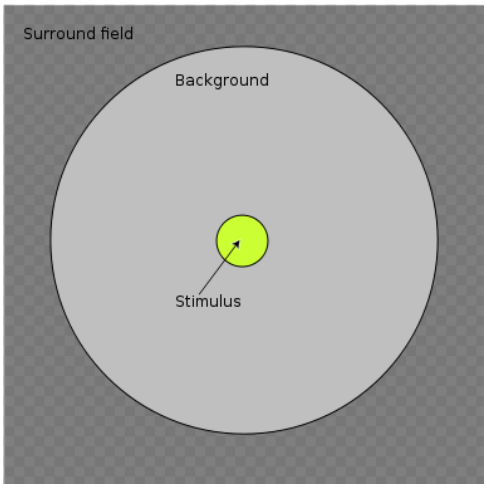


FIGURE 3.5 – Conditions d'observation.



FIGURE 3.6 – Image brute : couleur physique (gauche), Image après adaptation chromatique : couleurs perçues, comparaison entre la couleur physique et la couleur perçue du tapis (droite).

$$\begin{bmatrix} L_a \\ M_a \\ S_a \end{bmatrix} = \begin{bmatrix} 1/L_W & 0 & 0 \\ 0 & 1/M_W & 0 \\ 0 & 0 & 1/S_W \end{bmatrix} \begin{bmatrix} L \\ M \\ S \end{bmatrix} \tag{3.10}$$

avec L,M,S sont les réponses initiales des cônes, L_a,M_a,S_a les réponses adaptées et L_w,M_w,S_w la couleur d'adaptation.

Après avoir rappelé brièvement les éléments physiques, photométriques et perceptuels de la lumière, nous décrivons dans les parties suivantes comment la lumière est abordée dans le domaine de la création d'image, en particulier dans le domaine de la photographie et du cinéma.

3.3 Dessiner avec la lumière : une intention

Le photographe, l'artiste, qui conçoit une image connaît le résultat qu'il souhaite obtenir : il a une intention, une émotion, un message qu'il veut transmettre. Son travail a pour but de contrôler les éléments, lui permettant d'atteindre son but. Ce dernier peut être décrit en terme de composition, de cadrage, de lumière, de température de couleurs, d'harmonie des couleurs, de netteté, de profondeur de champ, etc. Tous ces éléments correspondent à des choix techniques conscients fait à partir de l'esthétique/style souhaité. En effet, l'image est construite à partir des possibilités qu'offrent les différents éléments techniques concourant à la production de l'image [47] :

- composition, cadrage,
- lumière,
- couleur,
- netteté et profondeur de champ.

3.3.1 Composition et cadrage

Le photographe construit la scène en fonction de la composition souhaitée de l'image. Il choisit les objets, leur forme, leur couleur et leur apparence. Ils les placent les uns par rapport aux autres et choisit un point de vue sur la scène ainsi constituée. La composition de l'image est le résultat de la composition de la scène et du cadrage choisi. Les différents types de cadrage ou plans [48] sont classés en quatre grandes familles : plans larges, plans moyens, plans rapprochés et gros plans. Les règles techniques sur la composition des images sont communes à tous les arts graphiques et elles s'appliquent à la photographie, comme à la synthèse d'image, sans contraintes supplémentaires particulières [48]. L'une des plus anciennes règles de composition est la règle dite des tiers [49]. Elle préconise à placer les éléments importants sur les droites horizontales et verticales divisant l'image en tiers (Cf. Figure 3.7).

Le photographe choisit la (distance) focale de l'objectif qu'il utilise. La focale détermine l'angle de champ visible dans la photographie. Néanmoins, il est possible d'obtenir le même cadrage avec différentes focales. Pour un cadrage donné, le choix de la focale détermine les déformations perspectives entre la géométrie de la scène réelle et sa projection sur la photographie. Par exemple, le choix d'un objectif grand angle (faible focale) entraîne une forte déformation des objets et donner ainsi une force dramatique forte (à des bâtiments par exemple). Au contraire, le choix d'une focale longue écrase la perspective et peut, par exemple, faire ressortir les rythmes et les répétitions d'une façade d'un bâtiment comme dans la photo d'Andreas Gursky, « Paris Montparnasse », 1993. Le savoir-faire du photographe lui permet de déterminer quel cadrage/focale/composition choisir pour atteindre son but.



FIGURE 3.7 – Composition et règle des tiers : les éléments importants de la photo sont placés sur les droites correspondant aux tiers verticaux et horizontaux et sur leurs intersections (Crédit photo : armor-photo.com).

3.3.2 Lumière

Le créateur détermine la configuration de l'éclairage : il choisit la position, l'orientation et les autres caractéristiques des sources de lumières. Cette partie de la création de l'image est l'une des plus importantes car elle détermine directement le modelé et le contraste des objets dans l'image, mais également l'esthétique globale de la lumière dans l'image [50]. En dehors de la position et de l'orientation de la source de lumière, les autres caractéristiques clés sont sa puissance et la géométrie du flux de lumière. Le créateur a à sa disposition une large gamme de dispositif permettant de choisir la géométrie du flux de lumière sortant de la source : réflecteurs, boîtes à lumière et parapluies (Cf. Figure 3.8). Plus le flux est parallèle (« directionnel ») plus le contraste sur l'objet est important. Un flux de sortie très diffus, donne un contraste faible.

Dans le cadre de la photographie ou de la cinématographie de studio où tout l'éclairage est artificiel et contrôlé, des configurations types d'éclairage ont été formalisées. Le modèle le plus classique pour formaliser la gestion de la lumière est le modèle à trois sources de lumière [51] : *key light*, *fill light* et *back light* (Cf. Figure 3.9).

La lumière principale (*key light*) impose l'ombrage, le modelé sur l'objet qui est le sujet principal de la photographie. L'angle relatif entre la caméra, le sujet et la *key light* est un élément déterminant, il permet de mettre plus ou moins en évidence le volume général et les détails de la surface. Le terme éclairage *Rembrant* [52] est utilisé quand la *key light* est positionnée à 45˚.



FIGURE 3.8 – Formes de réflecteurs différentes permettant de contrôler le flux de sortie de la lumière (source : Elinchrom).

La seconde source (*fill light*), dont la puissance est moins que forte que celle de la *key light*, permet de modifier les zones d'ombres provenant de la *key light*. Cela permet ainsi de contrôler la valeur de contraste entre les zones claires (éclairée par la *key light*) et les zones sombres éclairées quant à elles par la *fill light*.

Une troisième source appelée (*back light*) est placée derrière le sujet et l'éclaire par l'arrière. Elle permet d'éclairer la silhouette (le contour) de l'objet. Cela permet de détacher le sujet du fond.

Les caractéristiques des trois sources (position, puissance, forme de flux de lumière sortant) permet au créateur de l'image de contrôler finement l'éclairage reçu par les différentes zones composant de l'image. Typiquement le ratio de puissance entre la *key light* et la *fill light* donne le contraste entre les zones claires (éclairées par la *key light*) et les zones sombres (éclairées par la *fill light*). Alors que la *back light* permet de détacher le sujet de la photographie du fond et met en valeur la silhouette du sujet (Cf. Figure 3.10).

L'équilibre entre les différentes zones de l'image a été formalisé par Ansel Adams dès 1949 [53, 54]. Il définit une esthétique de lumière, où toutes les parties de l'images sont correctement exposées et équilibrées les unes par rapport aux autres. Cette esthétique est classiquement appelé style *medium key*. D'autres esthétiques de la lumière en photographie ont été proposées [55], en particulier les esthétiques *high key* et *low key* :

- L'esthétique de lumière *high key* : des images globalement très lumineuses avec un faible contraste.
- L'esthétique de lumière *low key* : des images globalement sombres mais avec des forts contrastes [56].

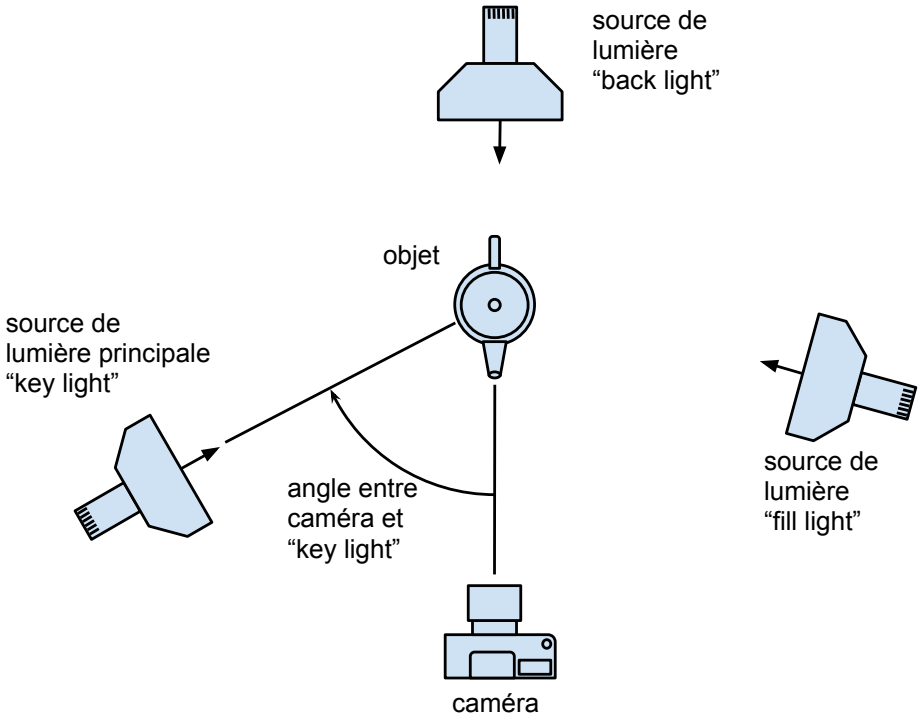


FIGURE 3.9 – Éclairage à 3 sources : *key light*, *fill light* et *back light*.

3.3.3 Couleur

En construisant la scène, le créateur prend en compte non seulement les formes, mais également les couleurs des objets et leur apparence. Cela lui permet de gérer l'harmonie des couleurs dans l'image finale. L'harmonie des couleurs est un domaine qui a été largement étudié etformalisé. Dès 1810, Johann Wolfgang Von Goethe [57] propose une théorie sur l'harmonie des couleurs basée sur le « cercle chromatique » (*Color Wheel*). L'harmonie des couleurs a ainsi donnés lieu à de nombreuses théories : [58, 59, 60]. Ces modèles d'harmonie ont été utilisés pour proposer des algorithmes permettant d'harmoniser automatiquement les photographies [61, 62], les vidéos [63] et les mondes virtuels [64]. La figure 3.11 illustre les 6 classes d'harmonie des couleurs de Morton [59].

En complément à l'harmonisation des couleurs, le créateur a également la possibilité de choisir la couleur des différentes sources de lumière : *key light*,*fill light* et *back light*. La *key light* étant la source principale, c'est elle qui donne l'ambiance générale. La couleur de la *fill light* contrôle directement la couleur des ombres provenant de la *fill light* et la *back light* permet de contrôler la couleur du contour du



FIGURE 3.10 – Éclairage à 3 sources : le ratio de puissance *key light*/*fill light* contrôle le contraste entre les zones claires (éclairées par la *key light*) et les zones sombres (éclairées par la *fill light*). La *back light* sépare le sujet de la photo du fond et met en valeur la silhouette en créant des éclats sur les contours. (Crédit photo : armor-photo.com)

sujet (Cf. Figure 3.12).

À l'ambiance générale de couleur de l'image est généralement associée la notion de température de couleur. La température d'une couleur est la température d'un corps noir idéal qui rayonne une lumière de cette la couleur. En fonction de la température, la gamme des couleurs va du bleu à l'orange (Cf. figure 3.13).

L'ambiance de couleur d'une image correspond à la dominante de couleur de l'image (couleur moyenne de l'image). En faisant l'hypothèse que la moyenne des réflectances des objets de la scène est achromatique, alors l'ambiance de couleur ne dépend que de deux éléments : la couleur de la *key light* et la couleur d'adaptation de la balance des blancs. La couleur d'adaptation permet de supprimer l'effet du à la couleur de la *key light*. Si la couleur d'adaptation x_a est identique à la couleur x_k de la *key light* alors la moyenne des couleurs de l'image sera achromatique. Pour le photographe ou l'opérateur de la caméra, c'est le ratio x_k/x_a qui déterminera l'ambiance de couleur de l'image finale.

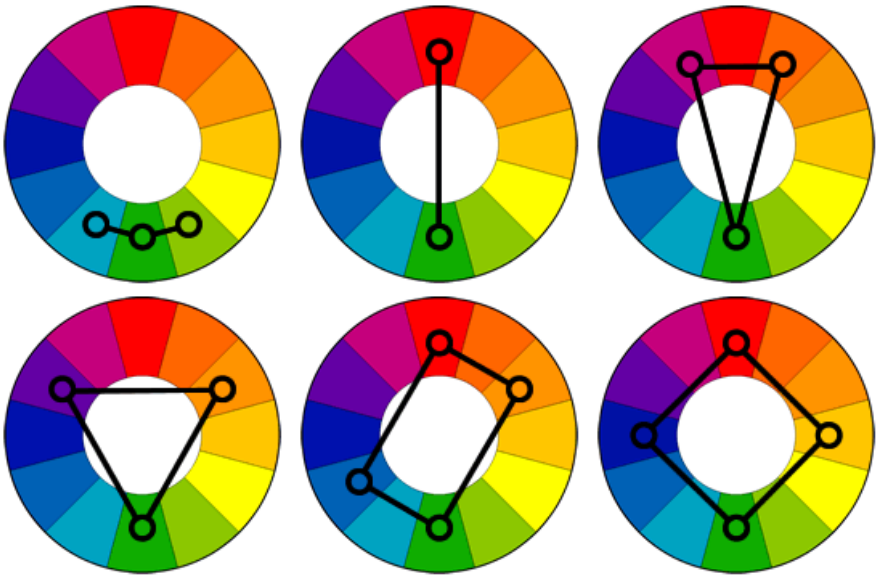


FIGURE 3.11 – Harmonies des couleurs : (haut-gauche) Analogue, (haut-centre) Complémentaire, (haut-droite) Complémentaire à 2 tons, (bas-gauche) 3 tons, (bas-centre) Complémentaire à 4 tons, (bas-gauche) 4 tons.

3.3.4 Netteté et profondeur de champ

Le dernier élément de contrôle sur l'esthétique de l'image est la netteté et la profondeur de champ. Le créateur détermine quelle partie de l'image est nette et quelle est l'intervalle de profondeur de la scène qui est nette sur l'image. Pour contrôler cela, le créateur a deux paramètres, il choisit de faire la mise au point sur une partie de la scène et il règle la profondeur de champ en réglant l'ouverture du diaphragme de l'appareil photo ou de la caméra. Plus l'ouverture est grande, plus petite est la profondeur de champ.

Limiter la profondeur de champ permet de focaliser l'attention de l'observateur sur la partie nette de la photo. En effet le regard se concentre naturellement sur les zones nettes des images. La partie nette de la photo est donc mise en valeur.

3.4 Éléments de contrôle

En tant que créateur de logiciel de synthèse d'images physique nous devons prendre en compte les éléments de contrôle permettant au créateur/artiste d'obtenir l'image qu'il souhaite. Ces principaux éléments de contrôle, présentés dans les parties précédentes, sont :



FIGURE 3.12 – Influence des couleurs des sources de lumière : (à gauche) *fill light* de couleur verte qui contrôle la couleur des zones non éclairées par la *key light* – (à droite) *back light* de couleur bleue donnant un reflet bleu sur la silhouette du sujet. (Crédit photo : armor-photo.com)

- la composition de la scène,
- le cadrage de l'image,
- la lumière,
- la couleur et
- la netteté et profondeur de champ.

Les outils proposés aux créateurs ne doivent ni imposer un style ou une esthétique unique, ni une méthode de travail. Les outils développés doivent en effet pouvoir s'intégrer dans différents approches ou différents environnements de créations et/ou de production, et ils doivent également laisser la liberté aux créateurs de produire les images avec l'esthétique/style souhaité.

Les outils interactifs de modélisation ne pose pas de problème particulier pour créer des scènes relativement complexes et choisir le cadrage désiré. Les outils de modélisation atteignent leur limite pour des scènes très complexes comme une forêt ou une ville par exemple. La principale limite des outils de modélisation interactifs concerne le paramétrage des matériaux dont les réglages peuvent être peu intuitif.

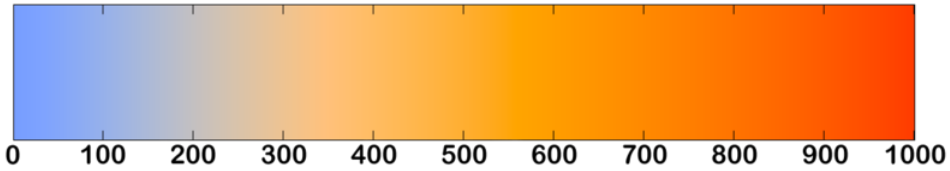


FIGURE 3.13 – Des couleurs froides aux couleurs chaudes (en mireds).

Le réglage de la lumière et des couleurs s'avère beaucoup plus difficile. Le contrôle des sources de lumière ne pose pas de problème en tant que telle. Mais du fait des limites du moteur de rendu, du temps de calcul des images et de la phase d'adaptation des gammes de luminances et des couleurs, le réglage des paramètres est une tâche peu intuitive qui nécessite de nombreux essais/erreurs. De plus comme les paramètres peuvent être redondant rien n'assure au créateur d'atteindre l'image qu'il souhaite sans une certaine maîtrise de l'ensemble des paramètres de la chaîne de production de l'image. La figure 3.14 illustre la difficulté du réglage des paramètres. L'objectif est d'obtenir une image de style *high key* (image très lumineuse avec peu de contraste) de la statue. La configuration des sources de lumières virtuelles correspond à un éclairage *high key* (ratio *key light*/*fill light* proche entre 1, *fill light* de couleur bleu). L'image obtenue avec les réglages par défaut de la chaîne de production (paramètres du moteur et de la réduction des gammes de luminances) produit une image trop sombre (Cf. figure 3.14 - image de gauche). L'augmentation de la puissance des sources de lumière virtuelle ne change pas le résultat. L'image de droite de la figure 3.14 correspond à l'image souhaitée, elle est obtenue en changeant uniquement les paramètre de l'opérateur de réduction des gammes de luminances.

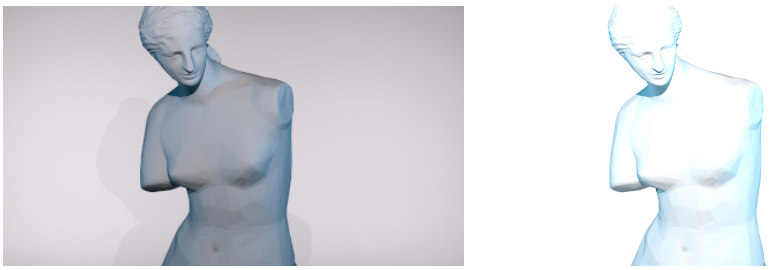


FIGURE 3.14 – Paramètre de rendu (à droite) image obtenue avec les paramètres par défaut - (à gauche) image obtenue avec un changement des paramètres de l'opérateur de réduction des gammes de luminances.

Le réglage de la profondeur de champ d'une caméra virtuelle, s'il est disponible, peut se faire de manière interactive et donc facilement. Dans ce cas la limite vient du contrôle de ce paramètre pour des applications interactives. Il s'agit alors de définir, dans un environnement 3D dans lequel l'utilisateur est libre de choisir son point de vue, sur quel objet (ou quelle partie de l'objet) le point doit être fait.

En conclusion, une problématique importante pour la synthèse d'images physique est d'offrir aux créateurs, d'une part, une grande liberté de création et, d'autre part, des outils lui permettant d'atteindre rapidement l'image qu'il souhaite. Si le but n'est pas de créer une image mais une séquence d'images, la dimension temporelle ajoute une complexité plus importante, il est alors nécessaire d'assurer la cohérence spatio-temporelle des images produites. Et si le but est de créer des applications 3D interactives (comme des jeux vidéos, des applications de visites virtuelles, des applications de réalités virtuelle), la problématique devient encore plus complexes.

Chapitre 4

Modélisation automatique de scènes 3D

4.1 Introduction

La description de la scène 3D est l'entrée indispensable au moteur de rendu dont le rôle est de calculer une image de cette scène. La modélisation la plus utilisée pour décrire une scène 3D est la modélisation surfacique. Elle consiste à spécifier les surfaces des objets de la scène en terme de géométrie et en terme de comportement de la surface vis à vis de la lumière. La modélisation géométrique des surfaces peut se faire de différentes manières : numérisation, modélisation interactive, catalogues d'objets paramétriques et modélisation procédurale. La problématique pour le créateur de la scène 3D est de choisir l'outil de modélisation le mieux adapté pour créer rapidement et simplement la scène qu'il souhaite obtenir.

Dans le contexte de la thèse CIFRE avec France Télécom R&D de Noémie Esnault [1], la problématique posée était la création d'une scène 3D mettant en scène des données issues d'une base de données relationnelle ou sémantique. Le contenu de la base de données pouvant évoluer, la scène 3D doit donc être calculée dynamiquement afin de prendre en compte les modifications de la base. Seule la structure de la base de données est constante. L'exemple typique est une base de données de films cinématographiques incluant les informations suivantes : titres, réalisateurs, styles, acteurs, etc. Le but est alors de construire une scène 3D dans la laquelle un utilisateur peut se balader et découvrir les différents films présents dans la base. Cette problématique est clairement située dans le contexte de la modélisation procédurale. En effet, il s'agit de calculer automatiquement une scène 3D à partir d'informations qui ne sont pas des informations représentant directement la géométrie d'un monde 3D.

En s'inspirant des architectures de créations de pages web dynamiques dont le contenu est extrait de bases de données, puis mis en forme, nous avons proposé une

solution de modélisation automatique de scène 3D à partir d'informations extraites dynamiquement d'une base de données. Notre solution construit dans une première étape une structure d'exploration hiérarchique, appelée *génotype*, des données. Le *génotype* est calculé à partir de la base de données et de la description de la stratégie d'exploration des données : le *génome*. Le *génone* ne dépend que de la topologie de la base de données, alors que le *génotype* prend en compte les éléments présents dans la base de données. Dans une deuxième étape, le monde 3D est calculé. La description de la métaphore visuelle est décrite dans le *phénome*. Ce dernier est appliqué sur le *génotype* afin de calculer le monde 3D appelé *phénotype*.

Les caractéristiques clés de l'approche que nous avons proposée sont :
- la compatibilité avec différents systèmes d'accès aux données : base de données relationnelle, sémantique, résultat de moteurs de recherches.
- la possibilité d'explorer de différentes manières (*génomes* différents, même *phénome*) des données ayant un même style de représentation 3D,
- la possibilité de de proposer des représentations différentes (*phénomes* différents) tout en conservant la même structure de navigation dans les données (même *génome*).

La contribution majeure de notre modèle est la séparation forte entre :
- les données,
- la structure de navigation dans les données et
- la visualisation de ces données

Notre modèle propose un cadre formel qui est indépendant de la base de données et de sa structure : il est générique. Notre modèle favorise la réutilisabilité en permettant au créateur soit de réutiliser des métaphores de structuration/navigation/exploration, soit de réutiliser des métaphores visuelles et des styles de visualisation.

La partie 4.2 présente un bref panorama des travaux en relation avec le contexte du travail présenté ici : visualisation 3D des bases de données, structuration des mondes 3D et technologies Web permettant la séparation entre les données et leur représentation. Notre modèle de construction automatique de monde 3D à partir des éléments extraits dynamiquement d'une base de données est présenté dans la partie 4.3. Notre solution est illustrée sur un exemple concret dans la partie 4.4. Enfin la partie 4.5 conclut sur cette partie des contributions présentées dans ce document.

4.2 Travaux précédants connexes

Peu de travaux portent sur la création d'un environnement générique pour la visualisation 3D de bases de données quelconques. Les travaux proposent habituellement une solution adaptée pour une base de données particulière.

Ainsi Alessio Bosca et al. [65] proposent une solution pour la visualisation des bases de données sémantiques représentant des ontologies. Les éléments de la base

de données sont visualisés en tirant partie des clés visuelles (couleurs) spécifiques pour chaque classe de concept de l'ontologie. La métaphore de visualisation, sphères interconnectées, ne peut pas être modifiée.

Dans [66], les auteurs proposent, quant à eux, une visualisation 3D des différents parcours d'une université. Les différents éléments extraits de la base de données sont : département, module, cours, etc. Ils sont représentés par des cubes dont la taille et la couleur sont déterminées par les informations associées : durée du module, nombre de crédits, etc.

Walczak et al. proposent une méthode permettant la création automatique d'expositions dans des musées virtuels. Le nombre de pièces, leurs caractéristiques et les objets présentés sont extraits d'une base de données. La méthode proposée permet uniquement la personnalisation de la géométrie générée.

Dans [67], les auteurs proposent un environnement logiciel basé sur XML pour naviguer l'intérieur d'un graphe de données sémantiques. L'environnement 3D est calculé automatiquement et représente les nœuds et les liens de l'ensemble sémantique.

Les nouveaux standards du WEB promeuvent une forte séparation entre les données,leur présentation et leur mise en page. Dans cette approche les données sont structurées dans des dialectes XML. Le style graphique des éléments (textes, graphiques2D et images) et la mise en page sont définis par des feuilles de style *CSS* (Cascading Style Sheets, [68]). Les feuilles de style sont décrites en dehors des documents contenant les données. Elles sont appliquées dynamiquement par le navigateur Web sur le document contenant les données. La prochaine norme de *CSS* (actuellement en phase de spécification) comprendra un ensemble limité de transformations 3D (*CSS 3D Transform*), permettant aux éléments 2D d'être rendus dans un espace 3D.

Les solutions proposées ne sont adaptées que pour une base de données particulière (dans les approches présentées ci-dessus : des cursus universitaires, une base de données d'un musée). La visualisation et la navigation proposées sont spécifiques. Pour conclure, il manque une solution générique de création automatique d'environnements 3D pouvant être simplement spécifiés et personnalisés. Dans ce contexte, nous proposons une nouvelle solution permettant de construire automatiquement des mondes 3D à partir de données extraites dynamiquement de bases de données.

4.3 Environnement de création automatique de monde 3D

Notre objectif est de fournir un outil permettant la définition et la construction automatique d'un environnement 3D offrant différentes métaphores de visualisation

d'un ensemble de données. L'ensemble de données d'entrée peut être n'importe quel type de données stockées dans une base de données ou des données sémantiques (comme des données *RDF* [69] ou des cartes topiques (*topic maps* [70]). La seule contrainte est de pouvoir spécifier une structure hiérarchique à partir des données.

4.3.1 Vue d'ensemble

La solution a été fortement inspirée par la philosophie des outils de création Web. Elle est basée sur la séparation entre la structure des données et la métaphore visuelle qui permet la visualisation des données. La séparation entre la structure des données et la métaphore de visualisation est assurée par deux composants (en référence à [71]) :

- le *génome* : définition la structuration des données,
- le *phénome* : définition du style visuel de la représentation des données.

La construction de la structure d'exploration est appelée *construction du génotype*. Elle consiste en la construction d'un graphe acyclique orienté, appelé *génotype*, qui décrit la structure d'exploration des données extraites de la base. La construction du génotype est réalisée à partir d'une spécification de la structure hiérarchique d'exploration appelée *génome*. Le *génotype* est le résultat de l'application du *génome* sur une base de donnée :

$$Genome \times BdD \longrightarrow Genotype \tag{4.1}$$

La construction de la scène 3D, appelée *construction du phénotype*, organise le *génotype* dans l'espace 3D et habille les données avec des modèles 3D. Ceci est fait en utilisant des grammaires procédurales décrite dans le *phénome* :

$$Phenome \times Genotype \longrightarrow Phenotype \tag{4.2}$$

4.3.2 Calcul du génotype

Le module de calcul de *génotype* a pour objectif de sélectionner les informations pertinentes dans une base de données afin de créer la structure hiérarchique d'exploration des données (*génotype*). Cette structure hiérarchique est la topologie exploratoire du monde 3D. La structure hiérarchique abstraite du *génotype* est décrite par le *génome*. Celui-ci décrit la succession des niveaux d'exploration (mot clés : $EXPLO$) et les informations (ATT) à extraire de la base de données qui seront associés à chaque élément appartenant à un niveau d'exploration.

La grammaire simplifiée, qui ne tient pas compte de sa mise en forme en dialecte XML, du *génome* est :

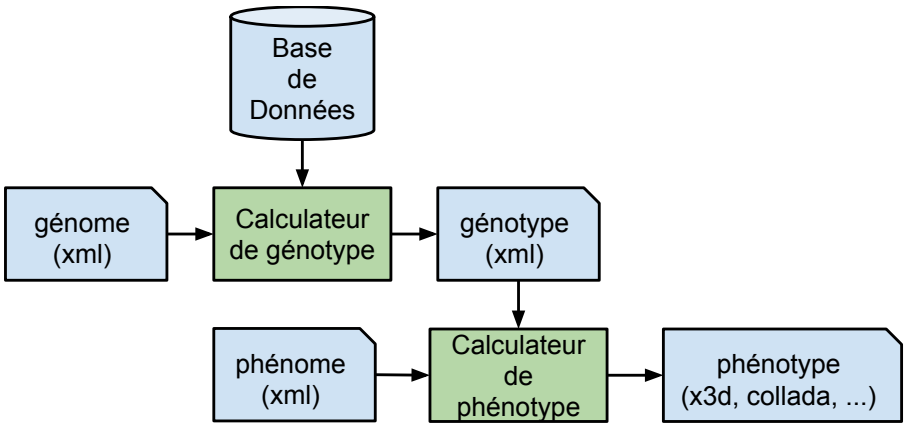


FIGURE 4.1 – Architecture de l'environnement de création automatique de monde 3D à partir de bases de données : séparation entre structuration des données (génome et génotype) et visualisation (phénome et phénotype).

$genome ::= GENOME\ racine\ explo$
$explo ::= exploDef \mid exploRef$
$exploDef ::= EXPLO\ id\ PATH\ chemin\ [attributs]\ [explo]$
$exploRef ::= EXPLO\ ref$
$racine ::= EXPLO\ id$
$attributs ::= ATT\ id\ type\ PATH\ chemin$
$type ::= STRING|IMAGE|...$

Le *génome* est défini par une *racine* et par une hiérarchie de niveaux d'exploration. A chaque niveau d'exploration est associé un identificateur *id*, le *chemin* permettant d'accéder aux éléments associés dans la base de données et des *attributs* extraits de la base de données. Les *attrbiuts* sont typés, nommés et possèdent le *chemin* permettant de les extraire dans la base de données.

L'exemple décrit par la figure 4.2 illustre le rôle génome. La base de données relationnelle (partie supérieure de la figure 4.2) contient des films, des genres, des réalisateurs, des acteurs et des dates de sortie. Cette base de données ne contient pas de structure hiérarchique, elle relie des données. Dans cet exemple, le créateur décide de créer un monde 3D permettant d'explorer cette base de données. Il choisit la structure d'exploration suivante : par genres, puis par réalisateurs, puis par films. Il associe des attributs à chaque niveau d'exploration : le nom du genre pour les genres, le nom du réalisateur pour les réalisateurs, le titre du film et l'affiche du film pour les films.

Le calcul du *génotype* s'effectue en extrayant les éléments de la base de données correspondant aux niveaux d'exploration définis dans le *génome*. Les éléments extraits sont organisés dans un arbre où chaque niveau de l'arbre correspond à un niveau d'exploration (Cf. Figure 4.3). Ce processus de construction du *génotype* est

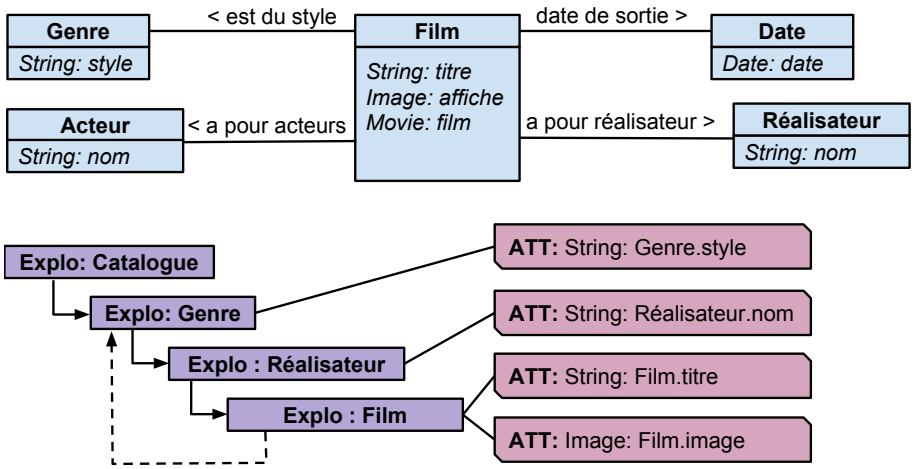


FIGURE 4.2 – Définition de la structure hiérarchique d'exploration par le génome.

réalisé à l'aide de transformation XSLT et de requête XQuery [72, 73].

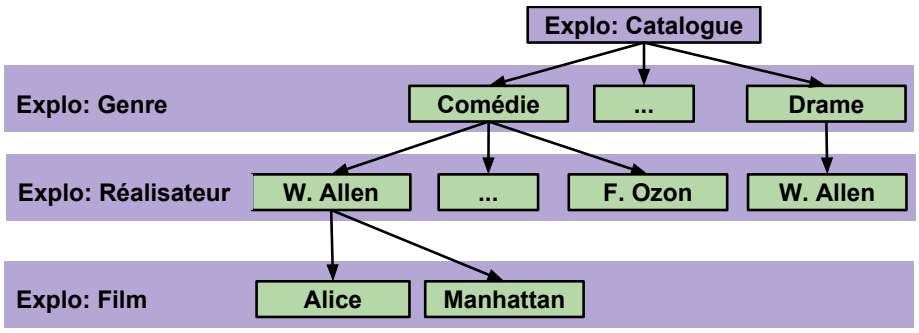


FIGURE 4.3 – Exemple de *génotype* extrait de la base de données à partir des niveaux d'exploration définis par le *génome*.

Le *génotype* se présente sous la forme d'un fichier XML qui structure de manière hiérarchique les éléments extraits de la base de données. La hiérarchie est conforme aux niveaux d'exploration du *génome*. Le *génome* ne dépend que de la structure de la base de données (chemin d'accès aux données). Il est indépendant des éléments présents dans la base de données. Le *génotype* correspond à l'extraction et à la structuration des informations (éléments) de la base de données.

4.3.3 Calcul du phénotype

Le *phénotype* est un monde 3D structuré sous la forme d'un graphe de scène décrit dans un dialecte XML générique qui est traduit dans un format de monde 3D,

comme X3D [74], WebGL [75] ou Collada [76], pour son affichage. Le dialecte XML du *phénotype* comporte les fonctionnalités 3D classiques des langages de description de mondes 3D :

- transformation géométriques : translation, rotation, mise à l'échelle,
- définition de géométrie 3D,
- textures,
- définition de matériaux,
- gestion du graphe de scène,
- gestion des interactions.

Le *phénotype* organise en 3D les *éléments* extraits de la base de données et structurés (hiérarchie) par le *génotype*. Le style (positionnement, géométrie, aspect, ...) des éléments et leur organisation dans l'espace sont décrits dans le *phénome*. Les caractéristiques du *phénotype* sont les suivantes : son style est conforme au *phénome*, sa topologie (structure en graphe de scène) est conforme au *génotype* (Cf. figure 4.4).

A chaque niveau d'exploration correspond un certain nombre d'*éléments*. Le *phénome* définit le modèle 3D associé à un *élément* en fonction de son niveau dans la hiérarchie du *génotype*. Le modèle associé à chaque *élément* peut être personnalisé en fonction : des *attributs* de l'*élément*, des *attributs* de son *élément* père dans la hiérarchie, des *attributs* de ses *éléments* fils et de propriétés comme le nombre de fils. Ainsi si l'*élément* a un *attribut* image, cette dernière peut être utilisée comme une texture dans le modèle 3D associé à l'*élément*.

Chaque *élément* a un certain nombre de fils (niveau d'*exploration* suivant), le *phénome* décrit comment sont distribués les fils : distribution circulaire ou linéaire, fonction mathématique, etc. Afin de permettre un contrôle fin de la distribution des *éléments* fils, le langage du phénotype propose les structures de contrôle classiques : boucle et conditionnelle.

En complément de création et de la personnalisation de modèle 3D, le *phénome* permet également de spécifier des interactions. Le modèle d'interaction proposé par le *phénome* est celui de VRML [77]. Les interactions sont décrites à l'aide de trois famille de composants : les capteurs (qui évaluent les conditions d'activation), les expressions qui permettent de faire des calculs et des animations (animations prédéfinies représentées par des images-clés).

Un extrait de la grammaire du *phénome* (grammaire partielle simplifiée qui ne prend pas en compte la syntaxe XML du langage) est :

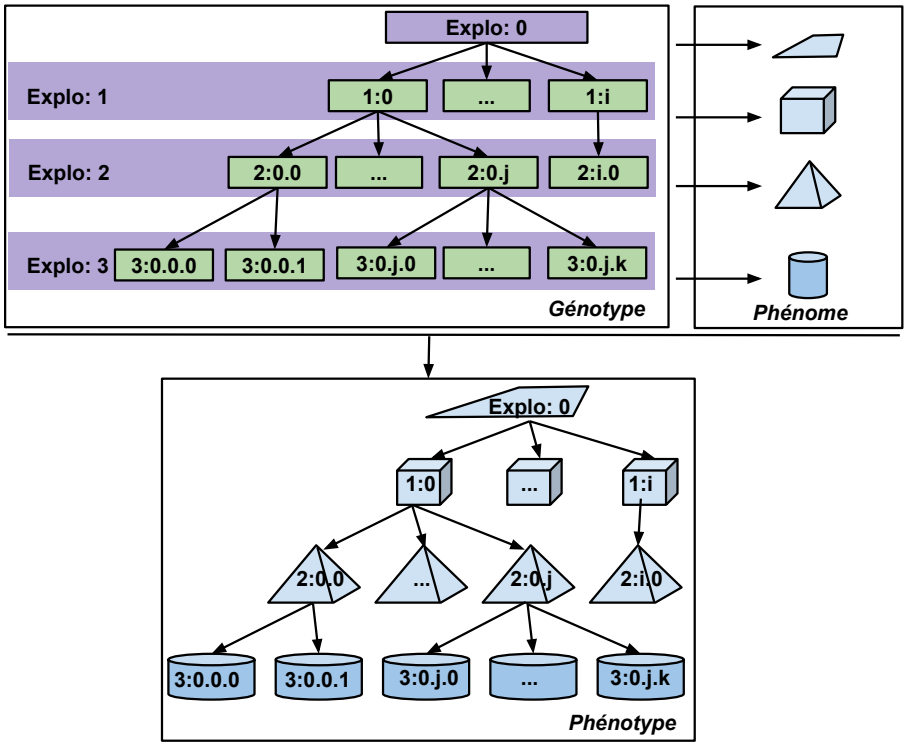


FIGURE 4.4 – Le *phénome* décrit l'aspect et l'organisation des éléments du *génotype*.

$phenome ::= PHENOME\ [defs]\ LEVEL\ rules\ [level]$
$level ::= LEVEL\ rules\ [level]$
$rules ::= RULE\ exp \mid RULE\ id$
$defs ::= DEF\ id\ exp$
$exp ::= cond \mid iter \mid expT$
$cond ::= IF\ expT\ THEN\ expT\ [ELSE\ expT]$
$iter ::= FOREACH\ expT\ DO\ expT$
$expT ::= ...$

Cette extrait de la grammaire montre qu'un *phénome* est constitué de :

- d'un ensemble de définition (def) permettant la simplification de l'écriture,
- d'une hiérarchie ($LEVEL$) de règles ($rules$), les règles d'un niveau hiérarchique s'applique à un niveau d'exploration du *génotype*,
- les règles sont des expressions (exp) qui peuvent contenir des structures de contrôle, (IF, $FOREACH$),
- les expressions des règles permettent de :
  - gérer des modèles 3D,
  - gérer des transformations géométriques

– gérer des interactions
– accéder à des attributs et des propriétés des éléments du *génotype*

Un *phénome* est appliqué à un *génotype* afin de calculer un *génotype*. Ce dernier est un monde 3D exprimé dans un format permettant une traduction rapide vers les formats classiques de monde 3D comme VRML, WebGL, X3D ou Collada. Seule la partie concernant les interactions peuvent ne pas être traduite si le langage cible ne prend pas en compte les interactions. Dans ce cas l'utilisateur peut seulement se balader dans le monde 3D. Le calcul du *génotype* est mis en œuvre comme un ensemble de règle XSLT.

Il n'y a pas de condition forte limitant l'application d'un *phénome* à un *génotype*. Si les profondeurs des hiérarchies (*génotype* et du *phénome*) sont différentes, le monde 3D est créé avec la profondeur la plus petite. De même si une *règle* du *phénome* fait appel à un *attribut* d'un *élément* du *génotype* qui n'existe pas, alors la *règle* est ignorée. Ce système assure une grande indépendance entre la description du style visuel (*phénome*) et les données du *génotype*.

4.4 Résultats

Dans ce document, nous ne présentons qu'un seul exemple. D'autres exemples pourront être trouvés dans les publications relatives à ces travaux : [1, 5, 3, 4, 2]. L'exemple présenté ici correspond à l'exemple de la base de données de films (Cf. section 4.3.2 et figure 4.2).

Il s'agit d'une base de données de films contenant : des films, des réalisateurs et des genres (drame, comédie, etc.). Le créateur du monde 3D à choisit d'explorer la base selon trois niveaux d'exploration :

1. les genres,
2. les réalisateurs,
3. les films.

Les *éléments* du *génotype* sont les genres, les réalisateurs et les films présents dans la base de données.

Dans le *phénome*, le créateur associe un modèle 3D personnalisable à chaque niveau de la hiérarchie. Il définit également un modèle 3D pour la racine du monde. Le tableau 4.1 indique le modèle 3D pour chaque niveau d'exploration.

Le *phénome* décrit également comment sont distribués spatialement les *éléments* du *génotype*, par exemple distribution circulaire des cinémas (*éléments* correspondant aux différents genres dans le *génotype*). À partir des règles de constructions définies dans le *phénome* et des *éléments* du *génotype* (*éléments* extraits de la base de données en fonction de la stratégie d'exploration définie dans le *génome*), le monde 3D (i.e. le *phénotype*) est calculé (Cf. figure 4.11).

Niveau	0	Place	associé à la racine (C. Figure 4.5)
Niveau	1	Cinéma	associé à un genre (C. Figure 4.6)
Niveau	2	Salle	associé à un réalisateur (C. Figure 4.7)
Niveau	3	Affiche	associé à un film (C. Figure 4.7)

TABLE 4.1 – Association des modèles 3D en fonction de niveau d'exploration.



FIGURE 4.5 – La place 3D (niveau d'exploration 0).

Les différents tests menés avec des infographistes dans le cadre de la thèse ont démontré la facilité de contrôle, la flexibilité, de l'approche proposée. A Partir d'une même base de données, il est facilement possible de proposer des métaphores visuelles très différentes : *cover-flow*, carrousel 3D ou monde 3D.

4.5 Conclusion et Perspectives

Dans le cadre de cette thèse, nous avons proposé une solution originale et innovante pour la création automatique (procédurale) de monde 3D permettant d'explorer des données, extraites dynamiquement d'une base de données. La solution proposée est basée sur deux caractéristiques/éléments principaux :

- le module de calcul de génotype qui permet la structuration des données explorées,
- le module de calcul du phénotype qui permet de définir l'aspect visuel des environnements 3D calculés

Cette séparation entre les données et leur visualisation permet une forte réutilisation des éléments créés : *génome* et *phénome*.

Notre solution a été développée dans le cadre de la visualisation 3d de base de données. Mais elle n'est pas limitée à cette utilisation. En effet, la solution proposée est un outil de création/construction de mondes 3D complexes. A ce titre, notre solution est plus généraliste que les systèmes de créations procédurales de ville [78, 79] mais en contrepartie est limitée à une structure topologique hiérarchique.



FIGURE 4.6 – Le cinéma (niveau d'exploration 1 - les genres).

Il serait donc envisageable d'étendre le modèle de règles du *phénome* pour prendre en compte des graphes acycliques orientés.

Pour le créateur, la solution que nous proposons permet de construire un monde 3D mettant en scène les éléments d'une base de données. La créationde monde 3D à partir de base de données est fonctionnalité innovante. En ce sens, notre contribution élargit le champ des possibles, repoussent les limites de la création.

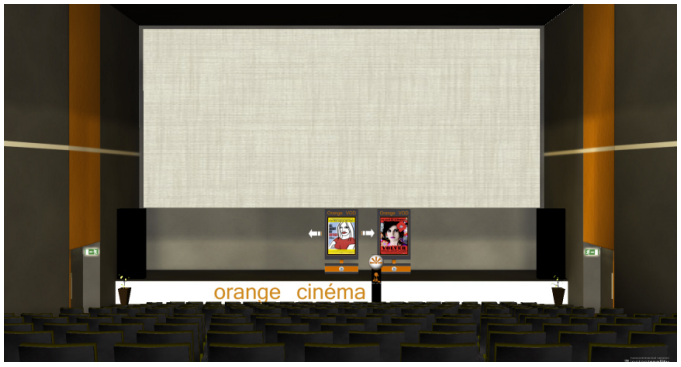


FIGURE 4.7 – La salle et les affiches de films (niveau d'exploration 2,3 - les réalisateurs, les films).



FIGURE 4.8 – La place avec les cinémas - un cinéma par genre.



FIGURE 4.9 – L'intérieur du cinéma avec les salles (accès par l'escalier) - une salle par réalisateur.



FIGURE 4.10 – La salle de cinéma et les affiches des films (en bas de l'écran) - une affiche par film.

FIGURE 4.11 – Le monde 3D (*phénotype*) construit automatiquement à partir de la base de données de films.

Chapitre 5

Rendu des milieux participatifs en illumination globale

Dans la chaine de production des images, le moteur de rendu est l'élément central. Il conditionne en grande partie la qualité des images produites. En effet l'image calculée diffère en fonction des propriétés physiques de la lumière prises en compte dans la simulation. La prise en compte de la diffusion de la lumière dans les milieux participatifs (brumes, brouillard, fumée, etc.) est essentielle pour restituer certaines atmosphères lumineuses naturelles. La figure 5.1 illustre l'apport de la prise en compte des milieux participatifs dans l'atmosphère lumineuse qui peut être restituée en synthèse d'images : il s'agit de la même scène (à gauche) sans milieux participatif et (à droite) avec un milieu participatif.



FIGURE 5.1 – Même scène 3D avec (à droite) prise en compte dans la simulation des milieux participatifs - (à gauche) sans milieux participatifs.

Dans ce contexte nous avons travaillé sur la prise en compte des milieux participatifs en illumination globale. Nous avons proposé une nouvelle approche de rendu progressif par *photon mapping* avec prise en compte des milieux participa-

tifs. Notre approche utilise deux *Kd-trees* : un premier pour stocker les *faisceaux* cylindriques qui correspondent aux chemins des rayons émis depuis la caméra et un second pour les *points visibles* depuis les rayons(i.e. fin des chemins de rayons). Cela permet de ne pas sauvegarder la carte des photons et résout ainsi le problème de l'occupation mémoire. Notre méthode gère les scènes contenant des objets diffus, spéculaires, transparents (avec réfraction) et les milieux participatifs homogènes ou hétérogènes. Note méthode peut s'adapter à toutes les techniques de rendu qui utilisent l'approche suivi de particules comme celle proposée dans [80].

5.1 Introduction

Le calcul des diffusions multiples au seing des milieux participatifs reste un défi important. Il existe des méthodes précises non biaisées (*path tracing* par exemple) mais elles sont trop gourmandes en terme de ressources de calcul et de mémoire pour être réellement utilisées. Des méthodes moins précises mais plus rapides et progressives comme [81] rendent réellement utilisables la prise en compte les milieux participatifs dans l'illumination globale. Elles étendent l'approche de rendu progressif (*progressive photon mapping*) [82] au *volume photon mapping* [83].

Notre contribution consiste à proposer une nouvelle approche au *progressive photon mapping* pour les scènes contenant des surfaces (diffuses, spéculaires et transparentes) et des volumes (milieux participatif hétérogènes ou homogènes). Les caractéristiques principales de notre approche sont :

- les photons émis ne sont pas mémorisés, ils sont détruits dès qu'ils ont contribué à un *faisceau* ou à un *point visible*, en conséquence le nombre de photons émis par passe n'est pas limité par des problèmes mémoires,
- les étapes de prétraitement (construction des *Kd-trees*) sont effectués toutes les N passes, cela réduit de manière très efficace les problèmes d'aliasing,
- les chemins de lumières sont guidés par une stratégie de *Metropolis* en fonction de la visibilité [84], cela nous permet de gérer des environnements 3D où les chemins de lumières contributifs sont complexes.

La partie 5.2 dresse en bref état de l'art des différentes approches utilisant le *photon mapping* dans les milieux participatifs. La partie 5.3 présente l'architecture globale de notre algorithme, quelques détails sont décrits dans la partie 5.4. La partie 5.5 présente des images obtenues avec notre algorithme avant de conclure sur cette contribution en 5.6.

5.2 État de l'art

La première méthode utilisant le *photon mapping* pour les volumes fut le *Volume Photon Mapping* [83]. Cette méthode permet d'obtenir des images de bonne qualité mais elle souffre des limitations inhérentes au *photon mapping*. En effet dès que la scène devient complexe, il faut un très grand nombre de photons et donc beaucoup de mémoire. En pratique, la mémoire disponible et le coût de calcul élevé

limitent drastiquement l'utilisation de cette méthode. La méthode *Beam Luminance Estimate* [85] réduit de manière significative le coût de calcul dans le cas de scènes 3D peu complexes. Mais son efficacité diminue dès que le rendu nécessite un grand nombre de photons.

Les approches progressives comme *Progressive Photon Mapping* [82] s'attaquent au problème du coût mémoire qui limite le nombre de photons utilisés pour le calcul de l'image. Le calcul se fait en plusieurs passes. A chaque passe, un nombre limité de photons sont émis et produisent une image bruitée. La carte de photons gourmande en mémoire est détruite entre chaque passe. Les images calculées à chaque passe sont cumulées afin de réduire le bruit. *A Probabilistic Approach* [81] étend la méthode progressive aux scènes contenant des milieux participatifs.

La méthode appelée *Progressive Photon Beams* [80] remplace les photons par des faisceaux de photons émis à partir des sources de lumière. Cela diminue la mémoire nécessaire pour le rendu. Mais ce gain devient vite nul si le nombre de sources augmente. De plus, dans ce cas la vitesse de convergence devient également lente.

Les méthodes précédentes produisent des images de qualité mais sont limités en fonction de la complexité de la scène. En effet, plus la scène est complexe, plus le coût mémoire (pour stocker les photons ou les faisceaux de photons) est important. Il est alors nécessaire d'augmenter de manière importante le nombre de passes. La vitesse de convergence diminue donc du fait des étapes de construction des structures de données accélératrices reconstruites à chaque passe.

Notre méthode ne stocke pas les photons. Ils sont détruits dès qu'ils ont contribué à l'image de la passe. Le problème de mémoire est donc négligeable. De plus les *Kd-trees* utilisés dans notre méthode sont conservés entre les passes. Ils ne sont reconstruits que toutes les N que pour réduire les problèmes d'aliasing.

5.3 Architecture de l'algorithme

Dans cette partie nous présentons l'architecture de notre algorithme `Algorithm 1`. Notre algorithme possède une étape de pré-traitement dont le but est de construire les deux *Kd-trees* associé aux chemins des rayons (*beam Kd-tree*) partant de la caméra et aux point visibles par ses rayons (*view Kd-tree*).

Les rayons sont émis depuis la caméra et passent par les pixels de l'image. Quand un rayon traverse le milieu participatif, il lui est associé un faisceau (cylindre dont le rayon diminue à chaque passe). Le faisceau subit les réflexions/réfractions s'il touche une surface réfléchissante (spéculaire parfait) ou transparente. Les faisceaux sont stockés dans un *Kd-tree*, appelé *beam Kd-tree*. Quand le rayon , après d'éventuelles réflexions/réfraction, touche une surface diffuse ou brillante, le point d'intersection (point visible) est mémorisé avec une zone d'influence (disque), dont le rayon diminue après chaque passe. Ces points visibles sont stockés dans un *Kd-tree*,

appelé *view Kd-tree*.

Cette étape de pré-traitement est réalisée toutes les N passes. À chaque étape de pré-traitement, les rayons sont légèrement perturbés (fonction `Jitter-View-Rays`) afin de réduire les problèmes d'aliasing.

À chaque passe un ensemble de photos sont lancés (fonction `SendPhoton`). Dès qu'un photon touche/traverse une surface/un faisceau, nous calculons l'apport du photon (fonction `UpdateCurrentLuminance`). Les photons ne sont pas sauvegardés. À la fin de cette étape de lancé de photons, l'image de la passe est calculée (fonction `Render`). La principale originalité de notre approche dans cette étape est de calculer directement l'apport du photon au lieu de le stocker dans une structure (*photon map*) qui utilise beaucoup de mémoire. Dès lors le nombre de photons par passe peut être très grand.

Finalement, à l'issue de la passe, l'image globale et les rayons des faisceaux et des disques sont mis à jour : fonctions `UpdateImage` et `UpdateRadii`.

Algorithm 1 main()

```
1: for i = 1 to nbPass do
2:   if i%N==0 then
3:     Jitter-View-Rays() ; // Perturbations des rayons primaires
4:     Preprocess() ; // Toutes les N passes
5:   end if
6:   for j = 1 to nbPhoton do
7:     SendPhoton() ;
8:     UpdateCurrentLuminance() ;
9:   end for
10:  Render() ;
11:  UpdateImage() ;
12:  UpdateRadii() ;
13: end for
```

5.4 Implémentation

Nous ne détaillons ici que le cas où des photons sont émis depuis les sources. Néanmoins notre méthode fonctionne également avec des faisceaux des photons comme proposés dans [80]. Les notations utilisés dans cette partie sont résumées dans la table 5.1.

Table 5.1 – Définitions de notations et grendeurs utilisées

Notation	Description
Φ_i	Flux du photons i
σ_a	Coefficient d'absorption
σ_s	Coefficient de diffusion
σ_t	Coefficient d'extinction
$Tr(x \leftrightarrow x')$	Transmittance entre les points x and x' $Tr(x \leftrightarrow x') = e^{-\int_0^{\|\|x-x'\|\|} \sigma_t(x+t\vec{\omega})dt}$
$p(x, \vec{\omega}, \vec{\omega'})$	Fonction de phase normalisée
$fr(x, \vec{\omega}, \vec{\omega'})$	BRDF
L_s, L_m	*Luminance* à un point visible, respectivement, à un faisceau.
L_s^t, L_m^t	*Luminance* cumulée à un point visible, respectivement, ou à un faisceau pour l'itération en cours.
N_s, N_m	Nombre de photons dans un disque (point visible) ou dans un faisceau depuis le début du rendu
N_s^t, N_m^t	Nombre de photons dans un disque (point visible) ou un faisceau pour l'itération en cours.
R_s, R_m	rayon associé au disque d'un point visible ou d'un faisceau .
α	Paramètre utilisateur permettant de contrôler la vitesse de convergence.

5.4.1 Pré-traitements

Cette étape consiste en l'initialisation des deux *Kd-tree* utilisées. À partir de la caméra, nous lançons des rayons qui sont potentiellement réfléchis et/ou réfractés. Le trajet complet du rayon est appelé *chemin du rayon*. Quand le rayon traverse un milieu participatif lui est associé un faisceau cylindrique. Un *chemin de rayon* est donc constitué de segments (en dehors des milieux participatifs) et de faisceaux cylindriques (dans les milieux participatifs). Quand le chemin a une intersection une surface diffuse ou brillante, on associe au point d'intersection (point visible) une zone d'influence (disque) (Cf. figure 5.2).

Pour un chemin ne traversant aucun volume, on associe à son point visible la *luminance* courante et la *luminance* cumulée L_s and L_s^t. Pour les faisceaux sont mémorisées les informations suivantes : la transmittance Tr, sa *luminance* courante et la *luminance* cumulée L_m and L_m^t ainsi que la fin du faisceau.

Les points visibles et les faisceaux sont structurés et mémorisés dans deux *Kd-trees* : *view Kd-tree* et *beam Kd-tree*. La stratégie de construction des *Kd-tree* utilise la méthode de Havran et al. [86]. Elle est détaillée dans [8].

5.4.2 Émission des photons dirigée selon la visibilité

Afin de gérer des conditions éclairements complexes, nous utilisons une stratégie de *Métropolis* [84] basée sur la visibilité pour guider l'émission des photons. Dans cette approche nous utilisons une stratégie de mutation pour les surfaces et les mi-

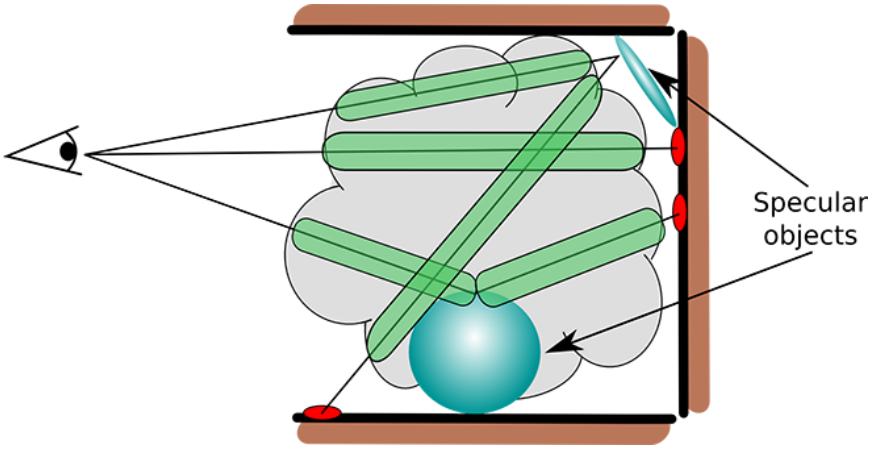


FIGURE 5.2 – Chemin de rayon : cylindre quand le rayon traverse un volume (en vert) et point visible (*view point*) (en rouge) sur les surfaces diffuses et brillantes (non spéculaires).

lieux participatifs mais avec des paramètres différents. Ces derniers sont déterminés de manière automatique avec la technique proposée dans [84]. Nous utilisons deux chaînes de Markov : une pour les surfaces et une pour les milieux participatifs.

5.4.3 Mise à jour des *luminances*

Quand un photon intersecte un faisceau, il ajoute sa contribution à la *luminance* cumulée au faisceau selon :

$$L_m^t += K(x_i, r_i) Tr(x \leftrightarrow x_i') \sigma_s(x_i') p(x_i, \vec{\omega}, \vec{\omega_i}) \Phi_i \tag{5.1}$$

Quand un photon touche un disque associé à un point visible, de la même manière, le photon contribue à la *luminance* cumulé du point visible, selon l'équation :

$$L_s^t += \frac{fr(x_i, \vec{\omega}, \vec{\omega_i})(\vec{n} \cdot \vec{\omega_i})\Phi_i}{\pi R_s^2} \times Tr(x \leftrightarrow x_s) \tag{5.2}$$

5.4.4 Mise à jour de l'image

Quand les photons de l'itération en cours ont été émis, l'image est mise à jour. Cette étape consiste à mettre à jour, premièrement, les *luminances* des points visibles et des faisceaux : L_s et L_m et deuxièmement mettre à jour les valeurs des rayons des disques (point visibles) et des faisceaux.

La mise à jour des luminances (associés aux faisceaux et aux point visibles) s'effectue selon l'algorithme proposé dans *Progressive Photon Mapping* [82] :

$$L_{(s|m)} = (L_{(s|m)} + \frac{1}{N^t_{(s|m)}} L^t_{(s|m)}) \frac{N_s + N^t_s(1-\alpha)}{N_s} \tag{5.3}$$

Pour l'image, la mise à jour se fait, en cumulant les contributions associées au chemin du rayon :

$$L = Tr_{nbB} L_s + \sum_{i=1}^{nbB} Tr_{i-1} L_{m,i}, \tag{5.4}$$

avec $L_{m,i}$ la *luminance* L_m du $i^{ème}$ faisceau du chemin, Tr_i la transmittance associée à ce faisceau ($Tr_1 = 1$, si le chemin commence depuis la caméra), nbB le nombre de faisceaux du chemin, et $Tr_{nbB} = \prod_{i=1}^{nbB} Tr_i$. Si $nbB = 0$ alors $L = L_s$ (le chemin n'a traversé aucun milieu participatif).

5.4.5 Mises à jours des rayons

Finalement pour assurer la convergence de l'algorithme global, la valeur des rayons (disques associés aux points visibles et cylindres des faisceaux) doit être réduite après chaque itération/passe. Cela se fait comme indiqué dans [81] :

$$R_s = R_s \sqrt{\frac{n+\alpha}{n+1}}, R_m = R_m \sqrt[3]{\frac{n+\alpha}{n+1}}, \tag{5.5}$$

avec n le numéro de l'itération courante.

5.5 Résultats

Nous présentons ici les résultats obtenus avec notre méthode utilisant ou non la stratégie de Métropolis [84] pour le guidage des photons ou des faisceaux de photons (proposé par [80]). Nos résultats sont comparés avec ceux obtenus par la méthode [80] (dénommé PPB) et [81] (dénommé APA). Pour les différentes variantes de notre méthode nous utilisation les dénominations suivantes :

PPT notre approche avec des photons émis depuis les sources

PPBT notre approche avec des faisceaux photons émis depuis les sources

PPT_metro notre approche version **PPT** avec Métropolis

PPBT_metro notre approche version **PPBT** avec Metropolis

Notre méthode dans ses différentes variantes (PPT, PPBT, PPT_metro et PPBT_metro) ainsi que PPB et APA ont été mises en œuvre avec le moteur de rendu Mitsuba [87]. Les paramètres de rendu pour chaque scène de test sont donnés dans la table 5.2 ($\alpha = 0.7$ pour toutes les méthodes et scènes).

Scène	*Breakfast hall*	*Kitchen*
polygones	1600k	250k
résolution	1080x1920	720x1280
photons volumiques	100k	100k
faisceau de photons	10k	10k
photons surfaciques	100k	500k
temps de rendu	10h	10h

TABLE 5.2 – Paramètre du rendu : chaque méthode utilise le même temps de rendu (dernière ligne). *photons volumiques* = nombre de photons émis par passe dans APA et PPT ; *faisceaux de photons* = nombre de faisceaux de photons dans PPB et PPBT.

Les résultats sont obtenus sur la configuration matériel bi-processeur 2.4 GHz Intel Xeon E5645 CPU (12 cœurs).

Pour chaque scène : *Breakfast hall* et *Kitchen*), nous avons calculé une image de référence en utilisant une technique de suivi de chemins (*path tracing*).

Les scènes *breakfast hall* et *kitchen* (Cf. figures 5.4 et 5.5) ont des conditions d'éclairement complexes. Dans la scène *breakfast hall*, la lumière provient de la lumière de l'extérieur et doit traverser les fenêtres. La pièce est remplie par un milieu participatif homogène. Cette scène est un cas défavorable pour Metropolis. C'est pourquoi notre technique PPT sans Métropolis est la plus efficace. Elle est plus efficace que APA et PPB (Cf. figure 5.3).

La scène de la cuisine (*kitchen*) est éclairée par le ciel et le soleil dont la lumière passe par un double vitrage. La cuisine est remplie par un milieu participatif hétérogène. Dans ce cas notre méthode avec Métropolis (PPT_Metro) donne les meilleurs résultats (Cf. figure 5.3).

5.6 Conclusion

Nous avons proposé une méthode progressive (rendu en plusieurs passes) d'illumination globale qui prend en compte les surfaces diffuses, brillantes, spéculaires, réfractives et les milieux participatifs homogènes et hétérogènes. Notre méthode utilise deux structures (*Kd-tree*) pour les points visibles et les faisceaux (émis depuis la caméra) traversant les milieux participatifs. Ces structures nous permettent de ne pas avoir à stocker les photons (surfaciques et volumiques). Dès lors le nombre de photons émis par passe n'est pas limité par des contraintes de mémoire. Les résultat montrent que notre méthode converge plus rapidement que APA [81] et PPB [80].

La prise en compte des milieux participatifs est le seul moyen pour la créateur de restituer des ambiances où la lumière est diffusée par des milieux participatifs

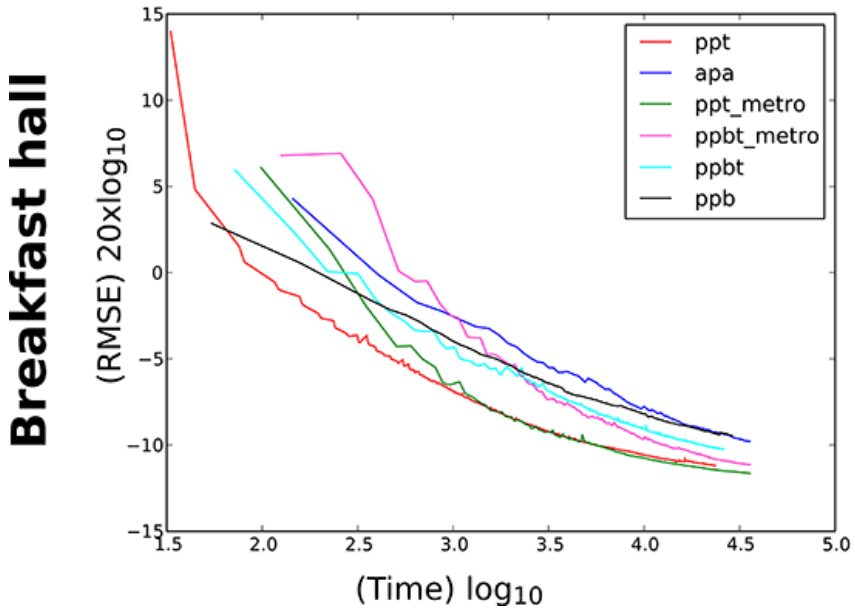


FIGURE 5.3 – Courbe de RMSE (entre l'image de référence et les différentes méthodes) en fonction du temps pour la scène Breakfast hall.

(Cf. figure 5.6). Cette possibilité n'a de sens que si les temps de calcul restent raisonnables pour obtenir des images de qualité et que la configuration matérielle requise soit abordable. La méthode que nous proposons contribue directement sur ces deux points : meilleur temps de calcul pour obtenir une qualité donnée et faible consommation de mémoire.

La méthode que nous proposons est progressive, elle permet donc d'avoir une image après chaque itération. Cela permet au créateur de vérifier au fur et à mesure du rendu (10 heures dans les cas des scènes présentées dans la partie 5.5) si l'image converge vers l'image qu'il souhaite obtenir. L'approche utilisée consiste à ne pas émettre l'ensemble des photons, puis à ne pas stocker la carte de photons mais au contraire de calculer pour chaque photon sa contribution. Cela permet de concevoir un logiciel de rendu encore plus interactif. Il est en effet possible d'interrompre le processus (même pendant une itération/passe) et demander d'afficher l'image calculée à ce moment du calcul. Dès lors le rendu devient interactif permettant au créateur de voir l'image se raffiner progressivement. Cela permet au créateur d'interrompre le calcul à tout moment, soit parce que l'image en cours de calcul ne lui convient pas, soit au contraire dès que la qualité requise est atteinte.

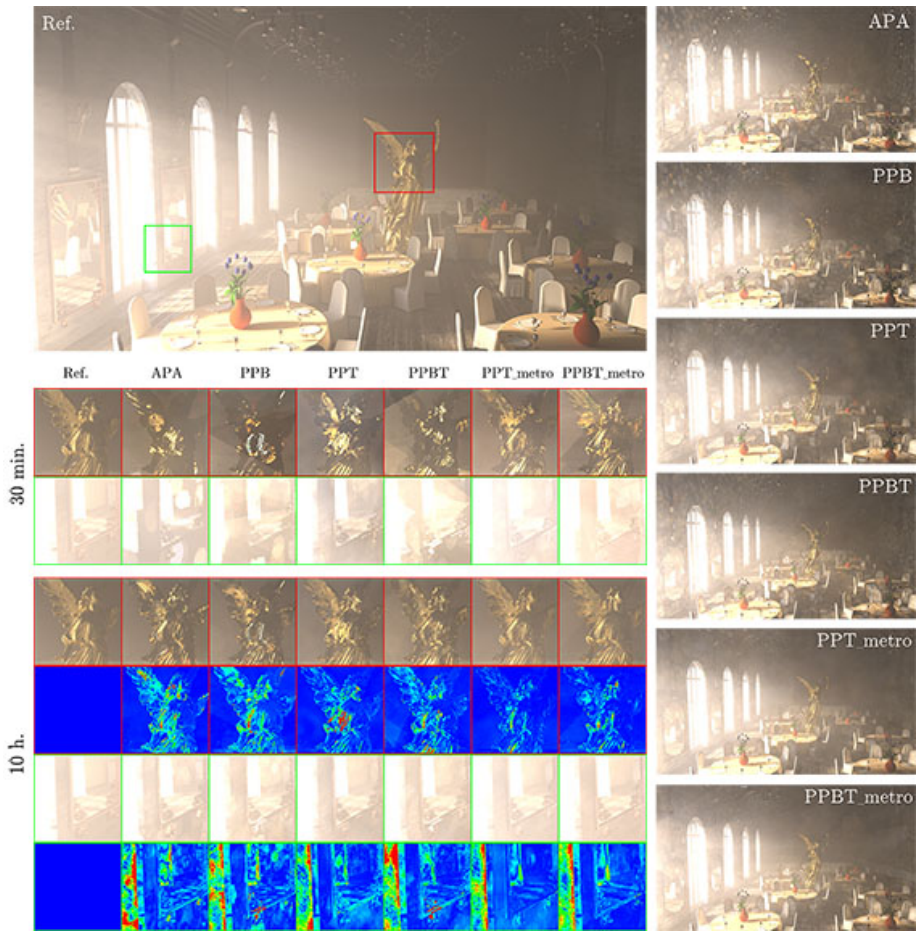


FIGURE 5.4 – La scène Breakfast hall (avec l'autorisation de Greg Zaal). (En haut à gauche) image de référence. (Colonne de droite) images obtenues avec les autres méthodes 10h. Détails sur les images obtenues avec les différentes méthodes après : 30min. et 10h. Et images de différences par rapport à l'image de références.



FIGURE 5.5 – La scène Kitchen (avec l'autorisation de Jay-Artist). (En haut à gauche) image obtenue avec *PPT_metro* (10h). (Colonne de droite) images obtenues avec les autres méthodes 10h. Détails sur les images obtenues avec les différentes méthodes après : 30min. et 10h.



FIGURE 5.6 – Même scène 3D avec (en bas) prise en compte dans la simulation des milieux participatifs - (en haut) sans milieux participatifs.

Chapitre 6

Post-traitements

Comme déjà mentionné dans les chapitres précédents, le résultat de calcul d'un logiciel de synthèse d'images physique est une image de luminance : pour chaque pixel de l'image est calculé la luminance pour un ensemble de longueurs d'onde. Cette luminance physique n'est pas directement compatible avec la représentation des formats d'image classiques. L'image de luminance physique n'est donc pas directement affichable sur un écran. En complément de l'image de luminance physique, le logiciel de synthèse d'image peut fournir des informations supplémentaires. Afin de pouvoir être affichée, l'image de luminance physique doit être transformée dans un format d'image classique. Elle doit donc subir des étapes de post-traitements :

- adaptation/réduction de la gamme de luminance (*tone mapping*)
- adaptation/réduction de la gamme de couleur (*color mapping* ou *color grading*)

Il est également possible d'appliquer d'autres traitements à l'image brute afin de modifier son aspect et ainsi fournir l'image finale souhaitée par le créateur. Les traitements appliqués peuvent demander des informations supplémentaires que peut calculer du moteur de synthèse d'image. Par exemple, Thomas Luft et al.[88] modifient le contraste de l'image en utilisant les informations de profondeur des pixels.

Dans ce contexte, nous avons travaillé sur la création d'un modèle d'attention visuelle qui exploite les informations provenant du rendu. Ce modèle évalue le point le plus probablement regardé par l'utilisateur immergé dans le monde 3D. La détermination du point observé permet de mettre en œuvre des post-traitements sur l'image qui nécessite cette information, l'ajout d'un flou de profondeur par exemple. Notre modèle d'attention visuelle est présenté dans la partie 6.1.

Nous avons également travaillé sur la réduction des gammes de luminance. Nous nous sommes focalisés les séquences d'images. Notre contribution porte sur le maintient de la cohérence spatiale et temporelle entre la séquence brute (luminances physiques) et la séquence aux gammes de luminances réduites (adaptée à l'affichage sur un écran classique). Cette contribution est détaillée dans la partie 6.2.

Nous avons également travaillé sur l'estimation de la couleur d'adaptation dans le contexte de l'illumination globale. L'estimation de la couleur d'adaptation permet de modifier avec précision l'équilibre colorimétrique de l'image et de restituer les couleurs telles quelles seraient perçues par un observateur. Ce travail est présenté dans la partie 6.3.

6.1 Attention visuelle et applications

Le cadre général de la thèse CIFRE avec France Telecom R&D de Sébastien Hillaire [10] était l'amélioration de la qualité visuelle des images produites en temps réel par illumination globale (jeux vidéos, réalité virtuelle, etc.). La qualité visuelle s'entend ici par restituer des images qui paraissent plus naturelles, qui procurent à l'utilisateur une meilleure immersion dans le monde 3D. Les moteurs de rendu ne tiennent généralement pas compte de la perception ni de l'attention visuelle humaine pour calculer l'image. Ceci résulte en des images qui n'apparaissent parfois peu ou pas naturelles. L'inconvénient de cette apparence non naturelle est un faible sentiment d'immersion des utilisateurs. Nous cherchons donc à déterminer automatiquement l'attention visuelle. Ainsi, après de la phase de rendu (dans les post-traitements), il est possible d'appliquer des effets visuels basés sur la perception humaine (comme la profondeur de champ) qui utilisent l'attention visuelle afin d'améliorer le retour visuel.

6.1.1 Introduction

La prise en compte de la perception visuelle humaine dans la production des images nécessite la détermination des éléments clés de cette perception. En premier lieu, du fait des aspects locaux de la perception, il est capitale de connaître le point de focalisation de l'utilisateur. Le point de focalisation est le point d'attention visuelle de l'utilisateur. Ils existent des solutions matérielles et logicielles afin d'estimer l'attention visuelle. Les solutions matérielles de suivi du regard peuvent être utilisés pour cette tâche. Cependant, ces solutions sont très souvent couteuses et intrusives, elles sont ainsi réservées pour des utilisations industrielles ou de recherche.

Les solutions logicielles consistent en la détermination algorithmique de distribution de l'attention visuelle humaine. Dans ce but elles simulent plusieurs composantes attentionnelles et cognitives du système visuel humain. Un point de focalisation final est calculé en fonction de la distribution de l'attention visuelle estimée. Cependant, ces modèles ont plusieurs limites fortes :

- manque de précision du fait du faible de composantes attentionnelles et cognitives prises en compte dans la simulation,
- trop générale car elles n'exploitent pas le scénario d'utilisation du monde 3D,
- trop particulière optimisée pour un cas d'utilisation précis.

En conséquence, nous avons travaillé sur les modèles d'attention visuelle et nous avons proposé un modèle d'attention visuelle plus performant qui utilise à la fois des informations bas niveaux issues de l'image (régions saillantes), des informations haut niveaux provenant du monde 3D (scénario) et des actions de l'utilisateur [10, 13, 12, 14].

De manière complémentaire, nous avons exploité notre modèle d'attention visuelle pour améliorer la qualité des images calculées. Nous avons notamment travaillé sur la profondeur de champ [15, 16, 17, 11]. Habituellement le logiciel de rendu ne simule pas la profondeur de champ due à l'optique de la caméra. L'absence de profondeur de champ confère aux images un aspect peu naturel et ne met pas en valeur l'élément principal de l'image. Il est possible de simuler précisément l'effet des optiques [89] mais cette approche n'est pas compatible avec un rendu en temps réel. Une autre approche consiste à simuler la profondeur de champ en appliquant un flou dépendant de la profondeur des pixels sur l'image calculée [90]. Cette technique nécessite de connaître la distance du point sur lequel se focalise le regard.

6.1.2 Modèles d'attention visuelle précédents

L'attention visuelle représente la capacité d'un être humain à se focaliser sur un objet il est composé de trois composants [91] :

1. *bottom-up* : simulation des réflexes visuel humain,
2. *top-down* : simulation des processus cognitifs,
3. intégration des composants *bottom-up* & *top-down*.

La composante *bottom-up* représente les réflexes du système visuel humain. Les caractéristiques principales de la composante *bottom-up* sont : un comportement local du à champ angulaire de 2 degrés [92], sensibilité aux zones saillantes [93, 94], aux zones scintillantes [91], à la profondeur [95] et aux mouvements [91].

La composante *top-down* d'attention visuelle représente les processus cognitif qui contrôle la stratégie utilisée pour analyser la scène. Elle dépend de la tâche à effectuer [96], de la mémoire [97], de l'accoutumance [98] et de l'habituation [98].

Même si les modèles d'attention visuelle *bottom-up* [93] arrivent à prédire les zones attirant le regard humain. Sans prendre en compte le scénario (la tâche à accomplir par l'utilisateur), ces modèles *bottom-up* ne sont pas efficace pour déterminer le point observé [99]. Par conséquent, il est nécessaire pour un modèle d'attention visuelle de simuler deux composantes *bottom-up* et *top-down*.

Peu de recherches ont été consacrées la prédiction du point observé en temps réel lors de l'exploration de monde 3D. Seul le modèle d'attention visuelle de Lee et al. [95] a été spécialement conçu pour ce but. Mais il ne renvoie pas le point focalisé

mais uniquement l'objet ou un ensemble d'objets, qui pourraient être observés par l'utilisateur.

6.1.3 Notre modèle d'attention visuelle

Notre objectif est de proposer un nouveau modèle d'attention visuelle complet (incluant des composants attentionnelles *bottom-up* et *top-down*) et temps-réel capable de calculer la position du point de focalisation d'un utilisateur navigant dans un monde 3D. Dans ce contexte, comparé à l'existant [95], notre modèle est le premier capable de calculer une position de focalisation au lieu d'un ensemble d'objets potentiellement regardés. Nous avons également conduit des expériences afin d'évaluer les performances de ce modèle.

L'architecture de notre modèle est constituée de trois composantes principales (Cf. figure 6.1) :

1. la composante *bottom-up* (en orange sur la figure 6.1) qui utilise les données issues du moteur de rendu afin de calculer la carte de saillance,
2. la composante *top-down* (en jaune sur la figure 6.1) qui utilise les informations provenant de la scène 3D, du scénario (c'est à dire de ce que l'utilisateur doit effectuer dans le monde 3D) et les entrées de l'utilisateur afin d'estimer l'attention visuelle haut niveau,
3. la composante d'intégration (en rouge sur la figure 6.1) qui cumule les informations *bottom-up* et *top-down* pour calculer le point focalisé.

La composante *bottom-up* (rectangle orange dans la figure 6.1) simule les réflexes du système visuel humain. Pour cela, nous calculons des cartes de caractéristiques pour chaque pixel de l'image : l'intensité, les couleurs antagonistes, la profondeur et le mouvement. Un opérateur de différence centre-région-contournantes est appliqué sur ces cartes de caractéristiques afin de calculer des cartes (appelé cartes de conspicuité ou cartes de perception) simulant la réponse à ces stimuli du système visuel humain. Finalement, toutes ces cartes de conspicuité sont intégrées et normalisées pour former la carte de saillance [93]. Plus la saillance d'un pixel est élevé, plus celui-ci a des chances d'attirer le regard. Cette carte de saillance est intégrée avec le résultat de la composante *top-down* afin de déterminer le point focalisé.

La composante *top-down* (rectangle jaune dans la figure 6.1) simule, quant-à-elle, les processus cognitifs se déroulant dans le cerveau. Cette composante attentionnelle est calculée par objet et non par pixel. Les précédents modèles utilisent une représentation liée au maillage des objets. Mais ce niveau de précision ne permet pas une estimation précise de l'endroit observé sur l'objet. De plus, l'évaluation est trop dépendante de la précision géométrique du maillage. Pour résoudre ce problème, nous proposons une représentation des objets basée sur les éléments de surface, nommés *surfels*. Par cette méthode, initialement introduite dans le domaine du rendu [100], chaque objet est virtuellement subdivisé en plusieurs petits éléments de surfaces. Pour chaque élément de surface, les valeurs attentionnelles suivante sont calculées :

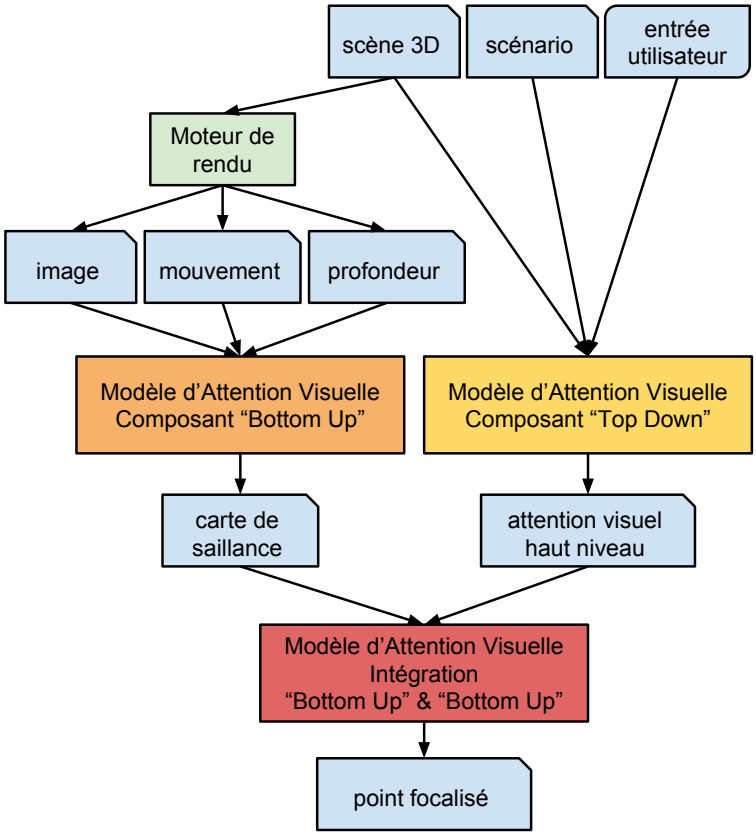


FIGURE 6.1 – Architecture de notre Modèle d'Attention Visuelle : composants *bottom-up*, *top-down* et intégration *bottom-up* & textittop-down.

- l'habituation [98] : l'attention privilégie les objets nouveaux et négliges les objets déjà connus et observés,
- le contexte spatial [95] : l'attention se porte sur les objets qui se rapprochent de l'observateur.

La composante *top-down* prend également en compte, premièrement, le fait que l'attention visuelle se porte majoritairement au centre de l'écran et, deuxièmement, la déviation de l'attention en fonction des entrées de l'utilisateur (dépendance vis à vis de la vitesse de rotation). Enfin la composante *top-down* prend également en compte le scénario en associant un poids sémantique plus important aux objets ayant une importance vis à vis du scénario. Ces différents éléments sont fusionnés pour donner une attention visuelle de haut niveau.

Les résultats des composants *bottom-up* et textittop-down sont intégrés/combinés (rectangle rouge dans la figure 6.1) afin de calculer le point focalisé. Ce dernier cor-

respond au point ayant la plus grande valeur d'attention visuelle.

Nous avons mis en œuvre notre modèle d'attention visuelle sur carte graphique en GLSL [101] afin d'atteindre des performances compatibles avec le temps-réel. Pour optimiser les temps de calcul, les cartes de caractéristiques sont calculées avec une résolution 256x256 pixels. Les calculs des valeurs maximales et moyennes des cartes de caractéristiques sont calculées par sous échantillonnage récursif. Les performances atteintes pour un environnement 3D simple (Cf. figure 6.2) sont : 170 images par seconde pour le rendu monde 3D sans modèle d'attention visuelle et 145 images par secondes avec le modèle d'attention visuelle. L'exécution du modèle d'attention visuelle dure : 1,01 millisecondes (soit 0,6 % du temps de rendu). Ces performances sont obtenues sur l'environnement matériel suivant : Intel Core 2 2.5Ghz, nVidia GeForce3700M et 4Gb de RAM. Notre modèle d'attention visuelle est donc compatible avec les applications 3D temps comme les jeux vidéo ou la réalité virtuelle.



FIGURE 6.2 – Monde 3D utilisé pour les tests du Modèle d'Attention Visuelle : (à gauche) vue de l'environnement 3D - (à droite) rendu pour l'utilisateur.

6.1.4 Résultat du modèle d'attention visuelle

Nous avons conduit des expériences avec des utilisateurs afin d'évaluer les performances de notre modèle d'attention visuelle. Les utilisateurs naviguaient dans des mondes 3D, soit librement, soit avec une tâche à accomplir (ramasser des objets particulier). Nos résultats sont comparés avec le modèle d'attention visuelle de Lee et al. [95]. Nos résultats montrent une bonne estimation du point de focalisation : 75,6% du temps avec une précision de 8 degré d'angle visuel contre 39,7% pour le modèle de Lee et al. Nous obtenons donc un gain de plus de 90% pour le calcul de du point regardé par l'utilisateur à un instant donné. Nous avons aussi comparé les performances des composantes *bottom-up* et *top-down*. Les résultats montrent que les performances de chacun des composantes prises séparément sont significativement inférieures aux performances du modèle complet (qui intègre les deux composants).

Ces résultats montrent que la simulation de l'attention visuelle humaine est suffisamment fiable pour de nombreuses applications. Nous avons également montré le bénéfice de combiner les composants *bottom-up* et *top-down*.

6.1.5 Algorithme de flou de profondeur avec autofocus

En photographie ou au cinéma, le flou de profondeur est du aux optiques des caméras, il est utilisé par le créateur d'images pour attirer l'attention de l'observateur vers la zone nette (le regard fuit naturellement les zones floues). La prise en compte du flou de profondeur permet de créer des images qui paraissent plus naturelle et qui mettent en valeur le sujet principal [90, 102]. L'utilisation des cartes graphiques permet l'utilisation des algorithmes de flou de profondeur en temps réel [103, 104]. Dans le cadre des applications temps réels, la détermination de la zone où faire la mise au point est la problématique clé. Nous proposons pour cela une technique qui simule les systèmes d'autofocus des caméras.

La problématique de la détermination de la zone où faire le net n'est pas directement celle d'un système d'attention visuelle. En effet, il ne s'agit de déterminer où regarde l'observateur mais au contraire de diriger son regard vers la zone pertinente. Il s'agit donc de contrôler et non de déterminer le point où regarde l'utilisateur. Le système d'autofocus doit déterminer quelle est la zone de l'image où l'on doit diriger le regard de l'observateur. Les deux problématiques sont liées, mais un modèle d'attention visuelle ne répond pas directement à la problématique donnée.

Notre système d'autofocus intelligent reprend une partie de la composante cognitive de notre modèle d'attention visuelle : la partie concernant le scénario. Une valeur sémantique est associée aux objets en fonction de leur importance dans le scénario. La seule valeur sémantique ne suffit pas car plusieurs objets de même importance peuvent être visibles en même temps sur l'écran. Nous privilégions alors la zone centrale de l'écran correspondant à 25% de l'écran. Cela est conforme avec le fait que les utilisateurs de jeux vidéo ont leur regard qui se focalise très majoritairement (82%) sur le centre de l'écran [105].

La détermination de la distance de focalisation se fait en plusieurs étapes. Premièrement, avec le rendu de la scène nous calculons la carte de profondeurs et également une carte contenant la valeur sémantique des objets visibles à l'écran. La valeur sémantique représente l'importance des objets de la scène en fonction du contexte et de la tâche courante.

Deuxièmement, à l'intérieur de la zone d'autofocus (correspondant aux 25% centraux de l'image), nous calculons la moyenne des profondeurs des pixels pondérée par la valeur sémantique des objets. Ainsi, même si un objet sémantiquement important n'occupe que peu de pixels de la zone d'autofocus, son poids dans la moyenne sera augmenté. La moyenne pondérée des distances est utilisée comme la distance où faire le net (Cf. Figure 6.3).

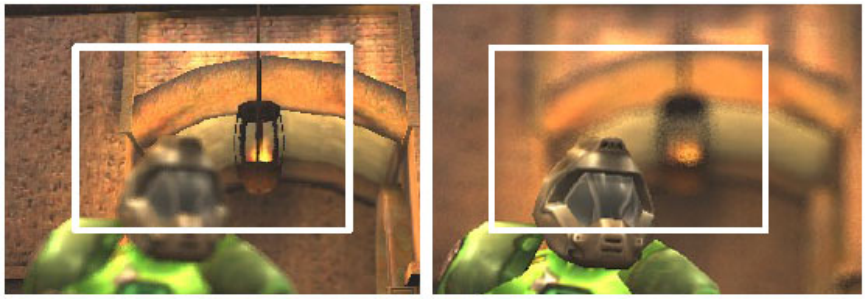


FIGURE 6.3 – Zone d'auto-focus (rectangle blanc) et influence du poids sémantique : (à gauche) auto-focus sans prise en compte de la carte sémantique - (à droite) auto-focus avec prise en compte de la sémantique : le net est fait sur le personnage.

Troisièmement, cette distance est filtrée par un filtre passe-bas afin de simuler le phénomène d'accommodation du système visuel humain et pour éviter les changements brusques de distance de focalisation.

La dernière étape consiste à calculer l'image floutée. Pour cela, nous utilisons la méthode proposée par Riguer et al. [106]. Nous avons amélioré cette méthode en appliquant une rotation aléatoire par pixel du noyau d'échantillonnage afin de remplacer les artefacts de bande de couleur par un bruit de haute fréquence automatiquement filtré par le système visuel humain (Cf. Figure 6.4).



FIGURE 6.4 – Flou de profondeur : (à gauche) flou de profondeur sans rotation du noyau d'échantillonnage - (au milieu) flou de profondeur avec rotation du noyau d'échantillonnage - (à droite) image finale avec flou de profondeur.

6.1.6 Conclusion

Dans le cadre de ces travaux, nous avons proposé un modèle d'attention visuelle qui intègre des composantes *bottom-up* et *top-down*. Ce modèle inclut une nouvelle représentation des objets visuels basée sur les éléments de surface. Cette nouvelle représentation est plus appropriée pour le calcul de certaines composantes attentionnelles telle que l'habituation. De plus, ce modèle compatible avec le temps réel est plus précis de 90% par rapport aux modèles précédents. Avec une précision de l'ordre de 70%, il est utilisable pour une grande gamme d'applications.

Nous avons également travaillé sur la perception visuelle dans le but d'améliorer de contrôler la qualité des images produites en illumination globale en temps réel. Nous avons en autre proposé un autofocus intelligent (qui exploite les connaissances sémantiques de la scène relativement au scénario) qui permet de déterminer la distance de focalisation et permet ainsi d'introduire un flou de profondeur aux images produites. Le flou de profondeur focalise l'attention visuelle de l'observateur sur les objets importants de la scène.

Cette approche consistant d'une part à déterminer une caractéristique de la perception des images (ici l'attention visuelle) et, d'autre part, à proposer des outils permettant de contrôler cette même caractéristique (ici le système d'autofocus intelligent) est une approche clé pour le contrôle des images produites. Disposant d'un modèle d'évaluation de l'attention visuel, il est possible de l'utiliser afin de contrôler l'éclairage dans le but que l'éclairage mette en valeur les objets importants de la scène 3D.

6.2 Adaptation des gammes de luminances pour les séquences d'images

L'état de l'art propose un grand nombre d'opérateurs de réduction des gammes de luminance afin de transformer les images HDR en images LDR compatibles avec les dispositifs de restitution classiques : écran, projection et impression [30, 31]. Chaque opérateur de compression propose un résultat différent et cohérent en fonction de ses objectifs (Cf. figure 3.4). L'opérateur de réduction possède un certain de paramètres, des paramètres externes permettant à l'utilisateur de contrôler le rendu de l'image LDR finale et des paramètres internes calculés directement à partir des caractéristiques propres de l'image HDR.

Malgré les avancés en matière de capture de vidéos HDR [107, 108], peu de travaux ont porté spécifiquement sur les séquences d'images HDR et la compression de leur gamme de luminance pour les transformer en séquences LDR. En effet, les séquences d'images sont généralement vues comme une simple suite d'images HDR sur laquelle il suffit d'appliquer l'opérateur de compression de luminance sur chaque image. Dans le cadre de la thèse CIFRE de Ronan Boitard avec Technicolor, nous avons mis en évidence les problématiques spécifiques de la compression des gammes

de luminance pour les séquences d'images [19, 20] et nous avons proposé des méthodes pour régler ses problèmes indépendamment de l'opérateur de compression utilisé [21, 24, 23]. Nous avons également montré le bénéfice de la cohérence spatio-temporelle pour la compression des vidéos HDR [22].

6.2.1 Problématique de la réduction des gammes de luminances pour les vidéos

Nous proposons de classer les problèmes liés à la réduction des gammes de luminances des vidéos HDR en trois grandes catégories :

- Scintillement : changement rapide de la luminance LDR apparaissant entre des images LDR successives,
- Altération de la cohérence du contraste : la séquence LDR ne reproduit pas les changements de conditions d'illumination de la séquence HDR,
- Altération de la cohérence des objets : au cours de la vidéo, un objet change de luminance dans la vidéo LDR alors que sa luminance reste constante dans la vidéo HDR.

Ces problèmes arrivent lors de l'application d'un opérateur de réduction des gammes de luminance sur une séquence d'images. Ils proviennent de la logique de fonctionnement des opérateurs. En effet, les opérateurs de réduction des gammes de luminance cherchent à adapter l'intervalle des luminances HDR $[L_{min}, L_{max}]$ sur l'intervalle LDR $[l_{min}, l_{max}]$ (classiquement $[0, 255]$). Or l'intervalle des luminances HDR peut changer au cours du temps : $[L_{min}(i), L_{max}(i)]$ avec i le numéro de l'image dans la séquence. Si la logique de l'opérateur est, pour simplifier, de faire correspondre $L_{min}(i)$ avec $l_{min} = 0$ et respectivement $L_{max}(i)$ avec $l_{max} = 255$, alors il est logique qu'à une valeur de luminance HDR donnée L^k, il correspond des valeurs de luminances LDR différentes $l^k(i)$ dans la séquence réduite. Cela entraine une incohérence temporelle entre les luminances de la séquence HDR et LDR.

En résumé, l'opérateur de réduction fonctionne comme un caméscope en mode automatique qui s'adapte en continue aux gammes de luminance de ce qui est filmé. Néanmoins les trois types d'altération (scintillement, altération de la cohérence du contraste et altération de la cohérence des objets) proviennent de conditions spécifiques différentes.

Le scintillement est un changement soudain des luminances LDR survenant sur une image d'une séquence alors qu'il n'y a pas de changement visible dans la séquence HDR. La figure 6.5 montre trois images LDR consécutives avec un cas de scintillement. L'image intermédiaire est plus lumineuse que l'image antérieure et postérieure. Ce genre de problème provient du fait que l'opérateur de réduction s'adapte en fonction de valeurs statistiques calculées sur l'image HDR. Un petit changement dans les valeurs statistiques calculées entre deux images HDR consécutives peut entraîner un changement important dans la séquence LDR.



FIGURE 6.5 – Scintillement dans la séquence LDR : exemple de trois images consécutives, l'image du milieu est plus lumineuse.

Les altérations de la cohérence du contraste proviennent de changements dans les conditions d'illumination, par exemple l'allumage ou l'extinction d'une source de lumière. Ces changements d'illumination se traduisent par des changements dans l'amplitude de l'intervalle des luminances HDR. L'algorithme de réduction des gammes de luminance s'adaptant à l'intervalle des valeurs HDR, il ne conserve pas dans la séquence LDR produite ces changements de luminances. Il perd ainsi la cohérence temporelle des conditions d'illumination et par conséquence l'intention artistique du créateur. L'image HDR la plus lumineuse peut ne pas apparaître comme la plus lumineuse dans la séquence LDR produite. La figure 6.6 illustre ce problème : dans la séquence HDR, les conditions d'illumination change et les valeurs de luminances deviennent plus faibles (images plus sombres). La séquence LDR produite (l'opérateur utilisé est [109]) ne reflète pas ce changement de conditions d'illumination.

Les altérations de la cohérence des objets correspondent au manque de cohérence temporelle entre les luminances des images HDR et celles des images LDR. Cela se traduit par le fait qu'un objet ayant une luminance constante dans la séquence HDR a une luminance non constante dans la séquence LDR produite. Ce type de manque de cohérence temporelle est similaire à la problématique de la stabilisation tonale décrit dans [110]. La figure 6.7 illustre le problème des altérations de la cohérence des objets dans les séquences LDR produites. Elle montre que la luminance des objets dans la séquence LDR produite ne semble pas cohérente avec l'évolution de leur luminance HDR (l'opérateur utilisé dans cet exemple est [109]). La luminance de la partie inférieure du point est constante dans la séquence HDR (ligne du haut dans la figure 6.7) alors qu'elle apparait comme baissant dans la séquence LDR (ligne du bas à droite dans la figure 6.7).

Les schémas 6.8 résument les trois types d'altération de la cohérence des luminosités dans les séquences LDR lors de réduction des gammes de lumimances :

- Scintillement (en haut) : des petits changements dans la séquence HDR entrainent des grands changements dans la séquence LDR,
- Altération de la cohérence du contraste (au milieu) : un changement de condition d'illumination dans la séquence HDR n'est pas reflété dans la séquence LDR,
- Altération de la cohérence des objets (en bas) : les changements de la lumi-



FIGURE 6.6 – Altération de la cohérence du contraste : (en haut) Les images HDR en fausse couleur (bleu : faible valeur de luminance, rouge : grande valeur de luminance), (en bas) Les images LDR correspondantes - Le changement de condition d'illumination entraine des images HDR plus sombres, ce changement n'est pas respecté dans les images LDR produites.

nance des objets dans la séquence HDR ne sont pas restitués dans la séquence LDR, les luminances HDR et LDR d'un objet ne semblent pas corrélées.

6.2.2 Travaux précédants connexes

Peu de travaux de recherche se sont focalisés sur la problématique du maintient de la cohérence temporelle lors de la réduction des gammes de luminance des séquences HDR. Seul le problème du scintillement apparaissant dans les séquences produites a été traité.

L'approche utilisée pour réduire le scintillement (changement rapide de la luminance dans les images LDR produites) consiste à proposer des stratégies de filtrage passe bas temporel. Le filtrage peut se faire sur les différents éléments intervenant dans la réduction des gammes de luminances :

- sur la courbe définissant la réduction des gammes de luminances [111],
- sur un paramètre de contrôle de l'opérateur de réduction (moyenne logarithmique de l'image HDR [112, 113]),
- sur l'image LDR produite [114, 115].

Ces techniques sont efficaces pour éliminer les problèmes de scintillement. Mais elles ne réduisent pas les altérations temporelles de la cohérence du contraste et des

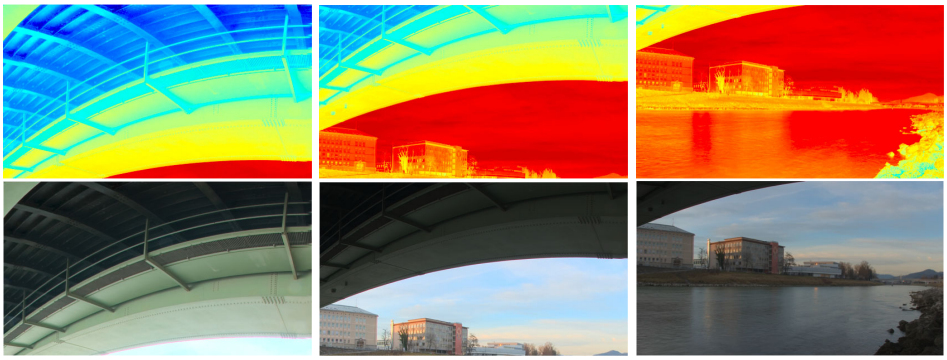


FIGURE 6.7 – Altération de la cohérence des objets : (en haut) Les images HDR en fausse couleur (bleu : faible valeur de luminance, rouge : grande valeur de luminance), (en bas) Les images LDR correspondantes - la luminance de la partie inférieure du pont évolue dans la séquence LDR alors qu'elle est constante dans la séquence HDR.

objets.

6.2.3 Algorithmes de maintient de la cohérence temporelle

Nous avons proposé une approche innovante pour maintenir la cohérence temporelle lors de la réduction des gammes de luminances des séquences HDR en séquence LDR [116, 24]. Dans ce travail nous faisons les hypothèses suivantes. Premièrement les séquences HDR ont été étalonnées : leur contenu et en particulier les changements de luminosités sont volontaires et correspondent aux intentions des créateurs. Ces changements doivent donc être reproduits au mieux dans la séquence LDR. Deuxièmement, la séquence est disponible dans son entier, elle correspond à l'entité temporelle durant laquelle la cohérence doit être maintenue. D'un point de vue technique, nous avons décidé de proposer des techniques indépendantes de l'algorithme de réduction utilisé, notre solution peut donc s'appliquer quelque soit l'algorithme utilisé pour réduire les gammes de luminance.

Première approche pour la conservation de la cohérence temporelle de la luminosité

Nous avons donc fait l'hypothèse que les changements de luminosités dans la séquence HDR doivent se retrouver dans la séquence LDR : nous voulons maintenir la cohérence temporelle de la luminosité. La *valeur clé* κ (*key value*), la moyenne logarithmique des luminances est un bon indicateur de la luminosité de l'image [109] :

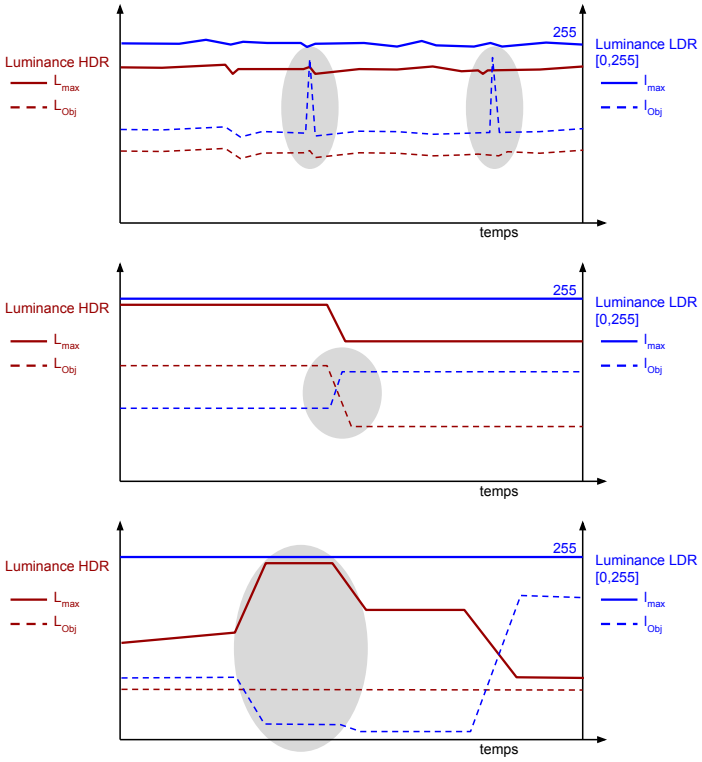


FIGURE 6.8 – Les différents types d'altération de la cohérence temporelle - (en haut) scintillement, (au milieu) altération de cohérence du contraste, (en bas) altération de la cohérence temporelle des objets.

$$\kappa = \exp\left(\frac{1}{N}\sum_{i=1}^{N}\log(\epsilon + L(i))\right) \tag{6.1}$$

avec $L(i)$ la luminance HDR du pixel i, N le nombre de pixels, ϵ une petite valeur (10^{-6}). Nous avons proposé d'exprimer la cohérence temporelle de la luminosité comme :

$$\frac{\kappa^{j}_{LDR}}{\kappa^{max}_{LDR}} = \frac{\kappa^{j}_{HDR}}{\kappa^{max}_{HDR}} \tag{6.2}$$

avec κ^{j}_{HDR} la *valeur clé* de l'image j dans la séquence HDR, respectivement κ^{j}_{HDR} dans la séquence LDR, κ^{max}_{HDR} la *valeur clé* de valeur maximale dans la séquence HDR et κ^{max}_{LDR} la *valeur clé* correspondante dans la séquence LDR.

A partir de ce ratio de *valeurs clés*, nous avons proposé un premier algorithme de

maintient de la cohérence temporelle [116] qui met à l'échelle les luminances LDR l^j de j-ème image de la séquence afin quelle respectent le ratio des *valeurs clés* :

$$l^j_{BC} = \left(u + (1-u)\frac{\kappa^{max}_{LDR}\kappa^j_{HDR}}{\kappa^{max}_{HDR}\kappa^j_{LDR}} \right) l^j \tag{6.3}$$

Cette première approche donne de bon résultat en terme de cohérence temporelle de la luminosité mais elle détériore le contraste des images LDR corrigées. En gardant la même logique, nous avons amélioré l'algorithme en conservant la cohérence par zones. Les zones décrivent la distribution des valeurs de luminances par image et au cours du temps.

Conservation par *zones* de la cohérence temporelle de la luminosité

Les zones sont définies en deux étapes : premièrement une segmentation des luminances par image. À chaque segment est associée sa *valeur clé*, deuxièmement une segmentation en *zones* des *valeurs clés* des segments de chaque images de la séquence. La cohérence temporelle (6.3) est alors appliquée par *zones*.

La segmentation des images de la séquence est effectuée dans le domaine des luminances. La segmentation dans le domaine des luminances a pour intérêt de fournir des segments de luminances directement corrélés avec la luminosité dont nous voulons assurer la cohérence. L'information spatiale qui pourrait extraites par une segmentation spatiale n'a pas d'intérêt par rapport à notre objectif. La segmentation est réalisée comme suit :

Algorithm 2 Calcul des segments()

1: **for** chaque image de la séquence **do**
2: Calcul de l'histogramme des luminances (domaine logarithmique) ;
3: Détermination des maximums ;
4: Filtrage des maximums (suppression des maximums trop proches) ;
5: Définition des limites des segments comme les minimas entre les maximums ;
6: Calculs des *valeurs clés* par segment ;
7: **end for**

La première étape correspond au calcul de l'histogramme des luminances exprimées en logarithme, l'histogramme est divisé en w_i intervalles. A partir de cet histogramme, nous trouvons les maximums locaux supérieurs à un seuil : L^S. Les maximums locaux sont filtrés, seuls sont conservés les maximums dont l'écart de luminances est supérieur à un seuil ΔL^S. Le minimum entre deux maximums nous donne la limite des segments. La *valeur clés* κ^i du i-ème segment est calculé. Nous obtenons donc pour chaque image j de la séquence et pour chaque segment i une *valeur clé* κ^j_i.

Les *valeurs clés* κ_i^j pour un segment i ne sont pas continues dans le temps. En effets, l'apparition ou la disparition d'objets dans la vidéo ainsi que des changements de conditions d'illumination entrainent des discontinuités temporelles importantes. Comme nous voulons des *zones* qui reflètent au mieux la distribution des luminances au cours de la séquences, nous appliquons donc l'algorithme 2 aux *valeurs clés* κ_i^j. Nous obtenons ainsi n *zones*. Chaque *zone* z_i est défini par ses limites $b_{(i-1,i)}$ et $b_{(i,i+1)}$ constante tout au long de la séquence et une *valeur clé* $\kappa_{z_i}^f$ pour la z_i pour l'image f de la séquence. L'algorithme de maintient de la cohérence 6.3 est alors appliqué par *zone*.

Chaque pixel p^f, de l'image f de la séquence, appartient a une *zone* z_i :

$$b_{(i-1,i)} \leq \log(L^{p^f}) \leq b_{(i,i+1)}$$

avec L^{p^f} la luminance HDR du pixel p de l'image f, et $b_{(i-1,i)}$ et $b_{(i-1,i)}$ les frontières de la *zone* z_i. $\kappa_{z_i,HDR}^f$ et $\kappa_{z_i,LDR}^f$ sont les *valeurs clés* HDR et LDR de la *zone* z_i pour l'image f et $\kappa_{z_i,HDR}^{max}$ et $\kappa_{z_i,LDR}^{max}$ les *valeurs clés* HDR et LDR maximales (pour toutes les images de la séquence) de la *zone* z_i. La luminance LDR l^{p^f} est mise à l'échelle pour conserver la cohérence temporelle :

$$l_{ZBC}^{p^f} = \left(u + (1-u)\frac{\kappa_{z_i,LDR}^{max}\kappa_{z_i,HDR}^f}{\kappa_{z_i,HDR}^{max}\kappa_{z_i,LDR}^f} \right) l^{p^f} \tag{6.4}$$

Cette équation peut être réécrite comme :

$$l_{ZBC}^{p^f} = s_i^f l^{p^f} \tag{6.5}$$

avec s_i^f, le facteur d'échelle :

$$s_i^f = u + (1-u)\frac{\kappa_{z_i,LDR}^{max}\kappa_{z_i,HDR}^f}{\kappa_{z_i,HDR}^{max}\kappa_{z_i,LDR}^f} \tag{6.6}$$

Pour éviter les discontinuités des pixels p (dont la luminance HDR est L_p^f) proches d'une frontière (de valeur de luminance HDR b_z) entre deux *zones* (z_m et z_n), les luminances de ces pixels sont mis à l'échelle par une pondération des deux facteurs d'échelle s_m^f et s_n^f correspond à ceux deux *zones* proches. La pondération

est donnée par :

$$s^f_{m,n} = s^f_m w_m + s^f_n w_n \tag{6.7}$$

avec (6.8)

$$w_m = \frac{\exp\left(-\frac{(\log(L^{p^f})-log(b_z-\delta/2))^2}{2\sigma^2}\right)}{s^f_m + s^f_n} \tag{6.9}$$

$$w_n = \frac{\exp\left(-\frac{(\log(L^{p^f})-\log(b_z+\delta/2))^2}{2\sigma^2}\right)}{s^f_m + s^f_n} \tag{6.10}$$

$$\sigma = \frac{\delta}{2\sqrt{2\log(3)}} \tag{6.11}$$

δ est la largeur de la frontière, c'est un paramètre défini par l'utilisateur. La figure 6.9 montre les fonctions de pondération w_m et w_n.

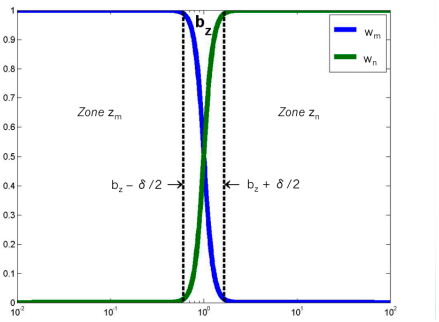


FIGURE 6.9 – Fonction de pondération des facteurs d'échelle.

6.2.4 Résultats

Nous présentons dans ce document deux résultats de notre algorithme de maintien par *zones* de la cohérence temporelle de la luminosité (dénoté ZBC). D'autres résultats sont disponibles dans [24].

La première séquence est la séquence HDR *Under Bridge* (Cf. figure 6.10). Ses gammes de luminances sont réduites par l'opérateur de Li [117] qui est un opérateur local. L'application de notre algorithme ZBC permet de garder une meilleure cohérence temporelle de la luminosité dans les images LDR. Dans la figure 6.10, la luminosité de la partie correspondant au dessous du pont reste pratiquement constante. L'algorithme ZBC maintient également le contraste de l'image. Dans la figure 6.10, la partie de l'image correspond aux bâtiments et au ciel reste très

contrasté.

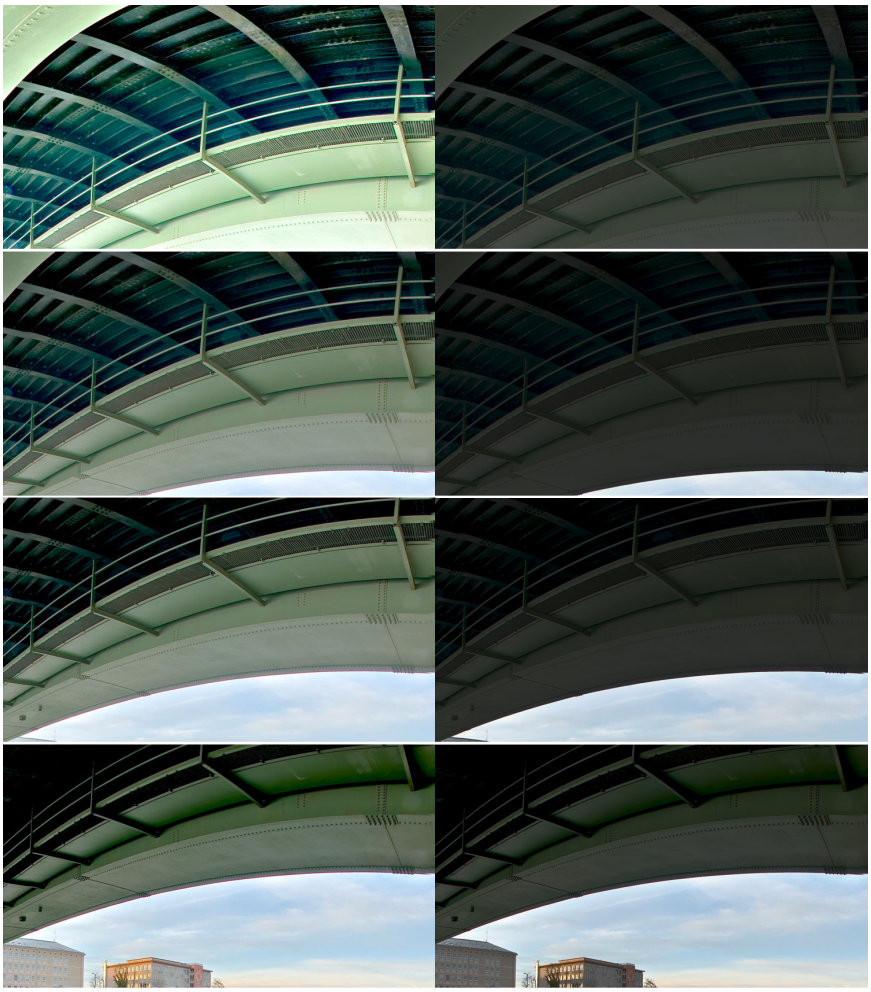


FIGURE 6.10 – La séquence HDR *Under Bridge* - (à gauche) des images de séquence LDR obtenu avec l'opérateur de réduction de Li [117] - (à droite) les images LDR avec ZBC : la cohérence temporelle de la luminosité est maintenue avec la séquence HDR et le contraste est conservé.

La deuxième séquence est la séquence HDR du tunnel (Cf. figure 6.11). L'algorithme de réduction des gammes de luminances est [109]. Dans cet exemple, les conditions d'éclairement du véhicule change. Alors que dans la séquence LDR produite initialement par [109], la dernière image de la séquence (colonne de gauche, en bas dans la figure 6.11) qui est la plus sombre de la séquence HDR apparait comme plus lumineuse que les précédentes, grâce à l'utilisation de ZBC elle retrouve une

cohérence temporelle de luminosité par rapport aux images précédentes.



FIGURE 6.11 – La séquence HDR *Under Bridge* - (à gauche) des images de séquence LDR obtenu avec l'opérateur de réduction [109] - (à droite) les images LDR avec ZBC : la cohérence temporelle de la luminosité est conservée malgré le changement de conditions d'illumination sur le véhicule principal.

Nous avons également montré que ZBC permet de conserver les fondus aux noir et les fondus aux blancs, présents dans les séquences HDR, dans les séquences LDR produites et corrigés par ZBC. Nous avons également des tests subjectifs qui montre que dans la majorité des cas les vidéos LDR corrigées par ZBC sont jugées comme ayant une meilleure qualité [24].

6.2.5 Conclusion et perspectives

Nous avons proposé une méthode permettant d'assurer la cohérence spatiale et temporelle de la luminosité dans les séquences LDR produites par l'application d'un algorithme de réduction des gammes de luminance. Notre méthode conserve également le contraste. Et elle peut s'appliquer sur des séquences HDR ayant un montage comme des fondus au noir ou au blanc. En résumé, cette méthode, publiée dans une revue internationale [24], a les caractétiques suivantes :

- elle est indépendante de l'opérateur de réduction des gammes de luminances choisi,
- elle s'applique comme un post-traitement sur la séquence LDR,
- elle assure la cohérence des luminosités entre la séquence HDR et la séquence LDR,
- elle fonctionne sur les fondu au noirs et au blanc.

La principale limite de notre approche est qu'elle nécessite la totalité de la séquence sur laquelle souhaite réduire la gamme de luminance. Elle ne fonctionne pas pour une diffusion temps réel d'un évènement filmé en HDR.

Nous travaillons sur une adaptation de cette méthode pour les mondes 3D interactifs (jeux vidéo, réalité virtuelle) rendu par un moteur d'illumination globale. Les contraintes supplémentaires de ce cadre d'application sont : (1) en fonctionnement en temps réel des post-traitements (mise à l'échelle des luminances HDR par *zones*), (2) les images HDR ne sont pas connues, ni disponibles (monde interactif) mais nous faisons l'hypothèse que le monde 3D est connu.

6.3 Gestion de la couleur

La première étape de la gestion/réglage de la couleur dans la création d'image est le réglage de la balance des couleurs. Il permet soit de restituer les couleurs telles qu'elles seraient perçues dans un être humain regardant directement la scène, soit volontairement modifier cette perception en donnant une colorimétrie plus froide (déviation des teintes vers le bleu) ou plus chaudes (déviation vers le rouge). Dans tous les deux cas, il est nécessaire d'estimer la couleur d'adaptation.

6.3.1 Notion de couleur d'adaptation

Les modèles d'adaptation chromatique prédisent la couleur perçue ou adaptée x_a, d'un échantillon de couleur x sous les conditions d'illumination x_W (couleur d'adaptation). Ces modèles peuvent se résumer à une fonction CAT de deux variables x and x_W :

$$x_a = CAT\,(x, x_W) \tag{6.12}$$

Ainsi les modèles d'adaptation chromatique procèdent en deux étapes : premièrement, l'estimation de la couleur d'adaptation x_W, puis deuxièmement, le calcul

de la couleur adaptée x_a (Cf. Figure 6.12). La qualité de l'adaptation dépend essentiellement de la précision de l'estimation de la couleur d'adaptation.

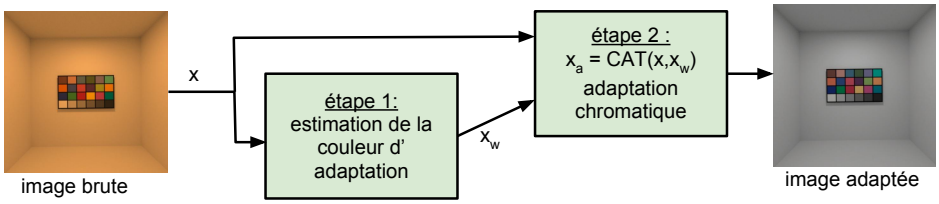


FIGURE 6.12 – Les deux étapes de l'adaptation chromatique

Dans le cas de la création d'images (photographie, cinéma, images de synthèse), le but n'est pas forcément de restituer les couleurs perçues, il peut être de donner une dominante de couleurs à l'image, également appelée température de couleur : ambiance froide avec une dominante des couleurs vers le bleu ou une ambiance chaude avec une dominante de couleur vers le rouge. Cette ambiance de couleur est donné par la couleur de blanc de référence : $[R_r, G_r, B_r]$. La transformation, dans sa forme linéaire, de couleurs pour obtenir cette dominante de couleur donnée s'exprime alors :

$$\begin{bmatrix} R_a \\ G_a \\ B_a \end{bmatrix} = \begin{bmatrix} R_r/R_W & 0 & 0 \\ 0 & G_r/R_W & 0 \\ 0 & 0 & B_r/R_W \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{6.13}$$

avec $[R_W, G_W, B_W]$ la couleur d'adpation, $[R, G, B]$, la valeur initiale du pixel et $[R_a, G_a, B_a]$ la couleur du pixel après choix de la température de couleur.

Dans ce cas, la couleur $x_r = [R_r, G_r, B_r]$ correspondant à la température de couleur choisit pour l'image adaptée. C'est un paramètre supplémentaire de la fonction CAT :

$$x_a = CAT\left(x, x_W, x_r\right) \tag{6.14}$$

Ce paramètre x_r est une paramètre de contrôle fournit par le créateur. Dans le cas où le but est de restituer les couleur perçues x_r vaut $[1, 1, 1]$ dans l'espcae de couleur LMS.

6.3.2 Estimation de la couleur d'adaptation

L'estimation de de la couleur d'adaptation est une problématique commune à différents domaines de recherche : traitements d'images, vision par ordinateur, apparence des couleurs, et synthèse d'images. Cette problématique commune provient du fait qu'un capteur d'appareil photographique (ou de caméra) ou que le moteur de synthèse d'image produit comme résultat brut une mesure des luminances physiques. Les couleurs physiques de ses luminances physiques sont le résultat des

interactions entre la couleur de l'éclairement arrivant sur un objet et la « couleur intrinsèque » de l'objet (réflectance). C'est à dire la couleur de l'objet sous une lumière blanche. L'estimation de la couleur d'adaptation puis l'adaptation chromatique permet donc soit de retrouver la « couleur intrinsèque » des objets, c'est à dire de restituer les couleurs telles qu'elles seraient perçues par un observateur (humain), soit de donner une température de couleur à une image.

De nombreuses méthodes ont été proposées, elles sont classifiées selon trois grandes approches [118] :

1. physique (*dichromatric models* [119, 120, 121, 122], *maxRGB* [123], *grey-world* [124, 125]),
2. statistique (*gamut mapping* [126], *bayesian* [127]),
3. haut niveau (*high level visual information*) [128, 125].

Ces diverses méthodes sont comparées dans [129].

Les méthodes statistiques et de haut niveau déterminent habituellement la couleur d'adaptation à partir d'un ensemble fini de sources de lumières correspondant à des sources artificielles (sources au tungstène, fluorescente, halogène, etc.) et à des conditions d'éclairage naturel (soleil à midi, levé dou couché de soleil, ciel nuageux, etc.). Cette base de données de conditions d'éclairage type n'a pas de sens dans le domaine de l'illumination globale. En effet le créateur peut définir les sources de lumières (dont les caractéristiques en terme de couleur) sans la contrainte de se limiter à un ensemble de sources réelles. Cette classe d'approches statistiques et de haut niveau ne s'applique pas dans le cadre de l'illumination globale.

Les méthodes de l'approche physique n'ont pas de conditions d'utilisation incompatible avec contexte de l'illumination globale. Ces méthodes ont des performances similaires [130, 131]. Cependant elles échouent, elles ne fournissent pas une estimation correcte de la couleur d'adaptation correcte, dans certains cas de tests et notamment dans le cas appelé « monde monochromatique » (*world of one reflectance*) [132]. Ce cas de test corrspond à deux scènes : (1) une scène complétement grise éclairée avec des sources de lumière colorée, (2) la même scène colorée (de la couleur des sources de la scène précédente) avec des sources de lumière blanche. Sans adaptation chromatique, les images des deux scènes sont identiques. Mais un observateur perçoit les couleurs intrinsèques des scènes : la scène grise est perçue comme grise alors que la scène colorée est perçue comme colorée. Comme les approches physiques sont uniquement basées sur les statistiques des images, elles ne font pas de distinction entre les deux images des deux scènes du cas de test « monde monochromatique ». Elles donnent donc la même couleur d'estimation et échouent.

Les modèles d'apparence des couleurs (*Color Appearance Model*) [133, 134, 135] prennent en compte l'adaptation chromatique, elles possèdent donc une étape d'estimation de la couleur d'adaptation. Le but des modèles d'apparence des couleurs

est de calculer l'apparence d'un stimulus coloré (petite surface colorée) sous des conditions d'illuminations données. Dans ce cas la couleur d'adaptation est donnée directement par les conditions de test. Ces modèles ont été étendus pour calculer l'apparence des images. Dans ce cas une adaptation chromatique est appliquée à chaque pixel de l'image. La couleur d'adaptation du pixel est calculée comme la couleur du pixel après application d'un flou gaussien sur l'image [134, 135] ou comme l'interpolation entre une couleur blanche et la couleur de la moyenne géométrique du voisinage du pixel [136]. Cette estimation de la couleur d'adaptation ne correspond pas à nos besoins. Nous cherchons la couleur d'adaptation permettant de récupérer les couleurs telles qu'elles seraient perçues et non l'apparence des pixels de l'image.

Le fait que les algorithmes d'illumination globale nécessitent une étape d'adaptation chromatique afin de fournir les couleurs perçues est largement connu [137], mais peu de travaux ont porté sur l'adaptation chromatique dans le contexte de l'illumination globale [137, 138, 139]. Pourtant la connaissance de la scène (géométrie, réflectances des objets, couleurs des sources) devrait permettre une estimation précise de la couleur d'adaptation.

Wilkie et Weidlich [139] proposent une méthode qui dépasse les limitations des méthodes précédentes [137, 138]. Leur méthode est capable de prendre en compte des conditions complexes d'éclairement comme de l'éclairage indirect, de l'éclairage passant au travers de surfaces transparentes colorées, etc. La méthode procède comme suit. Premièrement, elle utilise un moteur d'illumination globale pour calculer et sauvegarder l'illumination incidente sur toutes les surfaces. Deuxièmement, les BRDFs des surfaces sont remplacées par des réflectances achromatiques, puis ils calculent une image intermédiaire en multipliant les illuminations incidentes avec la scène achromatique. Troisièmement, un poids est associé à chaque pixel de l'image intermédiaire en fonction de sa réflectance d'origine (plus la réflectance est achromatique plus le poids est fort). Enfin, la couleur d'adaptation est la moyenne pondérée des couleurs des pixels de l'image intermédiaire. Cette méthode donne de bon résultat mais elle possède néanmoins deux limitations. Premièrement, la couleur d'adaptation est calculée à partir des pixels de l'image. Alors un petit déplacement de la caméra peut entrainer une forte variation de la couleur d'adaptation en particulier si le petit mouvement de la caméra a pour conséquence de faire entrer ou sortir de l'image une surface achromatique. Nous appelons ce défaut incohérence spatio-temporelle. Deuxièmement, la stratégie de pondération surestime la couleur de l'illumination arrivant sur les surfaces blanches au détriment de la couleur globale de l'éclairement. Par exemple un source de lumière n'éclairant le seul objet blanc de la scène voit sa contribution surestimée par rapport à l'éclairage global de la scène.

En résumé, les méthodes existantes sont spatio-temporellement incohérentes du fait qu'elles limitent l'estimation de la couleur d'adaptation aux surfaces visibles depuis la caméra. Les méthodes existantes donnent également des résultats qui ne

paraissent pas naturels. Cela provient des deux éléments suivants : premièrement, la stratégie de pondération de la moyenne donne une pondération trop forte aux surfaces blanches au détriment de l'éclairage global, deuxièmement, la stratégie ne prend pas compte de la distance des surfaces par rapport à la caméra, les surfaces très éloignées ont le même poids que les surfaces proches qui reflètent pourtant la couleur de l'éclairage local.

6.3.3 Notre méthode d'estimation de la couleur d'adaptation centrée sur l'oeil

Modèle du capteur hémisphérique virtuel

L'hypothèse fondamentale de l'adaptation chromatique est qu'un objet blanc (réflectance achromatique) est perçu blanc quelque soient les conditions d'illuminations. La couleur physique d'un point d'une surface blanche diffuse est la couleur de l'irradiance à ce point. Nous pouvons généraliser pour n'importe quelles surfacse diffuses, alors la couleur perçue $L_a(y)$ d'un point y est la couleur de la réflectance $\rho(y)$ à ce point. La couleur de la luminance $L(y) = [L_L, L_M, L_S]$, en y d'une surface dont la réflectance est $\rho(y) = [\rho_L, \rho_M, \rho_S]$ et dont l'irradiance est $E(y)$, est donnée par 6.15 :

$$L(y) = \frac{1}{\pi} [E(y)] \rho(y) \tag{6.15}$$

avec :

$$[E(y)] = \begin{bmatrix} E_L(y) & 0 & 0 \\ 0 & E_M(y) & 0 \\ 0 & 0 & E_S(y) \end{bmatrix} \tag{6.16}$$

L'équation 3.10 peut alors être réécrite :

$$L_a(y) = \rho(y) = \frac{1}{\pi} \begin{bmatrix} 1/L_W & 0 & 0 \\ 0 & 1/M_W & 0 \\ 0 & 0 & 1/S_W \end{bmatrix} [E(y)] \rho(y) \tag{6.17}$$

Alors, la couleur d'adaptation $x_W(y)$ est la couleur de l'irradiance :

$$x_W(y) = \frac{1}{\pi} \begin{bmatrix} E_L(y) \\ E_M(y) \\ E_S(y) \end{bmatrix} \tag{6.18}$$

Étant donné qu'une couleur d'adaptation unique est nécessaire pour toute l'image, celle-ci est habituellement calculée comme la moyenne pondérée des irradiances [138] ou des luminances [139] des surfaces visibles depuis la caméra. A notre avis, les limites des méthodes existantes proviennent, d'une part, de la stratégie de pondération de la moyenne et, d'autre part, des éléments qui rentrent dans le calcul de la moyenne (pixels de l'image).

Les stratégies de pondération de la moyenne donnent habituellement plus de poids aux surfaces dont la réflectance est achromatique. Cela a pour but d'assurer que les surfaces blanches soient perçues blanches après adaptation chromatique. Néanmoins cela surestime l'éclairage sur ces surfaces blanches au détriment de l'éclairage global. L'équation 6.18 montre qu'il n'y pas de justification à donner plus de poids à l'éclairage arrivant sur les surfaces blanches. La stratégie de pondération des méthodes existantes ne prend pas en compte la distance des surfaces depuis la caméra. Cela signifie que l'éclairage autour de la position de la caméra (de l'observateur) n'est pas plus de poids que l'éclairage lointain. La distance depuis la caméra est prise en compte uniquement par le biais de la perspective : une surface éloignée de la caméra a une projection recouvrant moins de pixels qu'une surface proche ayant la même aire. La méthode d'estimation de la couleur d'adaptation doit, à notre avis, plus privilégier la couleur de l'éclairage proche de l'observateur. Le calcul de la moyenne se limite généralement aux pixels de l'image, c'est à dire aux surfaces visibles depuis la caméra. Cela a deux conséquences : une incohérence spatio-temporelle et une dépendance vis à vis de l'angle de vue. Ainsi un petit mouvement de la caméra peut entraîner une forte variation de l'estimation de la couleur d'adaptation. C'est le cas notamment quand du fait du déplacement de la caméra une surface blanche rentre ou sort du champ de la caméra. Limiter l'estimation de la couleur d'adaptation à l'image provoque une dépendance de l'estimation à l'angle de vue. Ainsi réduire ou élargir le champ de vision peut changer la couleur d'adaptation. Cette dépendance ne nous paraît pas naturelle.

En résumé, nous pensons qu'une méthode d'estimation de la couleur d'adaptation doit avoir les propriétés suivantes :

1. indépendante des réflectances des surfaces visibles,
2. cohérente spatio-temporellement,
3. indépendante de l'angle de vue utilisée par calculée l'image.

Nous considérons que la couleur d'adaptation est la couleur de la lumière, et pour être plus précis de l'éclairement au niveau de l'oeil. Cette hypothèse est cohérente avec le modèle de J. Von Kries [39] qui fait l'hypothèse que ce sont les photorécepteurs qui effectuent l'adaptation chromatique. Elle s'accorde également à notre expérience de tous les jours qui nous montre que lorsque nous sommes dans une pièce et que nous regardons une autre pièce, munie d'un éclairage de couleur différente, nous pouvons percevoir la couleur de l'éclairage de la pièce observée. Cela montre que l'adaptation chromatique se fait plus en fonction de l'endroit où nous sommes plutôt que de là où nous regardons. Nous rendons ainsi l'adaptation chromatique indépendante de la réflectance des surfaces observées et seulement dépendante de l'éclairement arrivant à l'œil.

Nous faisons également l'hypothèse que toutes les directions (champ de vision de 180 $^\circ$) contribuent uniformément à l'adaptation. Cela signifie que l'adaptation chromatique ne se limite pas à l'angle de vue de la caméra, elle prend en compte

l'éclairement venant en dehors de l'angle de vue de la caméra. Nous justifions cela par le fait que la vision humaine a approximativement un champ de vision hémisphérique. Nous rendons ainsi l'estimation de la couleur d'adaptation indépendante de l'angle de vue de la caméra.

En résumé l'estimation de la couleur d'adaptation peut être modélisée comme un capteur hémisphérique situé au niveau des yeux. Ce capteur est isotrope et mesure la couleur de l'éclairement à sa surface. D'un point de vue mathématique, la couleur d'adaptation x_W^{eye} est la moyenne des couleurs de l'irradiance sur la surface d'une petite hémisphère virtuelle située au niveau des yeux et orientée dans la direction de la caméra (Fig. 6.13) :

$$x_W^{eye} = \frac{1}{2\pi r^2} \int_{S_\Omega} E(x)dx \tag{6.19}$$

avec $E(x)$ l'irradiance en x sur la surface de l'hémisphère S_Ω de rayon r.

L'irradiance $E(x)$ est calculée comme :

$$E(x) = \int_\Omega L_i(x, \omega_i) cos(n_x, \omega_i) d\omega_i \tag{6.20}$$

avec $L_i(x, \omega_i)$ la luminance incidente dans la direction ω_i au point x, et n_x la normale en x.

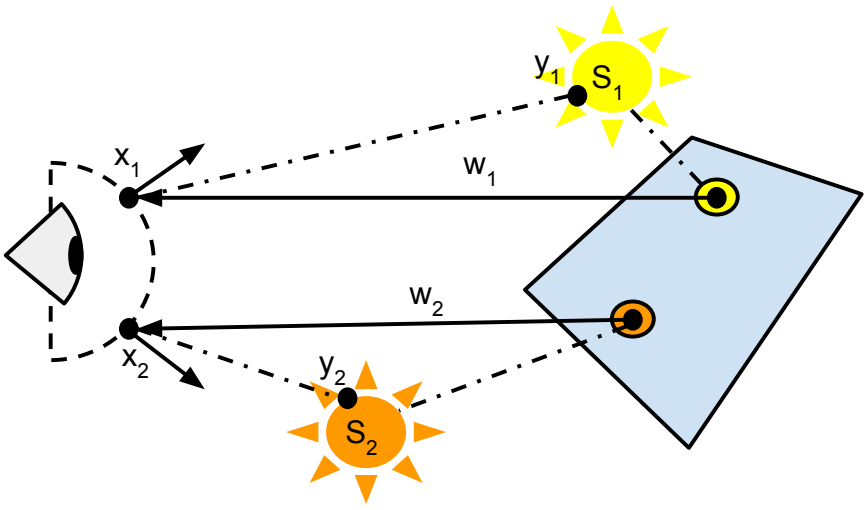


FIGURE 6.13 – Moyenne de la couleur de l'irradiance sur le capteur hémisphérique comme couleur d'adaptation.

Notre nouvelle approche pour l'estimation de la couleur d'adaptation a les caractéristiques suivantes :

1. notre couleur d'adaptation est calculée comme la moyenne des irradiances arrivant sur le capteur virtuel situé au niveau des yeux, de ce fait une plus grande importante est donnée à la couleur de l'éclairement au niveau des yeux,
2. notre couleur d'adaptation ne met aucune pondération particulière sur la luminance des surfaces blanches visibles, elle ne surestime donc pas l'éclairage local à ces surfaces blanches,
3. notre estimation de la couleur d'adaptation ne dépend pas de champ de vision de la caméra utilisée, elle est donc plus robuste vis à vis des déplacements de la caméra.

Le moteur de rendu calcule l'image brute et la moyenne des couleurs de l'irradiance sur l'hémisphère (Cf. Figure 6.14). Ensuite l'adaptation chromatique (*CAT : Color Adpatation Transform*) est appliquée sur l'image brute afin de d'obtenir une image adaptée dont les couleurs correspondent à celles qui serait perçues. Cette image peut ensuite être modifiée pour avoir une température de couleur déterminée par le créateur.

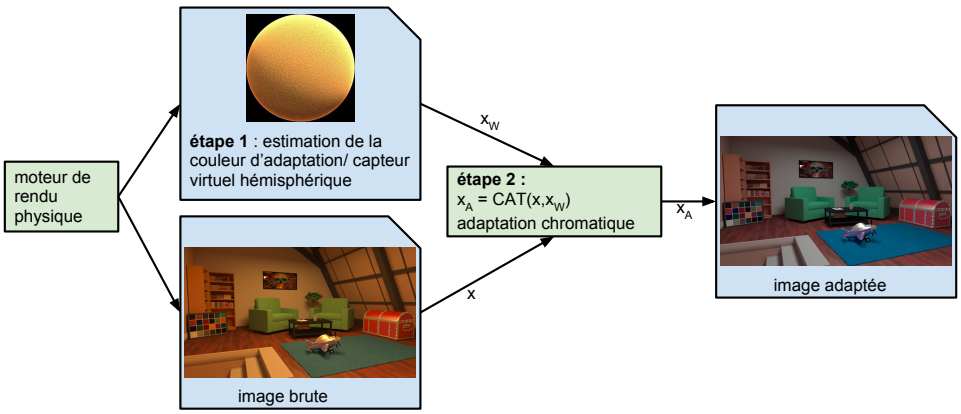


FIGURE 6.14 – Architecture overview of our chromatic adaptation process

Pour l'adaptation chromatique, nous utilisons la transformation linéaire de la Commission Internationale de l'Éclairage appelée CAT02 [134].

Mise en oeuvre

La mise en œuvre de notre modèle utilise le moteur de rendu physique Mitsuba [87] dans son mode « suivi de chemin » (*path tracing*). Afin de calculer la moyenne de l'irradiance sur l'hémisphère, nous séparons la composante directe et la composante indirecte de l'irradiance. Pour le calcul de la contribution directe, nous utilisons une méthode d'échantillonnage classique en prenant des points de manière aléatoire sur les sources de lumière. Pour la contribution indirecte, nous utilisons la technique

classique du suivi de chemin stochastique (*stochastic ray tracing*). Le détail de cette mise en œuvre est présenté dans [25].

Concernant la performance, l'évaluation des termes direct et indirect utilise un nombre d'échantillon N. Dans notre utilisation, nous avons utilisé $N = 0,1\%$ du nombre total d'échantillons utilisés pour calculer l'image. Ce nombre est suffisant pour une bonne estimation de la couleur d'adaptation et représente un surcoût négligeable.

Nous avons proposé également une mise en œuvre de notre méthode d'estimation de la couleur d'adaptation compatible avec des moteurs de synthèse d'images physique commerciaux pour lesquels il n'est pas possible d'intervenir dans le code. Dans ce cas nous procédons comme suit :

1. une petite hémisphère diffuse blanche est placé à la position de la caméra juste devant le plan de l'image (*front clipping plane*). L'hémisphère n'est pas visible dans l'image de la scène,
2. l'image brute de la scène est rendue. Le petit hémisphère ne perturbe pas beaucoup l'éclairage calculé,
3. une camera orthographique est ajoutée devant l'hémisphère. Elle regarde l'hémisphère,
4. l'image de l'hémisphère par la caméra orthographique est calculée avec le moteur,
5. à partir de l'image de l'hémisphère, la moyenne des luminances L_{av} sur l'hémisphère est calculée. Comme l'hémisphère est blanc et parfaitement diffuse, nous obtenons la moyenne de l'irradiance sur l'hémisphère (πL_{av}). La couleur moyenne donne la couleur d'adaptation.
6. l'adaptation chromtique est effectuée sur l'image brute.

6.3.4 Résultats

Nous avons montré [25] que dans les cas de tests standards (tel qu'ils sont proposés par [139] : éclairage direct, monde monochromatique [132], éclairage indirecte, etc) notre méthode calcule la bonne couleur d'adaptation comme le méthode *WCAM* (*W*ilkie *C*olor *A*daptation *M*éthode) [139]. En effet dans ces cas de test, la couleur de la lumière est homogène dans la scène alors prendre en compte uniquement le champ de vision ou un champ de 180 ° ne change pas le résultat .

Nous avons ajouté une autre série de cas tests plus complexes. Par complexe nous entendons : des grandes variations de couleurs de l'irradiances dans la scène 3D dues des sources de lumière multiples avec des couleurs différentes. Dans ces cas complexes, le résultat de l'adaptation chromatique n'est pas fourni par l'état de l'art. Par conséquence, nous pouvons seulement apprécier l'aspect plus ou moins naturel de l'image adaptée et mettre en évidences les artéfacts dus à l'estimation de la couleur d'adaptation. Nous montrons également le résultat sur une séquence

d'images correspondant à une ballade dans des pièces ayant des éclairages de couleurs différentes. Dans tous les cas testés [25] nous avons des résultats plus naturels, plus cohérents spatio-temporellement par rapport à ceux obtenue avec *WCAM*.

Nous ne présentons dans ce document que le cas de test complexe qui correspond en une scène 3D faite de deux pièces (Cf. Figure 6.15). La première est éclairée avec des sources de couleurs oranges alors que la seconde pièce est muni d'un éclairage bleutée. Pour cette scène nous calculons trois images (Cf. Figure 6.17) :

1. la première image correspond au champ de vision rouge dans la figure 6.17, dans cette image la seconde pièce n'est pas visible, une première statue blanche est dans le champ de vue,
2. la deuxième image, champ de vison vert dans la figure 6.17, permet de voir les deux pièces et les statues blanches de chaque pièce,
3. la troisième image correspond au champ de vision bleu dans la figure 6.17, dans cette image seule le seconde pièce, éclairée en bleu, est visible alors que la caméra se situe toujours dans la pièce éclairée avec des sources oranges.

Dans ce cas de test, *WCAM* calculent trois couleurs d'adaptation différentes car elle limite le calcul de la couleur d'estimation au champ de vision de la caméra. En revanche notre méthode calcule trois couleurs d'adaptation très proche les unes des autres comme le montre le diagramme de chromaticité, figure 6.16. Cela résulte du fait que notre méthode est plus sensible à la position de la caméra qu'à l'irradiance des objets visibles depuis la caméra.

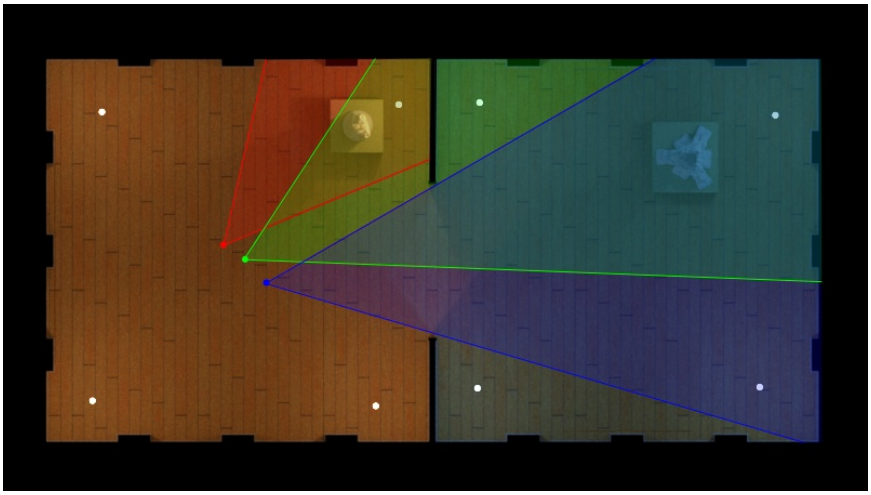


FIGURE 6.15 – Vue du dessus de la scène composée de deux pièces et visualisation des trois champs de vision correspondant au trois positions et orientations de la caméra (rouge, vert, bleue)

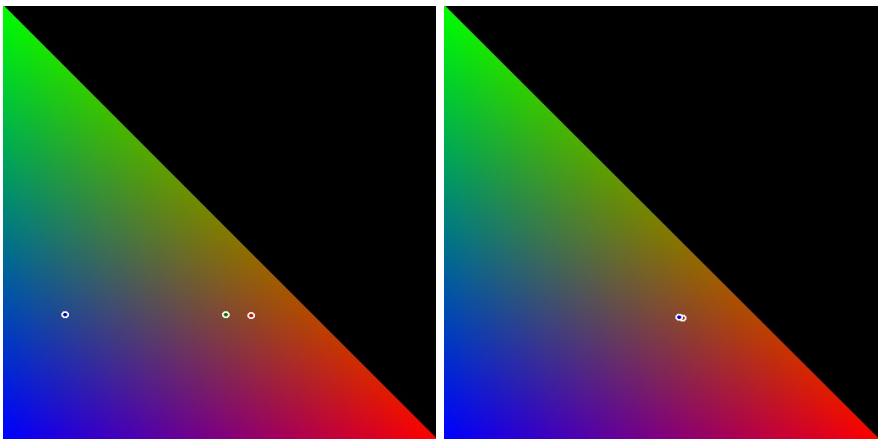


FIGURE 6.16 – Diagramme de chromaticité dans l'espace de couleur RGB. Les cercles de couleur (rouge, vert, bleu) correspondent à la couleur d'adaptation pour les trois positions et orientations de la caméra (Cf. Figure 6.15). La figure de gauche correspond à la méthode WCAM. La figure de droite montre les résulats pour notre méthode.

Les images brutes, adaptée avec *WCAM* et adaptée avec notre méthode (Cf. Figure 6.17) montrent bien la différence de rendu final obtenu entre les images adaptées par *WCAM* alors que notre méthode fournit des images spatio-temporellement cohérente : un petit déplacement de la caméra n'influence que peu l'estimation de la couleur d'adaptation. Les trois images adaptées avec notre méthode, du fait de leur cohérence, nous paraissent plus naturelles.

6.3.5 Conclusion sur l'estimation de la couleur d'adaptation

Nous avons proposé une méthode simple, précise et cohérente spatio-temporellement qui calcule automatiquement la couleur d'adaptation dans le contexte de la synthèse d'images physique. Notre modèle assimile l'estimation de la couleur d'adaptation à un capteur hémisphérique virtuel placé au niveau de la caméra qui mesure la couleur moyenne de l'irradiance à sa surface. Premièrement, nous avons montré que cette méthode surpasse les méthodes proposées dans l'état de l'art et notamment dans le cas de conditions d'illumination complexe. Deuxièmement, du fait de sa cohérence spatio-temporelle, notre méthode peut être utilisée dans le cas de séquences d'image. Troisièmement, elle peut se mettre en œuvre également avec des moteurs d'illumination globale propriétaires. Cette contribution a été publiée dans une revue internationale de premier plan [25].

Notre méthode d'estimation de la couleur d'adaptation permet, soit de restituer les couleurs telles qu'elles seraient perçues par un observateur ce qui donne des couleurs paraissant plus naturelles, soit de choisir la température de couleur des images



FIGURE 6.17 – Cohérence spatiale et temporelle de l'estimation de la couleur d'adaptation dans le cas de deux pièces ayant des éclairages de couleurs différentes. Image brute à gauche, image adaptée selon WCAM au centre, image adaptée avec notre méthode à droite.

ce qui donne alors l'ambiance de couleur choisie par le créateur (chaude ou froide par exemple) aux images. Le choix s'effectue lors de l'adaptation chromatique en choisissant la couleur du blanc de référence : $[R_r, G_r, B_r]$.

Chapitre 7

Discussion et Perspectives

La première partie 7.1 de ce chapitre dresse le bilan des contributions présentées dans ce document. Elle les replace du point de vue du créateur et des possibilités de contrôle qu'apportent nos contributions. La deuxième partie 7.2 présente les perspectives de recherche qui se focalise sur la prise en compte de l'intention de l'utilisateur. Nous proposons en effet d'aller plus loin en terme de capacité de contrôle par le créateur sur l'image calculée. Nous proposons de permettre au créateur de décrire l'esthétique de l'image qu'il souhaite calculer. Cette intention esthétique est prise en compte par le pipeline de rendu afin d'obtenir plus rapidement l'image souhaitée, c'est à dire celle qui correspond à l'intention du créateur. Enfin la partie 7.3 conclue sur nos contributions et notre projet de recherche. Nous en précisons les enjeux et les retombés attendus.

7.1 Discussion

Notre positionnement vis à vis des nombreux enjeux de la synthèse d'image physique est de se focaliser sur les moyens de contrôles proposés au créateur afin qu'il puisse obtenir rapidement l'image qu'il souhaite. Nos contributions portent sur différentes parties du pipeline concourant à la production des images : modélisation, moteur de rendu et post-traitements (Cf. figure 7.1). Nous allons discuter dans les parties suivantes de l'apport en terme de contrôle pour le créateur de nos contributions. Pour chaque contribution présentée dans ce document, nous rappelons brièvement la contribution et nous l'analysons du point de vue contrôle supplémentaire pour le créateur.

7.1.1 Modélisation automatique de scène 3D

Nous avons proposé un environnement de création automatique de monde 3D permettant de mettre en scène les données issues de bases de données. Premièrement, les données sont extraites dynamiquement et hiérarchisées selon la stratégie définie dans le *génome*. Ensuite, le monde 3D est construit, la topologie du monde

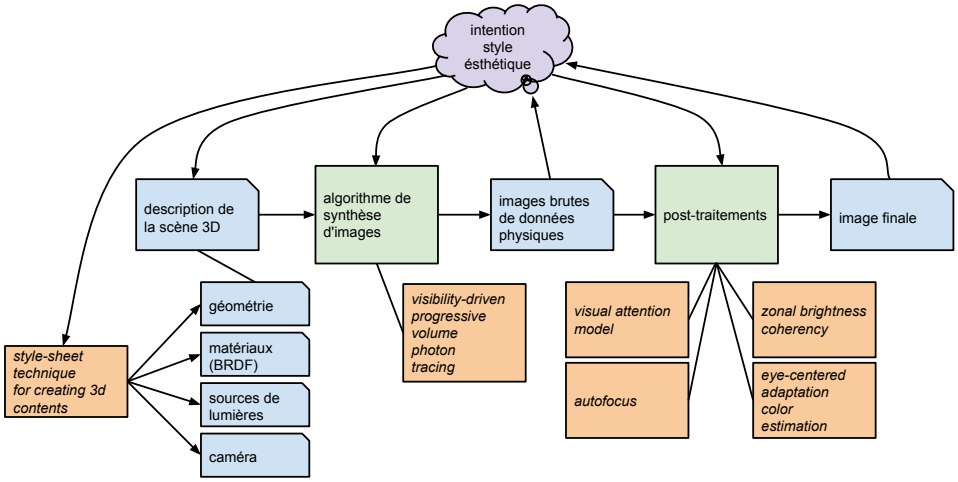


FIGURE 7.1 – Chaîne de la synthèse d'images physique avec nos principales contributions.

3D est conforme à la structure hiérarchique définie par le *génome*. Les données extraites sont habillées par des modèles 3D selon les règles écrites dans le *phénome*. Ces règles définissent également comment les différents modèles 3D sont positionnés dans le monde 3D.

Pour le créateur, notre environnement de modélisation lui permet de créer automatiquement et dynamiquement des mondes 3D à partir de bases de données. Aucun autre environnement n'offre cette fonctionnalité. Il contrôle la topologie du monde 3D via le *génome* et le visuel de monde 3D via le *phénome*. À partir des spécifications (*génome* et *phénome*), le monde 3D est construit automatiquement à partir des données présentes dans la base de données. Le créateur spécifie les règles de constructions et les modèles 3D utilisés, il n'a pas le contrôle direct sur le monde 3D construit. Mais cette limite est inhérente à l'aspect automatique et dynamique du système que nous avons proposé.

7.1.2 Rendu des milieux participatifs en illumination globale

Nous avons proposé une nouvelle approche pour le rendu des milieux participatifs en illumination globale. Cette fonctionnalité était déjà présente dans de nombreux moteurs de rendu. Notre approche a trois avantages pour le créateur. Premièrement, notre algorithme prend en charge tout type de milieu participatif (homogène, hétérogène) et de surfaces (diffus, brillant, spéculaire, transparent, etc.). Notre méthode n'impose pas aux créateurs des limites sur les caractéristiques des matériaux utilisés dans son monde 3D. Deuxièmement, notre algorithme a une consommation mémoire qui ne dépend pas de nombres de particules (photons ou faisceaux) utilisés. Le créa-

teur peut donc calculer des images de très haute qualité (beaucoup de particules) sur des configurations matérielles abordables. Troisièmement, notre algorithme de rendu est progressif. Une image montrant l'avancée du rendu est disponible à l'issue de chaque passe. Le créateur peut donc vérifier la progression le calcul de l'image en cours de rendu. Il peut donc arrêter le calcul s'il constate que l'image ne correspond pas à ce qu'il souhaite avant la fin du calcul complet de l'image.

7.1.3 Attention visuelle dans les environnements 3D interactifs

Nous avons proposé un modèle d'attention visuelle adapté aux environnements 3D interactifs qui combine des composantes attentionnelles *bottom-up* et *top-down*. Notre modèle est plus précis et performant que les modèles d'attention visuelle proposés antérieurement.

Pour le créateur, le modèle d'attention visuelle est un outil qui lui permet de déterminer automatiquement où le regard de l'observateur va être focalisé. Cela lui permet de vérifier que les éléments important de son monde 3d sont bien mis en valeur. Le modèle d'attention visuelle peut être également utilisé comme une entrée pour des traitements postérieurs comme le réglage de l'exposition en fonction de la zone observée.

7.1.4 Autofocus intelligent pour le flou de profondeur

De manière duale au modèle d'attention visuelle, nous avons proposé un système d'autofocus intelligent pour déterminer sur quelle partie de l'image faire le net. Notre système a été créé pour les mondes 3D interactif. Pour le créateur, ce système lui offre un outil de contrôle sur l'attention visuelle de l'observateur. Ainsi pour une application 3D interactive il peut définir simplement (poids sémantique associé aux objets de la scène 3D) quelle seront les objets qui attirerons l'attention visuelle de l'utilisateur de l'application.

7.1.5 Adaptation des gammes de luminance pour les séquences d'images

Nous avons proposé un algorithme qui permet de conserver la cohérence spatiale et temporelle des luminances entre les images *HDR* et *LDR* d'une séquence d'images. L'algorithme est compatible avec tous les opérateurs de réduction des gammes de luminance. De plus notre algorithme conserve les propriétés de l'opérateur de réduction des gammes de luminance utilisé.

Notre solution laisse la liberté au créateur de choisir l'opérateur de réduction des gammes de luminance qui lui convient afin de transformer les images *HDR* en images *LDR*. Le créateur est assuré que les réglages effectués sont valides sur l'ensemble de la séquence. Le créateur gère ainsi la séquence d'image aussi simplement

qu'une simple image. La séquence produite est cohérente, n'a pas de problème de scintillement et respecte les lumières des images brutes *HDR*.

7.1.6 Détermination de la couleur d'adaptation

Nous avons proposé un nouvel algorithme de calcul de la couleur d'adaptation dans le cadre de l'illumination globale. Notre algorithme donne de meilleurs résultats que les solutions précédentes, il est notamment plus stable temporellement que les autres algorithmes. Notre algorithme est, de plus, indépendant du moteur de rendu.

Pour le créateur, avoir une estimation précise de la couleur d'adaptation permet soit de fournir une image dont les couleurs correspondent à celles qui seraient perçues, soit de choisir la dominante de couleur qu'il souhaite. La cohérence temporelle de notre algorithme permet d'obtenir le même équilibre colorimétrique sur toute la séquence d'images.

7.1.7 Analyse

Nos différentes contributions (rappelées brièvement dans les sections précédentes) apportent plus de contrôle pour le créateur : nouvelle possibilité de création (modélisation automatique de monde 3D à partir de base de données), capacité de contrôler avant la fin de calcul l'image produite (méthode progressive pour le rendu des milieux participatifs), outil de contrôle de l'attention visuelle et garantie de la cohérence spatiale et temporelle dans les post-traitements (adaptation des gammes de luminances et de couleurs). Néanmoins dans ces algorithmes l'intention de l'utilisateur n'est pas explicitement présente. L'intention de l'utilisateur peut se décrire en terme d'esthétique de l'image ou de la séquence d'images produites. Par exemple, le créateur peut définir qu'il souhaite créer une image avec l'esthétique *high key*. Aujourd'hui, ni les outils de modélisation, ni les moteurs de rendu, ni les étapes de post-traitement ne prennent en compte des informations sur le style ou l'esthétique de l'image souhaitée par le créateur. Dès lors, le créateur ne reçoit pas d'aide de l'environnement. Il est seul à régler, vérifier, tester les paramètres et espérer ainsi obtenir l'image souhaitée.

La prise en compte de l'intention en terme d'esthétique dans le pipeline de création d'images a pour but de d'obtenir plus rapidement l'image souhaitée par le créateur en limitant drastiquement le nombre d'essais/erreurs/réglages fait par ce dernier avant d'obtenir le bon ensemble de paramètres correspondant à l'image avec l'esthétique désirée. Avant d'être en mesure de prendre en compte une intention esthétique dans le pipeline de rendu, il reste des verrous scientifiques à lever. Il faut notamment une description formelle de l'esthétisme et les algorithmes permettant d'évaluer la distance entre une image et une esthétique. Disposant de ces deux éléments, il est alors possible de modifier l'architecture du pipeline de rendu pour prendre en compte cette intention esthétique.

Premièrement la définition d'un modèle proposant un cadre formel pour la description de l'esthétique des images calculées, deuxièmement, la mise au point d'une métrique permettant de calculer la distance entre une image et une esthétique et, troisièmement, la prise en compte de l'intention du créateur sous la forme de l'esthétique souhaitée dans le pipeline de rendu forment les trois éléments clés des perspectives de recherche détaillés dans la partie 7.2.

7.2 Perspectives

7.2.1 Introduction

Nos perspectives de recherche se situent dans le contexte de la synthèse d'images physique (illumination globale). Elles ont pour but de proposer aux créateurs d'images et de séquences d'images un environnement de création dans lequel leurs intentions esthétiques sont prise en compte afin d'obtenir rapidement les images avec l'esthétique souhaitée. Dans ce projet de recherche l'élément central est la création d'un modèle permettant de décrire et de mesurer l'esthétique d'une image. Il ne s'agit pas ici d'évaluer la qualité esthétique d'une image. Mais il s'agit d'avoir un modèle permettant de décrire d es esthétiques et de calculer une distance entre une image et une esthétique. L'esthétique est ici synonyme de style. Avec ce modèle d'esthétique, le créateur peut décrire l'esthétique qu'il souhaite pour son image, cette description doit être considérée comme une entrée supplémentaire du pipeline de calcul des images. Le système doit alors tenir compte de cette entrée afin d'aider le créateur.

La partie 7.2.2 propose une vue d'ensemble de ce que pourrait être un pipeline de création prenant en compte l'intention esthétique du créateur. La partie 7.2.3 présente les premiers éléments sur lesquels nous avons travaillé afin de créer notre modèle d'esthétique. Dans la partie 7.2.4, nous proposons les approches que nous voulons explorer afin de régler automatiquement, c'est à dire d'optimiser, certains paramètres du pipeline de rendu (scène 3D, moteur, post-traitements) permettant ainsi d'atteindre l'esthétique souhaitée. La prise en compte de l'intention de l'utilisateur dans les phases de rendu et de post-traitement ne change pas les caractéristiques fondamentales de ces éléments (rendu et post-traitement). Néanmoins, il est important de pouvoir calculer et post-traiter rapidement une image. La partie 7.2.5 indique les axes de recherche sur lesquels nous travaillons dans le cadre du rendu et des post-traitements.

7.2.2 Vue d'ensemble

La figure 7.2 illustre ce que pourrait être le pipeline de rendu prenant en compte l'intention de l'utilisateur décrite sous la forme de l'esthétique souhaitée pour l'image calculée. Les éléments nouveaux sont les boîtes rouges numérotées de un à cinq.

Le premier élément supplémentaire (boîte rouge (1) dans la figure 7.2) est la description de l'intention sous la forme d'un style ou d'une esthétique. Associée à cette

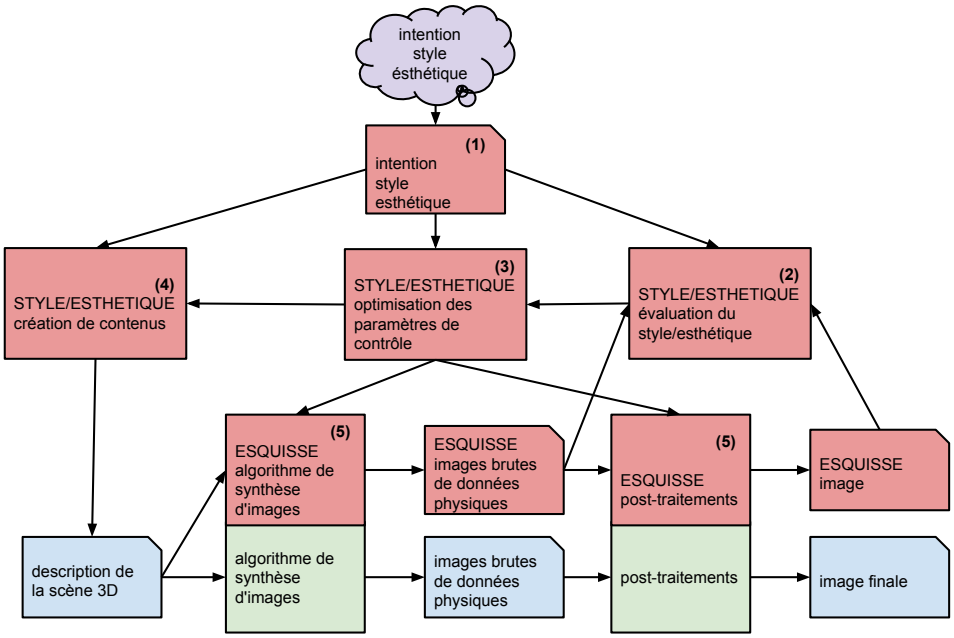


FIGURE 7.2 – Chaîne de la synthèse d'images physique avec prise en compte de l'intentien styliste/esthétique du créateur

description de l'esthétique souhaitée, l'élément d'évaluation de l'esthétique (boîte rouge (2) dans la figure 7.2) permet de calculer la distance entre l'image calculée et l'esthétique souhaitée. Les éléments de description et d'évaluation sont les éléments centraux. Ils permettent, premièrement, au créateur d'exprimer son intention et, deuxièmement, au pipeline d'évaluer l'image calculée par rapport à cette intention. L'approche choisie pour ces deux éléments est présentée par la partie 7.2.3.

Le but est d'aider le créateur. Cela peut se faire en adjustant les réglages de paramètres dans le pipeline. L'ajustement ou plus précisement l'optimisation de paramètres nécessite un bloc fonctionnel d'optimisation des paramètres du pipeline. C'est le rôle de la boite (3) dans la figure 7.2. Les paramètres optimisés peuvent être de différentes natures :

- paramètres relatifs à la scène 3D et notemment les paramètres des sources de lumières comme les ratios de puissances des différentes sources, leur géométrie, etc. (boîte (4)) dans la figure 7.2),
- paramètres relatifs au moteur de rendu et/ou post-traitement, typiquement les opérateurs de réduction de gamme de luminance ont des paramètres important influençant l'aspect de l'image finale.

Le créateur doit pouvoir spécifier les paramètres qu'il accepte de voir modifier automatiquement par le pipeline. Il s'agit ici de proposer des modes de contrôles

pouvant être similaires à ceux présents dans les appareils photos et caméras : manuel, priorité vitesse, priorité ouverture, automatique, automatique portrait, etc. Chaque mode de contrôle détermine clairement les paramètres imposés et fixés par le créateur et ceux qu'il laisse sous le contrôle du pipeline.

7.2.3 Description et évaluation de l'esthétique

La notion d'esthétique et de style a donné lieu à deux nombreux travaux dans deux domaines : premièrement, dans l'évaluation de la qualité esthétique des images et, deuxièmement, dans la problématique du transfert de style ou d'esthétique. Dans la suite, la notion d'esthétique et celle de style sont considérées comme équivalente, nous appelons cette notion *style/esthétique*.

L'évaluation de la qualité esthétique [140, 141, 142, 143, 144, 145] a pour but de calculer une note indiquant la qualité esthétique de l'image (c'est à dire de la beauté image). Dans l'évaluation de la qualité esthétique, on retrouve les règles classiques de composition, d'exposition, d'harmonie des couleurs, de netteté, etc. Nous avons montré que ces méthodes ne prennent en compte qu'un *style/esthétique* de lumière : l'esthétique *medium key*. Ces méthodes échouent pour les *style/esthétique low key* et *high key*. Elles les évaluent comme des images sous-exposées ou sur-exposées [146]. Nous avons proposé une méthode permettant de classer les images selon ces trois grandes classes de *style/esthétique* de la lumière : *low key*,*medium key* et *high key* [146]. Ce premier travail sur *style/esthétique* montre qu'il est possible d'évaluer l'appartenance d'une image à différentes classes de *style/esthétique* de lumière.

La problématique du transfert ou de l'amélioration du *style/esthétique* fonctionne généralement à partir d'images servant d'exemples du *style/esthétique* à transférer [147, 148, 149, 150]. Ces méthodes se concentrent sur le transfert des *styles/esthétiques* de couleurs. Le *style/esthétique* de couleur est calculé à partir de valeurs statistiques issues des images exemples. Ces travaux montrent qu'un *style/esthétique* de couleur peut être transféré d'une image à une autre. La limite de ces approches qui utilisent des images comme définition du *style/esthétique* est l'absence de définition explicite de ce qui est effectivement transféré. L'élément de *style/esthétique* transféré est implicite. Shacked et Lischinski [151] ont proposé pour la première fois une métrique de *style/esthétique* calculée à partir de valeurs statistiques bas niveau d'une image. Leur métrique est limitée à une seule esthétique de lumière. Nous avons étendu leur métrique afin de prendre en compte les esthétiques *low key*, *medium key* et *high key* [152]. Nous avons utilisé cette métrique de *style/esthétique* pour contrôler les ratios de puissances des sources de lumière dans une scène 3D afin d'obtenir automatiquement les paramètres des sources de lumière correspondant à des images *low key*, *medium key* et *high key*. La figure 7.3 illustre ce travail [152] : à partir d'une configuration initiale correspondant à l'image de gauche, le système a optimisé les paramètres des sources de lumières (ratio de puissance) afin d'obtenir une image (image de droite) avec un *style/esthétique low*

key. On remarquera la forte augmentation du contraste sur la statue entre les deux images.



FIGURE 7.3 – Optimisation automatique des paramètres des sources de lumière pour obtenir une esthétique *low key* : (à gauche) image initiale - (à droite) image après optimisation des paramètres des sources de lumières [152]

En résumé, nous avons déjà travaillé sur l'évaluation du *style/esthétique* en montrant qu'il est possible de déterminer la classe en terme de *style/esthétique* de lumière d'une image et qu'il est possible de régler automatiquement les ratios de puissances des sources de lumière afin d'atteindre un *style/esthétique* de lumière particulier. Nous souhaitons prolonger ce travail en prenant en compte plus d'éléments dans la caractérisation du *style/esthétique* notamment les aspects relatifs à la couleur et aux corrélations entre couleur et luminance. En effet, dans un modèle d'éclairage à trois types de sources (*key light*, *fill light* et *back light*), chaque source contrôle un intervalle de luminance et une couleur.

Avec la description et l'évaluation des *style/esthétique*, l'autre élément primordial du nouveau pipeline qui prend en compte l'intention de l'utilisateur est le processus d'optimisation des paramètres. Nous précisons dans la partie suivante 7.2.4 nos axes de recherche concernant ce processus d'optimisation.

7.2.4 Optimisation des paramètres

Les travaux sur la prise en compte du *style/esthétique* (évaluation, transfert) proviennent essentiellement du domaine du traitement d'images. Ces travaux ne prennent pas en compte le calcul de l'image par un moteur de rendu. A l'opposé, les quelques travaux provenant de l'informatique graphique agissent sur les paramètres de rendu mais excluent la partie post-traitement de l'image. Dans ce

contexte, l'originalité de notre approche est de considérer les deux éléments : rendu et post-traitement. Il nous paraît indispensable de prendre en compte les deux éléments et d'avoir une stratégie qui détermine quels paramètres de quels éléments (rendu et post-traitement) doivent être optimisés afin d'obtenir l'esthétique cible. Par exemple, si l'esthétique cible est une esthétique *high key* avec des ombres froides (dominante bleue dans les ombres), alors il est indispensable que la configuration des sources de lumière correspond à une image *high key*, que la lumière *fill light* soit bleue. Mais le contrôle du niveau moyen de luminance et l'ajustement du contraste est facilement effectué par des post-traitements. Cet exemple montre bien qu'il est capital d'avoir une stratégie d'optimisation déterminant les paramètres de l'esthétique à optimiser par le rendu et ceux qui doivent être pris en compte par les post-traitements.

La recherche sur la stratégie d'optimisation sera la suivante. Pour une esthétique cible donnée, quelles sont les caractéristiques (de l'esthétique) qui sont facilement contrôlées par les post-traitements. Dès lors il est possible de déterminer les caractéristiques clés de l'image avant les post-traitements. Ces caractéristiques clés doivent donc être atteintes par le rendu. Notre idée fondatrice est qu'aucun post-traitement ne peut modifier les caractéristiques de l'image provenant directement de la configuration géométrique de l'éclairage. Les post-traitements peuvent uniquement moduler les caractéristiques issues de la configuration de l'éclairage. En résumé, un post-traitement ne peut pas transformer une image *high key* en image *low key* car les configurations géométriques des sources entre ces deux esthétiques de lumières sont différentes.

7.2.5 Rendu et des post-traitements

La prise en compte de l'intention de l'utilisateur ne nécessite pas des transformations en profondeurs des moteurs de rendu et des algorithmes de post-traitements. Néanmoins, la qualité et la rapidité du rendu et des post-traitements restent des points clés afin que la synthèse d'images physique (illumination globale) soit un choix pertinent pour le créateur.

La prise en compte de l'intention de l'utilisateur en terme de *style/esthétique* nécessite le calcul d'une image afin de pouvoir évaluer l'*style/esthétique* de l'image calculée par rapport à l'*style/esthétique* souhaité. Il est donc important de pouvoir évaluer rapidement le *style/esthétique* de l'image avant la fin du calcul complet de l'image (rendu et post-traitement). En prenant en compte, les aspects statistiques des termes intervenant dans le calcul du *style/esthétique*, nous travaillons sur des méthodes dites d'*esquisses* permettant d'évaluer rapidement avec précision le *style/esthétique* sans nécessiter le calcul complet de l'image (rendu et post-traitement). Cette approche permet de régler automatiquement et rapidement les paramètres. Alors, le réglage automatique des paramètres n'induit qu'un surcoût négligeable pour l'utilisateur.

De notre expérience sur les algorithmes de rendu, nous avons constaté qu'une scène 3D se caractérisent par deux grandes classes d'espaces : les espaces *simples* et les espaces *difficiles*. Un espace *simple* est une partie de la scène dans laquelle la lumière arrive simplement : la géométrie des chemins de lumière jusqu'à cet espace sont simple et, dans des approches probabilistes, ces chemins de lumière sont fréquents. Les espaces dits *difficiles* sont quant à eux des espaces dans lesquels la lumière arrive difficilement (chemin de lumière complexe) et la probabilité de trouver un tel chemin est rare. Les algorithmes actuels gèrent de la même manière les deux types d'espace. En conséquence, la partie de l'image correspondant aux espaces *simples* atteint rapidement une bonne qualité alors que les parties de l'image correspondant aux espaces *difficiles* convergent très lentement. Malheureusement, à chaque nouvelle itération, la probabilité de trouver un chemin arrivant dans un espace *simple* reste toujours plus importante que celle de trouver un chemin aboutissant dans un espace *difficile*. Une nouvelle passe de rendu dans un algorithme progressif, ou un plus grand nombre de photons ou d'échantillons contribuent donc plus à améliorer les parties de l'image qui ont déjà atteint une bonne qualité (espaces *simples*) que les parties correspondant aux espaces *difficiles*. La stratégie de rendu n'est pas efficace. Nous travaillons sur la segmentation automatique de la scène en espaces *simples* et *difficiles* et nous mettons au point une stratégie de rendu qui prend en compte de manière efficace cette segmentation.

7.3 Conclusion

Aujourd'hui quand un créateur utilise la synthèse d'image physique pour créer une image, il connait son intention en terme de *style/esthétique* et il essaie d'obtenir l'image qu'il souhaite en réglant progressivement, suite à de nombreux essais, les paramètres du pipeline de rendu. Ce processus d'essais-erreurs est coûteux en terme de temps car, d'une part, il nécessite à chaque essai un calcul d'une image et, d'autre part, car le réglage des paramètres est difficile.

Dans ce cadre il est important de proposer des outils de contrôle aux créateurs. Nos contributions présentées dans ce document vont dans cette direction. Nous avons proposé un système pour créer automatiquement des mondes 3D à partir de bases de données : le créateur contrôle directement la topologie du monde 3D et son aspect visuel. Notre nouvel algorithme de rendu des milieux participatifs en illumination globale rend accessible la présence de tel milieu dans des scènes 3D. La caractéristique progressive de notre algorithme permet au créateur de contrôler l'image en cours de calcul. Il n'a pas besoin d'attendre la fin de calcul pour évaluer la qualité de l'image. Notre modèle d'attention visuelle permet au créateur de régler des effets en fonction de l'endroit où le regard de l'utilisateur de focalise. Il lui permet également de vérifier la visibilité des objets importants. Notre système d'autofocus offre au créateur la possibilité de diriger l'attention visuelle de l'utilisateur d'un monde 3D interactif. En assurant la cohérence spatio-temporelle des algorithmes de

réductions des gammes de luminance appliqués aux séquences d'images, nous permettons au créateur de contrôler simplement cette étape de post-traitement. Les réglages effectués seront automatiquement valides sur toute la séquence et l'esthétique de lumière de la séquence *HDR* sera respectée sur la séquence *LDR*. Nous avons également proposé un algorithme spatio-temporellement cohérent et précis d'évaluation de la couleur d'adaptation en illumination globale. Notre algorithme permet au créateur de contrôler simplement l'équilibre colorimétrique des séquences produites.

Ces contributions permettent plus de contrôle pour le créateur sur l'image calculée. Nous proposons de poursuivre dans cette voie en prenant en compte directement l'intention de du créateur en terme de *style/esthétique*. Cela 'nécessite de lever plusieurs verrous : (1) disposer d'un modèle et d'une métrique permettant d'évaluer une image en fonction du *style/esthétique* souhaité, (2) avoir une stratégie d'optimisation des paramètres du rendu permettant de converger vers des réglages (rendu et post-traitement) correspondant au *style/esthétique* souhaité et (3) posséder des algorithmes de rendus et de post-traitements rapides ou permettant d'évaluer le *style/esthétique* rapidement avant la fin de calcul de l'image.

L'enjeu principal du projet présenté ici est de rendre accessible la puissance de la synthèse d'images physique (illumination globale, imagerie à grande gamme de luminance, gestion des couleurs) aux créateurs. L'approche retenue est de proposer des nouvelles gammes d'outils de contrôle permettant aux créateurs de se focaliser sur ce qu'il souhaite obtenir et non sur le réglage de paramètres complexes et techniques. L'une des originalités forte de l'approche est de prendre en compte la diversité des *style/esthétiques* possibles et de proposer au créateur de décrire le *style/esthétiques*. L'approche que nous proposons est donc ouverte et n'impose pas un *style/esthétiques* unique correspondant à des valeurs optimales implicites. L'enjeu scientifique est donc de comprendre, de modéliser la notion de *style/esthétiques*.

Les retombés attendues sont donc naturellement une plus grande utilisation de la synthèse d'images physique (illumination globale, imagerie à grande gamme de luminance, gestion des couleurs) dans la création et la production d'images. Mais, les retombées sont plus larges. Le contrôle du *style/esthétiques* dans les images et dans leur modifications permet de garantir le respect du *style/esthétiques* lors de l'adaptation des images aux périphériques de restitution, lors des phases adaptations des gammes de luminances et de couleurs par exemple. Le modèle de *style/esthétiques* permet également d'améliorer les algorithmes d'évaluation de la qualité des images, de recherches dans les bases de données d'images, etc. Le champ d'application des travaux de recherche que nous nous proposons de mener est donc très large et comprend tous les domaines dans lesquels la notion de *style/esthétiques* rendre en compte.

Bibliographie

[1] N. Esnault : *Production automatique d'interfaces graphiques 3D pour le Web.* Thèse de doctorat, Université de Rennes 1, 2011.

[2] N. Esnault, J. Royan, , R. Cozot et C. Bouville : Métaphores de présentation de l'information basées sur des méthodes procédurales. *In Association Française d'Informatique Graphique (AFIG)*, Arles France, 11 2009.

[3] N. Esnault, J. Royan, R. Cozot et C. Bouville : A flexible framework to personalize 3d web users experience. *In 15th International Conference on Web 3D Technology*, page 35–44, New York, NY, USA, 2010. ACM.

[4] N. Esnault, J. Royan, , R. Cozot et C. Bouville : A style-sheet technique for creating 3d web contents. *In Web3Dw '10 : Proceedings of the International Conference Web Virtual Reality and Three-Dimensional Worlds.* IADIS, 2010.

[5] N. Esnault, J. Royan, , R. Cozot et C. Bouville : A style-sheet technique for creating interactive 3d web contents. *IADIS International Journal on WWW/Internet*, 9(1), 2011.

[6] Charly Collin : Carte de photons progressive dans des scènes contenant des milieux participatifs. Mémoire de D.E.A., IRISA/Université de Rennes 1, 2011.

[7] Charly Collin, Mickaël Ribardière, Rémi Cozot et Kadi Bouatouch : Progressive volume photon tracing. *In SIGGRAPH Talks*, pages 5–9, Los Angeles, CA, 2012. ACM.

[8] Charly Collin, Mickaël Ribardière, Adrien Gruson, Rémi Cozot, Sumanta N. Pattanaik et Kadi Bouatouch : Visibility-driven progressive volume photon tracing. *The Visual Computer*, 29(9):849–859, 2013.

[9] Adrien Gruson, Ajit Hakke Patil, Rémi Cozot, Kadi Bouatouch et Sumanta N. Pattanaik : Light propagation maps on parallel graphics architectures. *In* Hank Childs, Torsten Kuhlen et Fabio Marton, éditeurs : *Eurographics Symposium on Parallel Graphics and Visualization (EGPGV)*, pages 81–88, Cagliari, Italy, 2012. Eurographics, Eurographics Association.

[10] Sébastien Hillaire : *Contribution to the Study of Visual Attention Models and Gaze Tracking to Improve Visual Feedback in 3D Interactive Applications.* Thèse de doctorat, INSA de Rennes, 2011.

[11] Sébastien Hillaire, Anatole Lécuyer, Rémi Cozot et Géry Casiez : Using an eye-tracking system to improve camera motions and depth-of-field blur effects in virtual environments. *In IEEE Virtual Reality Conference 2008 (VR 2008)*, pages 47–50, Reno, Nevada, USA, 2008. IEEE.

[12] Sébastien Hillaire, Anatole Lécuyer, Tony Regia-Corte, Rémi Cozot, Jérôme Royan et Gaspard Breton : A real-time visual attention model for predicting gaze point during first-person exploration of virtual environments. *In* George Baciu, Rynson W. H. Lau, Ming C. Lin, Taku Komura et Qunsheng Peng, éditeurs : *ACM Symposium on Virtual Reality Software and Technology (VRST)*, pages 191–198, Hong Kong, China, 2010. ACM.

[13] Sébastien Hillaire, Gaspard Breton, N. Ouarti, Rémi Cozot et Anatole Lécuyer : Using a visual attention model to improve gaze tracking systems in interactive 3d applications. *Computer Graphics Forum (CGF)*, 29(6):1830–1841, 2010.

[14] Sébastien Hillaire, Anatole Lécuyer, Tony Regia-Corte, Rémi Cozot, Jérôme Royan et Gaspard Breton : Design and application of real-time visual attention model for the exploration of 3d virtual environments. *IEEE Transactions on Visualization and Computer Graphics (TVCG)*, 18(3):356–368, 2012.

[15] Sébastien Hillaire, Anatole Lécuyer, Rémi Cozot et Géry Casiez : Effets de flou visuel pour la navigation en environnements virtuels en vue à la première personnes. *In Association Française d'Informatique Graphique (AFIG)*, pages 83–90, Marne-la-Vallée, France, 2007.

[16] Sébastien Hillaire, Anatole Lécuyer, Rémi Cozot et Géry Casiez : Depth-of-field blur effects for first-person navigation in virtual environments. *In* Aditi Majumder, Larry F. Hodges, Daniel Cohen-Or et Stephen N. Spencer, éditeurs : *Proceedings of the ACM Symposium on Virtual Reality Software and Technology (VRST)*, pages 203–206. ACM, 2007.

[17] Sébastien Hillaire, Anatole Lécuyer, Rémi Cozot et Géry Casiez : Depth-of-field blur effects for first-person navigation in virtual environments. *IEEE Computer Graphics and Applications*, 28(6):47–55, 2008.

[18] Erik Reinhard, Timo Kunkel, Yoann Marion, Jonathan Brouillat, Rémi Cozot et Kadi Bouatouch : Image display algorithms for high and low dynamic range display devices. *Journal of the Society for Information Display*, 15(12):997–1014, 2007.

[19] Ronan Boitard, Dominique Thoreau, Rémi Cozot et Kadi Bouatouch : Temporal coherency in video tone mapping, a survey. *In HDRi2013 - First International Conference and SME Workshop on HDR imaging*, 2013.

[20] Ronan Boitard, Dominique Thoreau, Rémi Cozot et Kadi Bouatouch : Survey of temporal artifacts in video tone mapping. *In HDRi2014 - Second International Conference and SME Workshop on HDR imaging*, 2014.

[21] Ronan Boitard, Kadi Bouatouch, Rémi Cozot, Dominique Thoreau et Adrien Gruson : Temporal coherency for video tone mapping. *In SPIE*, volume 8499, 2012.

[22] Ronan Boitard, Dominique Thoreau, Rémi Cozot et Kadi Bouatouch : Impact of temporal coherence-based tone mapping on video compression. *In 21st European Signal Processing Conference (EUSIPCO 2013)*, Marrakech, Morocco, 2013.

[23] Ronan Boitard, Dominique Thoreau, Rémi Cozot et Kadi Bouatouch : Motion-guided quantization for video tone mapping. *In ICME 2014 - IEEE International Conference on Multimedia & Expo.*, 2014.

[24] Ronan Boitard, RÃ©mi Cozot, Dominique Thoreau et Kadi Bouatouch : Zonal brightness coherency for video tone mapping. *Signal Processing : Image Communication*, (0):–, 2013.

[25] A. Gruson, M. Ribardière et R. Cozot : Eye-centered color adaptation in global illumination. *Computer Graphics Forum*, 32(7):111–120, 2013.

[26] James T. Kajiya : The rendering equation. *SIGGRAPH Comput. Graph.*, 20(4):143–150, août 1986.

[27] Subrahmanyan Chandrasekhar : *Radiative Transfer*. Dover Publications, 1960.

[28] CIE : *Commission Internationale de l' Éclairage Proceedings*. Cambridge : Cambridge University Press, 1932.

[29] Mark D. Fairchild : *Color Appearance Model, Second edition*. John Wiley and sons, Ltd, 2005.

[30] Erik Reinhard, Greg Ward, Sumanta Pattanaik et Paul Debevec : *High Dynamic Range Imaging : Acquisition, Display, and Image-Based Lighting (The Morgan Kaufmann Series in Computer Graphics)*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2005.

[31] Francesco Banterle, Alessandro Artusi, Kurt Debattista et Alan Chalmers : *Advanced High Dynamic Range Imaging : Theory and Practice*. A. K. Peters, Ltd., Natick, MA, USA, 1st édition, 2011.

[32] Sumanta N. Pattanaik, Jack Tumblin, Hector Yee et Donald P. Greenberg : Time-dependent visual adaptation for fast realistic image display. *In Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '00, pages 47–54, New York, NY, USA, 2000. ACM Press/Addison-Wesley Publishing Co.

[33] Erik Reinhard, Michael Stark, Peter Shirley et James Ferwerda : Photographic tone reproduction for digital images. *ACM Trans. Graph.*, 21(3):267–276, juillet 2002.

[34] Rafal Mantiuk, Karol Myszkowski et Hans-Peter Seidel : A perceptual framework for contrast processing of high dynamic range images. *ACM Trans. Appl. Percept.*, 3(3):286–308, juillet 2006.

[35] F. Drago, K. Myszkowski, T. Annen et N. Chiba : Adaptive logarithmic mapping for displaying high contrast scenes. *Computer Graphics Forum*, 22: 419–426, 2003.

[36] Frédo Durand et Julie Dorsey : Fast bilateral filtering for the display of high-dynamic-range images. *ACM Trans. Graph.*, 21(3):257–266, juillet 2002.

[37] Raanan Fattal, Dani Lischinski et Michael Werman : Gradient domain high dynamic range compression. *ACM Trans. Graph.*, 21(3):249–256, juillet 2002.

[38] John J. McCann : Do humans discount the illuminant ? *In Proceedings of SPIE*, 2005.

[39] J. Von Kries : *Chromatic adaptation*, pages 109–119. MIT Press, 1970.

[40] D.H. Brainard et B.A. Wandell : asymmetric color matching : how color appearance depends on the illuminant. *Journal of the Optical Society of America A*, 9:1433–1448, 1992.

[41] E. J. Chichilnisky et B. A. Wandell : Photoreceptor sensitivity changes explain color appearance shifts induced by large uniform backgrounds in dichoptic matching. *Vision Research*, 53:239–254, 1995.

[42] Y. Nayatani, K. Takahama et H. Sobagaki : Formulation of anonlinear model of chromatic adaptation. *Color Research Application*, 6:161–171, 1981.

[43] Guth S. L. : Model for color vision and light adaptation. *Journal of the Optical Society of America A*, 8:976–993, 1991.

[44] E. J. Brenenam : Corresponding chromaticities for different states of adaptation to complex visual fields. *Journal of the Optical Society of America A*, 4:1115–1129, 1987.

[45] M. D. Fairchild : A model of incomplete chromatic adaptation. *In the 22nd Session of the CIE*, pages 33–34. CIE, 1991.

[46] M. D. Fairchild : Formulation and testing of an incomplete-chromatic-adaptation model. *Color Research Application*, 16:243–250, 1991.

[47] Martine Joly : *Introduction à l'analyse de l'image*. Armand Colin, 2009.

[48] Daniel Arijon : *Grammaire du langage filmé*. Dujarric, 1999.

[49] John Thomas Smith : *Remarks on Rural Scenery : With Twenty Etchings of Cottages, from Nature*. Nathanaiel Smith, 1797.

[50] Andrew Laszlo : *Every Frame a Rembrandt : Art and Practice of Cinematography*. Focal Press, 2000.

[51] Virginia Wissler : *Illuminated Pixels : The Why, What, and How of Digital Lighting*. Cengage Learning, 2012.

[52] Cecil B. DeMille : *The Autobiography of Cecil B. DeMille*. Prentice-Hall Inc., 1959.

[53] Ansel Adams : *The Negative : Exposure and Development*. Morgan and Lester, 1949.

[54] Ansel Adams : *Artificial Light Photography*. Morgan and Morgan, 1956.

[55] Kris Malkiewicz : *Film Lighting*. Fireside, 1992.

[56] Thomas M. Sipos : *Horror Film Aesthetics : Creating the Visual Language of Fear*. McFarland, 2010.

[57] Johann Wolfgang von Goethe : *Theory of Colours*. Dover Publications Inc., 2006.

[58] Michel Eugène Chevreul : *The Principles of Harmony and Contrast of Colours, and Their Applications to the Arts*. 1855.

[59] Jill L. Morton : *Color Logic*. Colorcom, 1998.

[60] Johannes Itten : *The Art of Colors*. Van Nostrand Reinhold Company, 1960.

[61] Daniel Cohen-Or, Olga Sorkine, Ran Gal et Tommer : Color harmonization. *ACM Transactions on Graphics (Proceedings of ACM SIGGRAPH)*, 25:1259–1282, 2006.

[62] Nikhil Sawant et Niloy J. Mitra : Color harmonization for videos. *In Proceedings of the 2008 Sixth Indian Conference on Computer Vision, Graphics & Image Processing (ICVGIP)*, pages 576–582. IEEE Computer Society, 2008.

[63] Xing Huo et Jieqing Tan : An improved method for color harmonization. *In Congress on Image and Signal Processing (CISP)*, 2009.

[64] Luaks Gruber, Denis Kalkofen et Dieter Schmalstieg : Color harmonization for augmented reality. *In IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 227–228, 2010.

[65] Alessio Bosca, Dario Bonino, Marco Comerio, Simone Grega et Fulvio Corno : A reusable 3d visualization component for the semantic web. *In Web3D '07 : Proceedings of the twelfth international conference on 3D web technology*, pages 89–96, New York, NY, USA, 2007. ACM.

[66] Lorenzo Sommaruga et Nadia Catenazzi : Curriculum visualization in 3d. *In Web3D '07 : Proceedings of the twelfth international conference on 3D web technology*, pages 177–180, New York, NY, USA, 2007. ACM.

[67] H. Halpin, D. Zielinski, R. Brady et G. Kelly : Exploring semantic social networks using virtual reality. *In Proceedings of the 7th International Conference on The Semantic Web*, page 614. Springer, 2008.

[68] B. Bos, H.W. Lie, C. Lilley et I. Jacobs : Cascading style sheets, level 2 CSS2 specification, 1998.

[69] Graham Klyne, Jeremy J. Carroll et Brian McBride : Resource description framework (rdf) : concepts and abstract syntax, 2004.

[70] Patrick Durusau, Steve Newcomb et Robert Barta : Iso/iec 13250 topic maps, 1999.

[71] Alvy Ray Smith : *Plants, fractals, and formal languages*. Proceedings of the 11th annual conference on Computer graphics and interactive techniques. ACM, 1984. 808571.

[72] Michael Kay : Xsl transformations (xslt) version 2.0, January 2007.

[73] Priscilla Walmsley : *XQuery*. O'Reilly Media, Inc., 2007.

[74] World Wide Web Consortium : X3d specifications, 2004.

[75] Catherine Leung et Andor Salga : Enabling webgl. *In Proceedings of the 19th International Conference on World Wide Web*, WWW '10, pages 1369–1370, New York, NY, USA, 2010. ACM.

[76] Remi Arnaud et Mark C. Barnes : *Collada : Sailing the Gulf of 3D Digital Content Creation*. AK Peters Ltd, 2006.

[77] Jed Hartman et Josie Wernecke : *The VRML 2.0 Handbook : The Official Guide to Constructing Virtual Worlds*. Addison-Wesley, Reading, MA, 1996.

[78] Yoav I. H. Parish et Pascal Müller : Procedural modeling of cities. *In Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '01, pages 301–308, New York, NY, USA, 2001. ACM.

[79] Graham Whelan, George Kelly et Hugh McCabe : Roll your own city. *In Proceedings of the 3rd International Conference on Digital Interactive Media in Entertainment and Arts*, DIMEA '08, pages 534–535, New York, NY, USA, 2008. ACM.

[80] Wojciech Jarosz, Derek Nowrouzezahrai, Robert Thomas, Peter-Pike Sloan et Matthias Zwicker : Progressive photon beams. *ACM Transactions on Graphics (Proceedings of ACM SIGGRAPH Asia 2011)*, 30(6), décembre 2011.

[81] Claude Knaus et Matthias Zwicker : Progressive photon mapping : A probabilistic approach. *ACM Trans. Graph.*, 30:25 :1–25 :13, mai 2011.

[82] Toshiya Hachisuka, Shinji Ogaki et Henrik Wann Jensen : Progressive photon mapping. *ACM Trans. Graph.*, 27:130 :1–130 :8, December 2008.

[83] Henrik Wann Jensen et Per H. Christensen : Efficient simulation of light transport in scences with participating media using photon maps. *In Proceedings of the 25th annual conference on Computer graphics and interactive techniques*, SIGGRAPH '98, pages 311–320, New York, NY, USA, 1998. ACM.

[84] Toshiya Hachisuka et Henrik Wann Jensen : Robust adaptive photon tracing using photon path visibility. *ACM Trans. Graph.*, 30(5):114 :1–114 :11, octobre 2011.

[85] Wojciech Jarosz, Matthias Zwicker et Henrik Wann Jensen : The beam radiance estimate for volumetric photon mapping. *Computer Graphics Forum (Proceedings of Eurographics 2008)*, 27(2):557–566, avril 2008.

[86] Vlastimil Havran, Jiří Bittner et Hans-Peter Seidel : Ray maps for global illumination. *In ACM SIGGRAPH 2004 Sketches*, SIGGRAPH '04, pages 77–, New York, NY, USA, 2004. ACM.

[87] Wenzel Jakob : Mitsuba renderer, 2010. http ://www.mitsuba-renderer.org.

[88] Thomas Luft, Carsten Colditz et Oliver Deussen : Image enhancement by unsharp masking the depth buffer. *ACM Trans. Graph.*, 25(3):1206–1213, juillet 2006.

[89] Matthias B. Hullin, Johannes Hanika et Wolfgang Heidrich : Polynomial Optics : A construction kit for efficient ray-tracing of lens systems. *Computer Graphics Forum (Proceedings of EGSR 2012)*, 31(4), juillet 2012.

[90] Michael Potmesil et Indranil Chakravarty : A lens and aperture camera model for synthetic image generation. *SIGGRAPH Comput. Graph.*, 15(3): 297–305, août 1981.

[91] L. Itti : Quantifying the contribution of low-level saliency to human eye movements in dynamic scenes. *Visual Cognition*, 12(6):1093–1123, Aug 2005.

[92] K. Cater, A. Chalmers et G. Ward : Detail to attention : Exploiting visual tasks for selective rendering. *In Proceedings of the 14th Eurographics Workshop on Rendering*, EGRW '03, pages 270–280, Aire-la-Ville, Switzerland, Switzerland, 2003. Eurographics Association.

[93] L. Itti, C. Koch et E. Niebur : A model of saliency-based visual attention for rapid scene analysis. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 20(11):1254–1259, 1998.

[94] Anne M. Treisman et Garry Gelade : A feature-integration theory of attention. *Cognitive Psychology*, 12(1):97–136, janvier 1980.

[95] Sungkil Lee, G. Jounghyun Kim et Seungmoon Choi : Real-time tracking of visually attended objects in virtual environments and its application to lod. *Visualization and Computer Graphics, IEEE Transactions on*, 15(1):6–19, 2009.

[96] A. L. Yarbus : *Eye Movements and Vision*. Plenum. New York., 1967.

[97] V. Navalpakkam et L. Itti : Modeling the influence of task on attention. *Vision Research*, 45(2):205–231, Jan 2005.

[98] Peter Longhurst, Kurt Debattista et Alan Chalmers : A gpu based saliency map for high-fidelity selective rendering. *In Proceedings of the 4th International Conference on Computer Graphics, Virtual Reality, Visualisation and Interaction in Africa*, AFRIGRAPH '06, pages 21–29, New York, NY, USA, 2006. ACM.

[99] Veronica Sundstedt, Efstathios Stavrakis, Michael Wimmer et Erik Reinhard : A Psychophysical Study of Fixation Behavior in a Computer Game. *In* Sarah Creem-Regehr et Karol Myszkowski, éditeurs : *Proceedings of the 5th Symposium on Applied Perception in Graphics and Visualization*, pages 43–50, Los Angeles, États-Unis, 2008. ACM.

[100] Hanspeter Pfister, Matthias Zwicker, Jeroen van Baar et Markus Gross : Surfels : Surface elements as rendering primitives. *In Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '00, pages 335–342, New York, NY, USA, 2000. ACM Press/Addison-Wesley Publishing Co.

[101] Ricardo Marroquim et André Maximo : Introduction to gpu programming with glsl. *In Proceedings of the 2009 Tutorials of the XXII Brazilian Symposium on Computer Graphics and Image Processing*, SIBGRAPI-TUTORIALS '09, pages 3–16, Washington, DC, USA, 2009. IEEE Computer Society.

[102] Brian A. Barsky : Vision-realistic rendering : simulation of the scanned foveal image from wavefront data of human subjects. *In in : First Symposium on Applied Perception in Graphics and Visualization, colocated with ACM SIGGRAPH*, pages 73–81, 2004.

[103] P. Rokita : Generating depth of-field effects in virtual reality applications. *Computer Graphics and Applications, IEEE*, 16(2):18–21, 1996.

[104] Joe Demers : *Depth of Field : A Survey of Techniques in GPU Gems.* Addison-Wesley, 2004.

[105] Alan Kenny, Hendrik Koesling et Declan Delaney : A preliminary investigation into eye gaze data in a first person shooter game. *In In Proceedings of European Conference on Modelling and Simulation, Addison-Wesley Professional*, 2005.

[106] Natalya Tatarchuk Guennadi Riguer et John Isidoro : *Real-time depth of field simulation in ShaderX2 : Shader Programming Tips and Tricks with DirectX 9.* Wordware, 2003.

[107] Michael D. Tocci, Chris Kiser, Nora Tocci et Pradeep Sen : A versatile hdr video production system. *ACM Trans. Graph.*, 30(4):41 :1–41 :10, juillet 2011.

[108] Joel Kronander, Stefan Gustavson, Gerhard Bonnet et Jonas Unger : Unified hdr reconstruction from raw cfa data. *IEEE International Conference on Computational Photography (ICCP)*, 0:1–9, 2013.

[109] Erik Reinhard, Michael Stark, Peter Shirley et James Ferwerda : Photographic tone reproduction for digital images. *ACM Trans. Graph.*, 21(3):267–276, juillet 2002.

[110] Zeev Farbman et Dani Lischinski : Tonal stabilization of video. *ACM Trans. Graph.*, 30(4):89 :1–89 :10, juillet 2011.

[111] RafaÅ, Mantiuk, Scott Daly et Louis Kerofsky : Display adaptive tone mapping. *ACM Transactions on Graphics*, 27(3):1, août 2008.

[112] SD Ramsey, JT Johnson III et Charles Hansen : Adaptive temporal tone mapping. *Computer Graphics and Imaging - 2004*, (3):3–7, 2004.

[113] Sing Bing Kang, Matthew Uyttendaele, Simon Winder et Richard Szeliski : High dynamic range video. *ACM Trans. Graph.*, 22(3):319–325, juillet 2003.

[114] Chul Lee et Chang-Su Kim : Gradient domain tone mapping of high dynamic range videos. *In 2007 IEEE International Conference on Image Processing*, volume 3, pages III – 461. IEEE, 2007.

[115] Benjamin Guthier, Stephan Kopf, Marc Eble et Wolfgang Effelsberg : Flicker reduction in tone mapped high dynamic range video. *In Proc. of IS&T/SPIE Electronic Imaging (EI) on Color Imaging XVI : Displaying, Processing, Hardcopy, and Applications*, 2011.

[116] Ronan Boitard, Kadi Bouatouch, Remi Cozot, Dominique Thoreau et Adrien Gruson : Temporal coherency for video tone mapping. *In Proc. SPIE 8499, Applications of Digital Image Processing XXXV*, octobre 2012.

[117] Yuanzhen Li, Lavanya Sharan et Edward H Adelson : Compressing and companding high dynamic range images with subband architectures. *ACM Transactions on Graphics*, 24(3):836, juillet 2005.

[118] Alain Tremeau, Shoji Tominaga et Konstantinos N. Plataniotis : Color in image and video processing : Most recent trends and future research directions. *EURASIP Journal on Image and Video Processing*, 2008, 2008.

[119] Kuk-Jin YOON, Yoo Jin CHOFI et In-So KWEON : Dichromatic-based color constancy using dichromatic slope and dichromatic line space. *In Image Processing, 2005. ICIP 2005. IEEE International Conference on*, volume 3, pages III–960–3, 2005.

[120] EDWIN H. LAND et JOHN J. MCCANN : Lightness and retinex theory. *J. Opt. Soc. Am.*, 61(1):1–11, Jan 1971.

[121] Jean-Michel MOREL, Ana B. PETRO et Catalina SBERT : Fast implementation of color constancy algorithms. *In Proc. SPIE 7241, Color Imaging XIV : Displaying, Processing, Hardcopy, and Applications, 724106 (January 19, 2009)*, 2009.

[122] Jean Michel MOREL, Ana Belén PETRO et Catalina SBERT : A pde formalization of retinex theory. *Trans. Img. Proc.*, 19(11):2825–2837, novembre 2010.

[123] E. H. LAND : The retinex theory of color vision. *ScientiPc American*, 237(6): 108–120, 1977.

[124] G.. BUCHSBAUM : A spatial processor model for object colour perception. *Journal of the Franklin Institute*, 310, 1980.

[125] Joost Van de WEIJER, Cordelia SCHMID et Jakob VERBEEK : Using high-level visual information for color constancy. *In IEEE Conference on Computer Vision (ICCV)*, 2007.

[126] G. D. FINLAYSON, S. D. HORDLEY et I. TASTL : Gamut constrained illuminant estimation. *Int. J. Comput. Vision*, 67(1):93–109, 2006.

[127] P.V. GEHLER, C. ROTHER, A. BLAKE, T. MINKA et T. SHARP : Bayesian color constancy revisited. pages 1–8, June 2008.

[128] V. CARDEI, B. FUNT et K. BARNARD : Estimating the scene illumination chromaticity using a neural network. *Journal of the Optical Society of America*, 19(12), 2002.

[129] S. D. HORDLEY : Scene illuminant estimation : past, present, and future. *Color Research and Application*, 31(4):303–314, 2006.

[130] Graham D. FINLAYSON et Elisabetta TREZZI : Shades of Gray and Colour Constancy. *In Color Imaging Conference*, pages 37–41, 2004.

[131] Arjan GIJSENIJ, Theo GEVERS et Joost WEIJER : Generalized gamut mapping using image derivative structures for color constancy. *Int. J. Comput. Vision*, 86(2-3):127–139, 2010.

[132] Alexa I. RUPPERTSBERG et Marina BLOJ : Reflecting on a room of one reflectance. *Journal of Vision*, 7(13):–, 2007.

[133] Timo KUNKEL et Erik REINHARD : A neurophysiology-inspired steady-state color appearance model. *Journal of the Optical Society of America A*, 26(4): 776–782, April 2009.

[134] Y. N. MORONE, M. D. FAIRCHILD, R. W. G. HUNT, C. LI, M. R. LOU et T. NEWMAN : The ciecam02 color appearance model. *In In Color Imaging Conference (2002), IS&T*, pages 23–27. Society for Imaging Science and Technology, 2002.

[135] J. Kuang, G. M. Johnson et M. D. Fairchild : icam06 : A refined image appearance model for hdr image rendering. *Journal of Visual Communication*, 2007.

[136] Erik Reinhard, Tania Pouli, Timo Kunkel, Ben Long, Anders Ballestad et Gerwin Damberg : Calibrated image appearance reproduction. *ACM Trans. Graph.*, 31(6):201 :1–201 :11, novembre 2012.

[137] Greg Ward et Elena Eydelberg-Vileshin : Picture perfect rgb rendering using spectral prefiltering and sharp color primaries. *In EGRW '02 : Proceedings of the 13th Eurographics workshop on Rendering*, pages 117–124, Aire-la-Ville, Switzerland, Switzerland, 2002. Eurographics Association.

[138] Laszlo Neumann, Francesc Castro, Attila Neumann et Mateu Sbert : Color appearance in multispectral radiosity. *In* G. Renner L. Szirmay-Kalos, éditeur : *Proceedings on the 2nd Hungarian Computergraphics and Geometry Conference*, pages 183–194, 2003.

[139] A. Wilkie et A. Weidlich : A robust illumination estimate for chromatic adaptation in rendered images. *In Eurographics Symposium on Rendering 2009*, 2009.

[140] Penousal Machado et Amílcar Cardoso : Computing aesthetics. *In Proceedings of the Brazilian Symposium on Artificial Intelligence, SBIA'98*, pages 219–229. Springer-Verlag, 1998.

[141] Florian Hoenig : Defining computational aesthetics. *In Proceedings of the First Eurographics Conference on Computational Aesthetics in Graphics, Visualization and Imaging*, Computational Aesthetics'05, pages 13–18, Aire-la-Ville, Switzerland, Switzerland, 2005. Eurographics Association.

[142] Ritendra Datta, Dhiraj Joshi, Jia Li et James Z. Wang : Studying aesthetics in photographic images using a computational approach. *In Proceedings of the 9th European Conference on Computer Vision - Volume Part III*, ECCV'06, pages 288–301, Berlin, Heidelberg, 2006. Springer-Verlag.

[143] Vera Rivotti, João Proença, Joaquim Jorge et Mário Costa Sousa : Composition principles for quality depiction and aesthetics. *In Proceedings of the Third Eurographics Conference on Computational Aesthetics in Graphics, Visualization and Imaging*, Computational Aesthetics'07, pages 37–44, Aire-la-Ville, Switzerland, Switzerland, 2007. Eurographics Association.

[144] S. Dhar, V. Ordonez et T. L. Berg : High level describable attributes for predicting aesthetics and interestingness. *In Proceedings of the 2011 IEEE Conference on Computer Vision and Pattern Recognition*, CVPR '11, pages 1657–1664, Washington, DC, USA, 2011. IEEE Computer Society.

[145] Shehroz S. Khan et Daniel Vogel : Evaluating visual aesthetics in photographic portraiture. *In Proceedings of the Eighth Annual Symposium on Computational Aesthetics in Graphics, Visualization, and Imaging*, CAe '12, pages 55–62, Aire-la-Ville, Switzerland, Switzerland, 2012. Eurographics Association.

[146] Julien Sicre : Classifying images by light aesthetics. Mémoire de D.E.A., Univeristé de Rennes 1, 2013.

[147] Erik Reinhard, Michael Ashikhmin, Bruce Gooch et Peter Shirley : Color transfer between images. *IEEE Comput. Graph. Appl.*, 21(5):34–41, septembre 2001.

[148] Baoyuan Wang, Yizhou Yu et Ying-Qing Xu : Example-based image color and tone style enhancement. *ACM Trans. Graph.*, 30(4):64 :1–64 :12, juillet 2011.

[149] Ahmet Oğuz Akyüz, Kerem Hadimli, Merve Aydinlilar et Christian Bloch : Style-based tone mapping for hdr images. *In SIGGRAPH Asia 2013 Technical Briefs*, SA '13, pages 23 :1–23 :4, New York, NY, USA, 2013. ACM.

[150] Nicolas Bonneel, Kalyan Sunkavalli, Sylvain Paris et Hanspeter Pfister : Example-based video color grading. *ACM Trans. Graph.*, 32(4):39 :1–39 :12, juillet 2013.

[151] Ram Shacked et Dani Lischinski : Automatic lighting design using a perceptual quality metric. *Computer Graphics Forum*, 20:2001, 2001.

[152] Vincent Leon : Global illumination and image appearance. Mémoire de D.E.A., Université de Rennes 1, 2013.

Table des figures

2.1 Éléments clés de la synthèse d'images physique. 7

3.1 Élément géométrique de l'équation 3.1. 17
3.2 Éfficacité lumineuse relative spectrale. 18
3.3 Fonctions $\bar{r}$, $\bar{g}$ et $\bar{b}$ de l'espace de couleur CIE RGB. 19
3.4 Résultats de la réduction des gammes de luminances avec des opérateur différents (de gauche à droite et de bas en haut : [34], [35], [36], [37], [32] et [33]). 21
3.5 Conditions d'observation. 22
3.6 Image brute : couleur physique (gauche), Image après adaptation chromatique : couleurs perçues, comparaison entre la couleur physique et la couleur perçue du tapis (droite). 22
3.7 Composition et règle des tiers : les éléments importants de la photo sont placés sur les droites correspondant aux tiers verticaux et horizontaux et sur leurs intersections (Crédit photo : armor-photo.com). 24
3.8 Formes de réflecteurs différentes permettant de contrôler le flux de sortie de la lumière (source : Elinchrom). 25
3.9 Éclairage à 3 sources : *key light*, *fill light* et *back light*. 26
3.10 Éclairage à 3 sources : le ratio de puissance *key light*/*fill light* contrôle le contraste entre les zones claires (éclairées par la *key light*) et les zones sombres (éclairées par la *fill light*). La *back light* sépare le sujet de la photo du fond et met en valeur la silhouette en créant des éclats sur les contours. (Crédit photo : armor-photo.com) 27
3.11 Harmonies des couleurs : (haut-gauche) Analogue, (haut-centre) Complémentaire, (haut-droite) Complémentaire à 2 tons, (bas-gauche) 3 tons, (bas-centre) Complémentaire à 4 tons, (bas-gauche) 4 tons. . . 28
3.12 Influence des couleurs des sources de lumière : (à gauche) *fill light* de couleur verte qui contrôle la couleur des zones non éclairées par la *key light* – (à droite) *back light* de couleur bleue donnant un reflet bleu sur la silhouette du sujet. (Crédit photo : armor-photo.com) . . 29
3.13 Des couleurs froides aux couleurs chaudes (en mireds). 30
3.14 Paramètre de rendu (à droite) image obtenue avec les paramètres par défaut - (à gauche) image obtenue avec un changement des paramètres de l'opérateur de réduction des gammes de luminances. . . 30

4.1 Architecture de l'environnement de création automatique de monde 3D à partir de bases de données : séparation entre structuration des données (génome et génotype) et visualisation (phénome et phénotype). 36
4.2 Définition de la structure hiérarchique d'exploration par le génome. . 37
4.3 Exemple de *génotype* extrait de la base de données à partir des niveaux d'exploration définis par le *génome*. 37
4.4 Le *phénome* décrit l'aspect et l'organisation des éléments du *génotype*. 39
4.5 La place 3D (niveau d'exploration 0). 41
4.6 Le cinéma (niveau d'exploration 1 - les genres). 42
4.7 La salle et les affiches de films (niveau d'exploration 2,3 - les réalisateurs, les films). 43
4.8 La place avec les cinémas - un cinéma par genre. 44
4.9 L'intérieur du cinéma avec les salles (accès par l'escalier) - une salle par réalisateur. 44
4.10 La salle de cinéma et les affiches des films (en bas de l'écran) - une affiche par film. 44
4.11 Le monde 3D (*phénotype*) construit automatiquement à partir de la base de données de films. 44

5.1 Même scène 3D avec (à droite) prise en compte dans la simulation des milieux participatifs - (à gauche) sans milieux participatifs. . . . 45
5.2 Chemin de rayon : cylindre quand le rayon traverse un volume (en vert) et point visible (*view point*) (en rouge) sur les surfaces diffuses et brillantes (non spéculaires). 50
5.3 Courbe de RMSE (entre l'image de référence et les différentes méthodes) en fonction du temps pour la scène Breakfast hall. 53
5.4 La scène Breakfast hall (avec l'autorisation de Greg Zaal). (En haut à gauche) image de référence. (Colonne de droite) images obtenues avec les autres méthodes 10h. Détails sur les images obtenues avec les différentes méthodes après : 30min. et 10h. Et images de différences par rapport à l'image de références. 54
5.5 La scène Kitchen (avec l'autorisation de Jay-Artist). (En haut à gauche) image obtenue avec PPT_metro (10h). (Colonne de droite) images obtenues avec les autres méthodes 10h. Détails sur les images obtenues avec les différentes méthodes après : 30min. et 10h. 55
5.6 Même scène 3D avec (en bas) prise en compte dans la simulation des milieux participatifs - (en haut) sans milieux participatifs. 56

6.1 Architecture de notre Modèle d'Attention Visuelle : composants *bottom-up*, *top-down* et intégration *bottom-up* & textittop-down. 61
6.2 Monde 3D utilisé pour les tests du Modèle d'Attention Visuelle : (à gauche) vue de l'environnement 3D - (à droite) rendu pour l'utilisateur. 62
6.3 Zone d'auto-focus (rectangle blanc) et influence du poids sémantique : (à gauche) auto-focus sans prise en compte de la carte sémantique - (à droite) auto-focus avec prise en compte de la sémantique : le net est fait sur le personnage. 64

6.4 Flou de profondeur : (à gauche) flou de profondeur sans rotation du noyau d'échantillonnage - (au milieu) flou de profondeur avec rotation du noyau d'échantillonnage - (à droite) image finale avec flou de profondeur. 64
6.5 Scintillement dans la séquence LDR : exemple de trois images consécutives, l'image du milieu est plus lumineuse. 67
6.6 Altération de la cohérence du contraste : (en haut) Les images HDR en fausse couleur (bleu : faible valeur de luminance, rouge : grande valeur de luminance), (en bas) Les images LDR correspondantes - Le changement de condition d'illumination entraine des images HDR plus sombres, ce changement n'est pas respecté dans les images LDR produites. 68
6.7 Altération de la cohérence des objets : (en haut) Les images HDR en fausse couleur (bleu : faible valeur de luminance, rouge : grande valeur de luminance), (en bas) Les images LDR correspondantes - la luminance de la partie inférieure du pont évolue dans la séquence LDR alors qu'elle est constante dans la séquence HDR. 69
6.8 Les différents types d'altération de la cohérence temporelle - (en haut) scintillement, (au milieu) altération de cohérence du contraste, (en bas) altération de la cohérence temporelle des objets. 70
6.9 Fonction de pondération des facteurs d'échelle. 73
6.10 La séquence HDR *Under Bridge* - (à gauche) des images de séquence LDR obtenu avec l'opérateur de réduction de Li [117] - (à droite) les images LDR avec ZBC : la cohérence temporelle de la luminosité est maintenue avec la séquence HDR et le contraste est conservé. 74
6.11 La séquence HDR *Under Bridge* - (à gauche) des images de séquence LDR obtenu avec l'opérateur de réduction [109] - (à droite) les images LDR avec ZBC : la cohérence temporelle de la luminosité est conservée malgré le changement de conditions d'illumination sur le véhicule principal. 75
6.12 Les deux étapes de l'adaptation chromatique 77
6.13 Moyenne de la couleur de l'irradiance sur le capteur hémisphérique comme couleur d'adaptation. 82
6.14 Architecture overview of our chromatic adaptation process 83
6.15 Vue du dessus de la scène composée de deux pièces et visualisation des trois champs de vision correspondant au trois positions et orientations de la caméra (rouge, vert, bleue) 85
6.16 Diagramme de chromaticité dans l'espace de couleur RGB. Les cercles de couleur (rouge, vert, bleu) correspondent à la couleur d'adaptation pour les trois positions et orientations de la caméra (Cf. Figure 6.15). La figure de gauche correspond à la méthode WCAM. La figure de droite montre les résulats pour notre méthode. 86

6.17 Cohérence spatiale et temporelle de l'estimation de la couleur d'adaptation dans le cas de deux pièces ayant des éclairages de couleurs différentes. Image brute à gauche, image adaptée selon WCAM au centre, image adaptée avec notre méthode à droite. 87

7.1 Chaîne de la synthèse d'images physique avec nos principales contributions. 89
7.2 Chaîne de la synthèse d'images physique avec prise en compte de l'intentien styliste/esthétique du créateur 93
7.3 Optimisation automatique des paramètres des sources de lumière pour obtenir une esthétique *low key* : (à gauche) image initiale - (à droite) image après optimisation des paramètres des sources de lumières [152] 95

Liste des tableaux

2.1 Les enjeux par domaines d'applications 11
2.2 Résumé des publications 14

3.1 Ordre de grandeur des luminances 20

4.1 Association des modèles 3D en fonction de niveau d'exploration. . . 41

5.1 Définitions de notations et grendeurs utilisées 49
5.2 Paramètre du rendu : chaque méthode utilise le même temps de rendu (dernière ligne). *photons volumiques* = nombre de photons émis par passe dans APA et PPT ; *faisceaux de photons* = nombre de faisceaux de photons dans PPB et PPBT. 52