



HAL
open science

Génération, visualisation et évaluation d'images HDR : application à la simulation de conduite nocturne

Josselin Petit

► **To cite this version:**

Josselin Petit. Génération, visualisation et évaluation d'images HDR : application à la simulation de conduite nocturne. Autre [cs.OH]. Université Claude Bernard - Lyon I, 2010. Français. NNT : 2010LYO10259 . tel-00707723

HAL Id: tel-00707723

<https://theses.hal.science/tel-00707723>

Submitted on 13 Jun 2012

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.

N° d'ordre : 259- 2010

Année 2010

THÈSE

présentée par

Josselin PETIT

en vue d'obtenir le grade de

Docteur en informatique de l'Université Claude Bernard Lyon I

Ecole doctorale : l'InfoMaths

Génération, visualisation et évaluation d'images HDR. Application à la simulation de conduite nocturne.

Date de soutenance : le 03 décembre 2010

Composition du jury :

Bernard PEROCHE , Université Claude Bernard Lyon I	Directeur de thèse
Karol MYSZKOWSKI , Max-Planck-Institut für Informatik	Rapporteur
Mathias PAULIN , Université Paul Sabatier	Rapporteur
Andras KEMENY , Renault	Examineur
Venceslas BIRI , Université de Marne-la-Vallée	Examineur
Roland BREMOND , LEPSiS	Examineur

Thèse préparée au Laboratoire Exploitation Perception Simulateurs et Simulations (LEPSiS, UMR INRETS-LCPC)

Remerciements

Je voudrais d'abord remercier les membres du jury. Les rapporteurs, Karol Myzkowsky et Mathias Paulin, pour leurs remarques et critiques constructives, et pour avoir parfaitement respecté les délais que je leur avais imposé. Un merci tout particulier à Karol pour avoir lu le mémoire et suivi la soutenance en français.

Je remercie également les examinateurs, Venceslas Biri et Andras Kemeny, pour avoir répondu présents.

Je remercie ensuite mon directeur de thèse, Bernard Péroche, pour ses précieux conseils et ses nombreux déplacements au cours de ces trois années.

Enfin, je remercie chaleureusement mon encadrant, Roland Brémond, pour son implication dans ma thèse, ses conseils, sa patience, son humour pince-sans-rire et son soutien de tous les instants, qui m'ont bien aidé.

Je voudrais ensuite remercier l'ensemble des personnes qui ont participé, de près ou de loin, au bon déroulement de ma thèse. Parmi elles, je pense aux personnes présentes à mes comités de suivi, Venceslas Biri, Roland Brémond, Eric Dumont, Stéphane Espié, Bernard Péroche et Jean-Philippe Tarel, qui tous les ans m'ont aidé à faire le point et à décider des meilleures approches à adopter pour la conduite de mes travaux.

Un merci tout particulier à Fabrice Vienne, avec qui j'ai porté la réalisation du rendu HDR et son intégration dans le simulateur du LEPSiS. Merci pour ta constante bonne humeur contagieuse.

Je voudrais également saluer Jean-Philippe Tarel, avec qui j'ai travaillé sur la saillance et avec qui je me suis toujours bien marré.

Ensuite, je voulais féliciter Eric Dumont pour son encadrement, car dans sa cheffitude, il a toujours su anticiper mes besoins matériels et y répondre avant même que je les formule. Je suis certain que tu continueras à faire un très bon chef à l'avenir.

Pour les conseils et les discussions que j'ai pu avoir avec eux, je voulais saluer Jean-Philippe Farrugia, Dominique Gruyer, Jean-Claude Iehl et François de Sorbier.

Je souhaite également bonne chance à David Callu pour la fin du développement du simulateur.

Ensuite, je souhaite remercier les infographistes Jacky Robouant et Jean-Sebastien Barreiro, pour avoir réalisé les modèles 3D, sans quoi je n'aurais jamais pu réaliser ce rendu HDR.

Merci également à toutes les personnes qui sont courageusement venues passer

mes expérimentations.

Merci aussi aux gestionnaires du LEPSiS, qui se sont toujours très bien occupé de moi, et se sont toujours appliqué à me simplifier mes démarches administratives.

Je veux également remercier tous les membres passés et présents du LEPSiS, avec qui j'ai eu de très bonnes relations pendant les trois années, et notamment Vanessa Bodard, qui m'a bien dépanné en m'apprenant le langage bash quand j'ai eu besoin de traiter des flux d'images HDR.

Je veux également saluer l'ensemble des membres de la Commission des Doct-rants, et notamment Omar Khemoudj et Alice Sadd, avec qui j'ai passé de très bons moments.

Enfin, je voulais adresser un énorme clin d'œil aux copains de bureau, Karine, Jérémie, Raouf et Romain, avec qui je n'ai partagé que des bons moments.

Un clin d'œil également à Hayet, qui est partie trop tôt du LCPC.

Et un dernier clin d'œil aux collègues de l'apéro du vendredi soir, que j'aimerais continuer à retrouver autour d'un picon bière.

Plus personnellement, je voulais remercier l'ensemble de mes amis, qui m'entourent et me soutiennent depuis bien longtemps. Un coucou à Erik, Flavien, Maxime et Thomas, avec qui j'espère bien continuer à passer de nombreuses soirées. Un coucou à Caro, Christophe, Claire, Jordane, Maelle, Olivier, Stéphanie, Stéphanie et Titoune, avec qui j'ai passé et continue de passer d'excellents moments. Un coucou à Galou, Lauriane, Marc, Mikk, Paul, Pascal, que j'ai plaisir à retrouver régulièrement pour capser et passer de supers moments.

Encore plus personnellement, je tiens à remercier l'ensemble de ma famille, notamment mes parents, mes grands-parents et ma sœur, pour son soutien de tous les instants, et notamment pour m'avoir aidé à faire tout ce qui me tenait à cœur, depuis toujours. Merci également pour m'avoir préparé un super pot de soutenance. Merci particulier à mes parents, pour tout. Enfin, merci à Julie, pour être ce qui tu es auprès de moi.

Résumé

Cette thèse se situe à l'interface de deux des sujets de recherche du LEPSi8S, la perception et la réalité virtuelle, appliqués aux transports routiers. L'objectif de la thèse est d'améliorer l'état de l'art concernant le rendu des images de synthèse pour les simulateurs de conduite. L'axe privilégié est le réalisme perceptif des images.

L'approche retenue propose un mode de rendu High Dynamic Range, qui permet de générer une image en luminance. La technique proposée permet de réutiliser des environnements virtuels classiques, avec un minimum d'informations supplémentaires concernant les sources lumineuses. Les textures et matériaux existants sont utilisés pour un rendu aussi proche physiquement de la réalité que possible.

Ensuite, l'image est traitée avec un opérateur de reproduction de tons, qui compresse la dynamique pour tenir compte des limites liées au dispositif d'affichage, tout en respectant autant que possible un réalisme perceptif du rendu. L'opérateur a été choisi de façon à ce qu'il soit adapté à la simulation de conduite, notamment pour les cas extrêmes (nuit, éblouissement, soleil rasant). Une simulation de l'éblouissement a également été implémentée. L'ensemble du rendu est temps réel, et a été intégré dans la boucle visuelle des simulateurs de conduite du LEPSiS.

Enfin, des comparaisons réel-virtuel ont permis de montrer la qualité du rendu HDR obtenu. Des expérimentations avec sujets, sur des photographies (avec une référence réelle) et sur des vidéos, ont de plus montré les meilleures performances d'un opérateur doté d'un modèle visuel humain pour la simulation de conduite, notamment par sa capacité à s'adapter temporellement aux variations de luminance.

Mots clés : image HDR, reproduction de tons, temps réel, adaptation, Irawan, évaluation.

Abstract

Rendering, visualization and evaluation of HDR images.
Application to driving simulation at night.

The LEPSiS is leading applied research on the transportation field. This PhD addresses perception and virtual reality, two research topics at the LEPSiS. The objective of my PhD was to improve the state of the art of the computer graphic image rendering for driving simulator applications. The main issue was the perceptual realism of the images, notably in high dynamic range conditions (night, glare).

The proposed approach puts forward a High Dynamic Range mode, allowing us to render images in luminance. We use classic virtual environments, with small additional information about the light sources. The textures and materials are used for a rendering as close as possible to physical reality.

Then, the image is processed by a tone mapping operator, which compresses the luminance dynamic, taking into account the limited range of the display device and the perceptual realism of the rendering. The chosen tone mapping is adapted to driving simulations, and especially to extreme situations (night, skimming sun). A glare simulation was also added. The entire rendering is real time, and is now included in the driving simulators of the LEPSiS.

Lastly, real-virtual comparisons assessed the quality of the obtained HDR rendering. Moreover, two psycho-visual experiments with subjects, on photographs (with a real reference) and on video (without reference), showed the relevance of a tone mapping with a human visual model, including temporal adaptation to changing luminance, for driving simulations.

Keywords : HDR image, tone mapping, real time, adaptation, Irawan, assessment.

Table des matières

1	Introduction	13
1.1	Synthèse d'image : modélisation et rendu	13
1.2	Sécurité routière et simulation de conduite	14
1.2.1	La sécurité routière en France	14
1.2.2	La simulation de conduite, un outil pour la prévention	16
1.2.3	Le rendu visuel au sein des simulateurs	17
1.2.4	La question de l'éclairage	17
1.3	Besoins identifiés	20
1.4	Objectifs de la thèse	21
1.5	Organisation du manuscrit	23
2	Lumière et couleur	25
2.1	Lumière	25
2.2	Radiométrie	25
2.3	Photométrie	27
2.3.1	Grandeurs photométriques	27
2.3.2	Efficacité lumineuse	30
2.4	Colorimétrie	31
2.5	Espaces de couleur	32
2.6	Diagramme de chromaticité	34
2.7	Propriétés des afficheurs	35
2.7.1	Luminance des afficheurs	35
2.7.2	Contraste des afficheurs	36
2.7.3	Gamma des afficheurs	37
3	Le système visuel humain	39
3.1	Anatomie de l'œil	40

3.1.1	L'œil humain	40
3.1.2	La rétine	40
3.1.3	Les bâtonnets	42
3.1.4	Les cônes	43
3.2	Perception visuelle	44
3.2.1	Perception des contrastes	45
3.2.2	Luminance d'adaptation	46
3.2.3	Adaptation temporelle	47
3.2.4	Éblouissement	51
4	Etat de l'art sur le rendu visuel	55
4.1	Le rendu HDR	56
4.1.1	Imagerie HDR par photographie	56
4.1.2	Imagerie HDR par synthèse d'images	58
4.1.3	Imagerie HDR par photographie pour la synthèse d'images	60
4.1.4	Manipulation d'images HDR	60
4.2	Les opérateurs de reproduction de tons	61
4.2.1	Symboles et définitions	63
4.2.2	Les opérateurs globaux statiques	64
4.2.3	Les opérateurs globaux dynamiques	71
4.2.4	Les opérateurs locaux statiques	84
4.2.5	Un opérateur local dynamique	88
4.2.6	Les opérateurs fréquentiels	89
4.2.7	Un opérateur gradient	89
4.3	Evaluation des opérateurs de reproduction de tons	90
4.4	Simulation de l'éblouissement	93
4.5	Saillance d'images et de vidéos	96
4.6	Conclusion	98
5	Rendu HDR	101
5.1	Génération d'images de synthèse HDR	102
5.1.1	Environnements virtuels	102
5.1.2	Description des sources de lumière	102
5.1.3	Description des objets éclairés	107
5.1.4	Modèle d'éclairage photométrique	108
5.1.5	Optimisation des calculs d'éclairage	110

5.1.6	Résultats	113
5.1.7	Performances	114
5.2	Visualisation réaliste d'images de synthèse HDR	116
5.2.1	Choix d'un algorithme de simulation de l'éblouissement	117
5.2.2	Implémentation de la simulation de l'éblouissement	117
5.2.3	Choix d'un opérateur de reproduction de tons	120
5.2.4	Implémentation de l'opérateur d'Irawan	123
5.2.5	Résultats	126
5.2.6	Performances	127
5.2.7	Conclusion	130
6	Evaluation	133
6.1	Evaluation du modèle HDR	135
6.2	Comparaison de simulations de l'éblouissement	139
6.2.1	Reproduction de l'éblouissement	140
6.2.2	Augmentation de la luminosité perçue	142
6.2.3	Conclusion sur la simulation de l'éblouissement	142
6.3	Evaluer la reproduction de tons	143
6.4	Evaluation de la reproduction de tons par une métrique	144
6.4.1	La saillance comme critère d'évaluation	144
6.4.2	Saillance d'images HDR par [Itti 00]	145
6.4.3	Modification de l'algorithme d'Itti et Koch [Itti 00]	147
6.4.4	Validation de l'algorithme d'Itti et Koch modifié	150
6.4.5	Conclusion	153
6.5	Evaluation de la reproduction de tons par des expérimentations	155
6.5.1	Description des expérimentations	156
6.5.2	Résultats	158
6.6	Conclusion	166
7	Conclusion	169

Table des figures

1.1	Evolution du nombre de morts sur la route, entre 1997 et 2009	14
1.2	Evolution désirée du nombre de morts sur la route, jusqu'en 2012	15
1.3	Rendus d'une autoroute avec et sans éclairage.	19
1.4	Rendu d'une autoroute, éclairée par un phare, de nuit	19
2.1	Spectre de la lumière blanche.	26
2.2	Illustration des relations entre les grandeurs radiométriques.	27
2.3	Illustration des relations entre les grandeurs photométriques.	28
2.4	Efficacités lumineuses pour la vision photopique et scotopique	30
2.5	Courbes $\bar{r}(\lambda)$, $\bar{g}(\lambda)$ et $\bar{b}(\lambda)$	31
2.6	Courbes $\bar{x}(\lambda)$, $\bar{y}(\lambda)$ et $\bar{z}(\lambda)$	32
2.7	Diagramme de chromaticité	34
3.1	Schéma des voies optiques du système visuel humain	39
3.2	Coupe sagittale d'un œil humain	41
3.3	Structure de la rétine	41
3.4	Distribution angulaire des cônes et des bâtonnets sur la rétine	42
3.5	Sensibilité spectrale des pigments visuels des cônes	43
3.6	Dynamique de luminance du monde réel	44
3.7	Seuils de détection des contrastes en fonction de la luminance et de l'état d'adaptation de l'œil	46
3.8	Evolution du seuil de détection des contrastes en fonction de l'adaptation temporelle des photorécepteurs	50
3.9	Variations du seuil de détection des bâtonnets lors d'une adaptation à la lumière	50
3.10	Variations du seuil de détection des cônes lors d'une adaptation à la lumière	51
3.11	Simulation des différents effets de l'éblouissement sur l'œil humain	52

3.12	Réduction du contraste par un éblouissement	53
4.1	Captures d'écran des modes HDR de différents moteurs de rendu . . .	59
4.2	Captures d'écran de différents jeux, avec et sans rendu HDR	59
4.3	L'objectif de la reproduction de tons	61
4.4	Histogramme et histogramme cumulé	69
4.5	Diagramme de l'opérateur de reproduction de tons de Pattanaik et <i>al.</i> .	74
4.6	Réponse des cônes et des bâtonnets en fonction de la luminance et de leur état d'adaptation	75
4.7	Modélisation du blanchiment par la lumière et de la régénération des bâtonnets et des cônes, en fonction du temps.	77
4.8	Courbe de réponse des photorécepteurs	82
4.9	Ajustement d'histogramme par la méthode d'Irawan et <i>al.</i>	85
4.10	Création de la luminance de voile, par Kawase	94
4.11	Image de stries et filtre pour l'éblouissement, d'après Spencer et <i>al.</i> .	95
5.1	Forme d'émission d'un lampadaire	105
5.2	Forme d'émission d'un phare	105
5.3	Texture initiale de phares éteints et image HDR de phares allumés . .	107
5.4	Vecteurs normalisés pour les calculs d'éclairage.	110
5.5	Pipeline du rendu HDR	111
5.6	Images de synthèse HDR	113
5.7	Luminances retenues pour participer à la simulation de l'éblouissement	118
5.8	Simulation de l'éblouissement avec 3 filtres gaussiens	119
5.9	Simulation de l'éblouissement, basée sur la <i>point-spread-function</i> . .	121
5.10	Pipeline pour la reproduction de tons d'Irawan et <i>al.</i>	126
5.11	Captures d'écran du rendu HDR sur route urbaine	127
5.12	Captures d'écran du rendu HDR sur autoroute	128
6.1	Rendu HDR réaliste d'environnements virtuels.	134
6.2	Images HDR réelles et leurs luminances	137
6.3	Images HDR virtuelles et leurs luminances	138
6.4	Comparaison de filtres pour l'éblouissement	140
6.5	Comparaison d'images avec et sans simulation de l'éblouissement . .	141
6.6	Evaluation d'opérateurs de reproduction de tons par la saillance. . . .	145
6.7	Comparaison de cartes de conspécuité	146
6.8	Filtre de Gabor pour la détection de contour	147

6.9	Scène HDR utilisée pour l'expérimentation.	150
6.10	Cartes de fixation des sujets sur la scène HDR.	151
6.11	Cartes de saillance de l'image HDR.	151
6.12	Cartes de saillance des images des opérateurs de reproduction de tons.	152
6.13	Images de l'expérimentation sur des images	157
6.14	Images de l'expérimentation sur des vidéos	159
6.15	Comparaison du comportement d'opérateurs face à une augmentation brusque de la luminance maximale	162
6.16	Comparaison du comportement d'opérateurs face à une baisse brusque des luminances moyenne et maximale	163
6.17	Comportement de l'opérateur de Durand et Dorsey	165

Liste des tableaux

2.1	Correspondance entre grandeurs radiométriques et photométriques . . .	28
3.1	Contribution à la dispersion de la lumière de chaque partie de l'œil . . .	53
4.1	Caractéristiques des principaux opérateurs de reproduction de tons . . .	62
5.1	Performances du rendu HDR en fonction du nombre de sources dif- fuses à l'écran	115
5.2	Performances du rendu HDR en fonction du nombre total de sources diffuses de la scène, en ayant 15 sources à l'écran.	115
5.3	Performances du rendu HDR en fonction de la résolution	116
5.4	Autres caractéristiques des principaux opérateurs de reproduction de tons	122
5.5	Performances de la simulation de l'éblouissement en fonction du nombre de filtres gaussiens utilisés.	128
5.6	Performances de la simulation de l'éblouissement en fonction de la taille des filtres gaussiens utilisés	128
5.7	Performances de l'opérateur d'Irawan et <i>al.</i>	129
5.8	Performances du rendu HDR, en fonction du nombre total de sources dans l'environnement virtuel	130
6.1	Comparaison des luminances entre réel et virtuel	136
6.2	Comparaison de différentes cartes de saillance.	149
6.3	Score moyen des algorithmes de reproduction de tons	160
6.4	Test HSD de Tukey pour l'expérimentation sur des images	161
6.5	Groupes d'appartenance des opérateurs (expérimentation sur des images)	161
6.6	Test HSD de Tukey pour l'expérimentation sur des vidéos	164
6.7	Groupes d'appartenance des opérateurs (expérimentation sur des vidéos)	164
6.8	Groupes d'appartenance des 8 opérateurs sur les 2 expérimentations. . .	165

Chapitre 1

Introduction

1.1 Synthèse d'image : modélisation et rendu

Les premières images de synthèse sont apparues dans les années 60. On ne pouvait alors avoir qu'une image de faible résolution, utilisant seulement quelques niveaux de gris. Mais les progrès en informatique ont permis d'atteindre une qualité photo-réaliste dès les années 90, d'abord pour les images fixes, puis pour les animations en temps différé. Il aura fallu attendre les années 2000 pour voir apparaître les premières animations d'images de synthèse en temps-réel, d'abord dans un mode de rendu assez dégradé, puis aujourd'hui dans un rendu en temps-réel atteignant la qualité photo-réaliste [CAD 07].

La génération d'images de synthèse nécessite deux étapes essentielles. La première est la modélisation de la forme, objet, personne ou scène à afficher (on parlera de *scène* par la suite). Cette modélisation correspond au stockage par l'ordinateur d'un ensemble de données géométriques et de propriétés photométriques décrivant la scène. La seconde étape est l'étape de rendu, qui consiste à illuminer la scène puis à l'afficher à l'écran, en exploitant les données prédéfinies lors de l'étape de modélisation.

En fonction des métiers, les utilisateurs ont des besoins en modélisation et en rendu différents. Par exemple, un designer produit a besoin de générer une modélisation qui simule au plus près ce que sera le produit fabriqué. En revanche, un architecte a davantage besoin de modéliser les aspects caractéristiques d'une construction, puis de la rendre dans son environnement en tenant compte de facteurs temporels et géographiques. Ce qui est commun à ces deux secteurs, contrairement aux domaines de l'audiovisuel et des jeux vidéo, c'est que dans les deux cas, qu'il s'agisse d'un produit ou d'une construction, on ne fait pas de l'image pour l'image mais bien pour simuler

une réalité en devenir. Il y a une énorme différence entre créer un rendu de phénomène physique qui paraît réel (comme pour le cinéma) et simuler un vrai phénomène physique.

Dans cette thèse, nous allons nous attacher à simuler des environnements routiers nocturnes, avec pour objectif que leur perception soit la plus proche possible devant un écran et dans la réalité.

1.2 Sécurité routière et simulation de conduite

1.2.1 La sécurité routière en France

4 262 personnes ont trouvé la mort en 2009, à la suite d'un accident de la route en France métropolitaine, soit en moyenne 12 tués par jour. 83 911 personnes ont été blessées (soit 230 par jour en moyenne), dont on estime approximativement à plus de 4 000 le nombre de blessés avec séquelles majeures : lésions avec perte de substance nerveuse (encéphale, moelle épinière, racines nerveuses), ou perte d'un membre ou destruction de grosses articulations [APR 10].

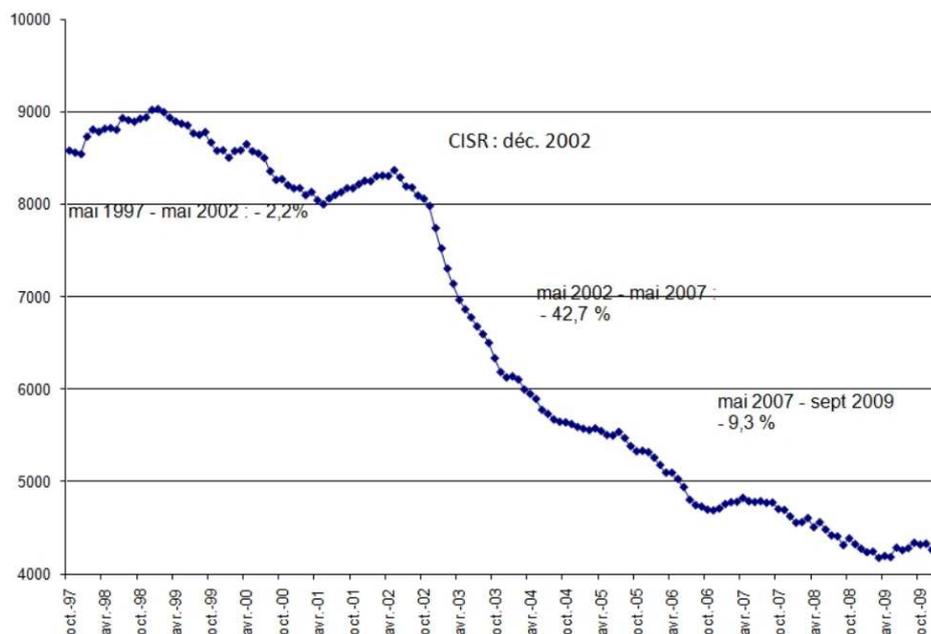


FIGURE 1.1 – Evolution du nombre de morts sur la route en France, entre 1997 et 2009. D'après [Got 10].



FIGURE 1.2 – Comparaison entre l'évolution du nombre de morts sur la route et l'objectif à atteindre, entre 2006 et 2012 . D'après [Got 10].

La situation s'est pourtant améliorée en France depuis 1998, où l'on dénombrait environ 9000 morts par an sur les routes (Figure 1.1). L'an 2000 a été un tournant pour la sécurité routière : elle est devenue une « Grande Cause Nationale », dans le but de favoriser une prise de conscience des accidents de la route ainsi qu'une large mobilisation (on se souvient du slogan : « Si chacun fait un peu, c'est la vie qui gagne. »).

Avec notamment un accroissement des contrôles par la mise en place du contrôle automatisé et l'aggravation des sanctions, les résultats pour la sécurité routière ont été exceptionnels : le nombre de morts a été divisé par deux entre 1998 et 2007 [Got 10].

Alors récemment élu, le nouveau président de la République fixa pour objectif que moins de 3000 personnes ne meurent par an sur les routes d'ici 2012. Le second graphique (Figure 1.2) détaille l'évolution au cours des trois dernières années et projette jusqu'en 2012 l'évolution régulière qui permettrait d'atteindre l'objectif de 3000 tués fixé par le président de la République. On constate qu'il y a un net décrochage entre les deux courbes.

Début 2010, le gouvernement a adopté les consignes suivantes pour continuer à faire baisser le nombre de tués sur la route [CISR 10] :

- combattre l'alcool au volant (plus d'éthylotests et plus de contrôles) ;
- dissuader les excès de vitesse (avec des radars) ;

- sanctionner plus sévèrement les grands délits routiers ;
- sensibiliser les futurs conducteurs (prévention dans les lycées) ;
- prévenir le risque professionnel.

Pour éviter les accidents de la route, le gouvernement agit sur deux leviers : la prévention et la répression. Les trois premières mesures du gouvernement prônent le développement des contrôles, des radars et l'aggravation des sanctions : ce sont des mesures répressives. La prévention concerne les deux dernières mesures, pour les futurs conducteurs et les agents qui sont amenés à conduire dans l'exercice de leur profession.

1.2.2 La simulation de conduite, un outil pour la prévention

Les simulateurs de conduite sont des simulateurs de véhicules terrestres (véhicules légers, poids lourds, motos, vélos... mais aussi trains et navires), dans lequel un conducteur réel conduit un véhicule virtuel [Wikipedia 10b]. Nous nous intéressons ici aux véhicules circulant sur route terrestre. L'intérêt du simulateur de conduite est de placer le conducteur dans un environnement comparable à celui d'une conduite réelle, mais où l'environnement et le véhicule sont totalement sous contrôle car ils sont représentés par des modèles logiciels.

En dehors des jeux vidéo, les simulateurs de conduite connaissent actuellement trois types d'utilisations, chacune contribuant à l'amélioration de la prévention pour la sécurité routière :

- les recherches et développements liés à la mise au point de nouveaux véhicules, de systèmes d'aide à la conduite ou de nouvelles infrastructures ;
- les études et recherches sur le comportement du conducteur et les facteurs humains liés à la conduite, comme la vigilance, l'attention et la perception ;
- la formation des conducteurs, tels que les professionnels de la conduite, ou les conducteurs apprenant la conduite économique ou éco-conduite. Les simulateurs sont également utilisés à l'école pour sensibiliser des futurs motards à la conduite des deux roues, notamment à travers l'analyse d'accidents survenus en simulation de conduite [MDEVO 10].

La simulation de conduite offre une alternative et une complémentarité aux recherches et à la formation à bord de véhicules réels. Les avantages du simulateur de conduite sont nombreux, à commencer par des aspects sécuritaires (possibilité d'étudier des situations dangereuses avec des sujets en toute sécurité), économiques (coût réduit de la mise en œuvre), organisationnels (facilité de mise en situation, possibilité

de reproduire exactement des scénarios), et environnementaux. Cependant, les simulateurs actuels ne permettent pas une représentation exacte de la réalité, et peuvent induire des biais dans l'apprentissage ou dans le comportement du conducteur par rapport à la situation réelle. Pour des performances optimales, il convient donc d'utiliser un simulateur le plus fidèle possible par rapport à la réalité.

1.2.3 Le rendu visuel au sein des simulateurs

Il existe de nombreux simulateurs de conduite de par le monde. Ce sont des outils complexes, qui nécessitent beaucoup d'investissements pour faire ressentir au conducteur l'impression d'être dans la réalité. La simulation des modèles de véhicules, de trafic, mais aussi la reproduction des mouvements de la caisse lors de la circulation en sont des composants importants, même s'ils ne sont pas toujours tous présents.

Comme pour la réalisation d'images de synthèse (cf section 1.1), l'affichage de scènes routières nécessite deux étapes. La première, de modélisation, consiste à créer la géométrie et l'aspect des routes, bâtiments et véhicules que l'on rencontre au cours de la simulation. C'est un travail d'infographiste, et c'est d'abord dans cette étape que les investissements pour le visuel d'un simulateur sont réalisés. La seconde étape, celle du rendu, consiste à éclairer la scène et à l'afficher à l'écran. Mais elle est souvent négligée, pour des raisons de coûts, de temps de développement, de savoir faire, de temps réel... ou tout simplement de besoin, car tous les simulateurs de conduite n'ont pas forcément pour objectif de reproduire le plus fidèlement possible à l'écran la réalité. Il existe cependant bon nombre d'applications où l'on cherche à avoir un rendu à l'écran aussi proche que possible de la réalité. Les études sur la perception de l'environnement routier ou de véhicules en circulation ont ce type de besoin sur le rendu visuel (Brémond, Dumont 2002), comme l'illustre la section suivante.

1.2.4 La question de l'éclairage

Cette section présente une étude réalisée sur l'impact de l'éclairage du réseau autoroutier, de nuit. Elle met en avant, dans ce cas précis, le besoin de réalisme visuel exact, ou tout du moins le plus proche possible de la réalité lors de la simulation.

La circulation nocturne ne représente qu'un cinquième du trafic total du réseau routier et autoroutier français, pourtant il y a autant de tués sur la route le jour que la nuit. En rase campagne comme en agglomération, l'insécurité routière est plus grande la nuit que le jour [Remande 09].

Un long débat - qui n'est pas terminé aujourd'hui - traite du rôle de l'éclairage public sur la sécurité des usagers. En France, deux parties s'opposent. D'une part, l'Association Française de l'Eclairage (AFE) et ses partenaires (Electricité De France et le Syndicat de l'Eclairage) avancent que l'éclairage améliore la visibilité, et ainsi favorise l'anticipation, sans pour autant que l'éclairage ne provoque une augmentation de la vitesse du conducteur qui pourrait être liée au sentiment de sécurité.

D'autre part, le gouvernement soutient qu'il n'a jamais été démontré que l'éclairage améliore la sécurité, car la plupart des études statistiques affirment le contraire ([Wanvik 10, Mayeur 10]). Elle craint également que l'éclairage ne donne un sentiment de sécurité au conducteur, qui peut alors être tenté d'augmenter sa vitesse et ainsi prendre plus de risques. Enfin, son dernier argument est d'ordre économique, car l'éclairage coûte cher ! (1 % de la consommation électrique française).

Cette interrogation sur l'influence éventuelle de l'éclairage public sur la réduction des accidents a poussé l'AFE et ses partenaires à financer une étude scientifique confiée au CNRS et au Centre d'Etude de Physiologie Appliquée de Strasbourg.

Une expérimentation était donc nécessaire. Mais une expérimentation sur site réel posait certains problèmes pour les raisons suivantes (voir toutefois [Koblentz 96]) :

- des situations accidentogènes sont provoquées lors de l'expérimentation, et peuvent mettre en danger les sujets ;
- les situations ne sont pas reproductibles d'un sujet à l'autre ;
- la différenciation des paramètres accidentogènes est difficile.

Une expérimentation sur simulateur de conduite a donc été choisie, car :

- elle ne présente aucun risque pour les sujets ;
- les situations sont reproductibles ;
- on peut faire varier un unique paramètre accidentogène à la fois ;
- les expérimentations peuvent être poussées jusqu'à l'accident.

La question de la validité du simulateur s'est ensuite posée. C'est une question importante pour tout type d'étude ou de recherche sur le comportement des conducteurs, et/ou les facteurs humains liés à la conduite (vigilance, attention, perception). Pour cette expérimentation, une évaluation subjective du simulateur utilisé a été demandée aux sujets, qui ont noté le « réalisme » du simulateur. Cependant, au vu des images présentées sur la figure 1.3, il y a une nette différence entre les images affichées et la réalité.

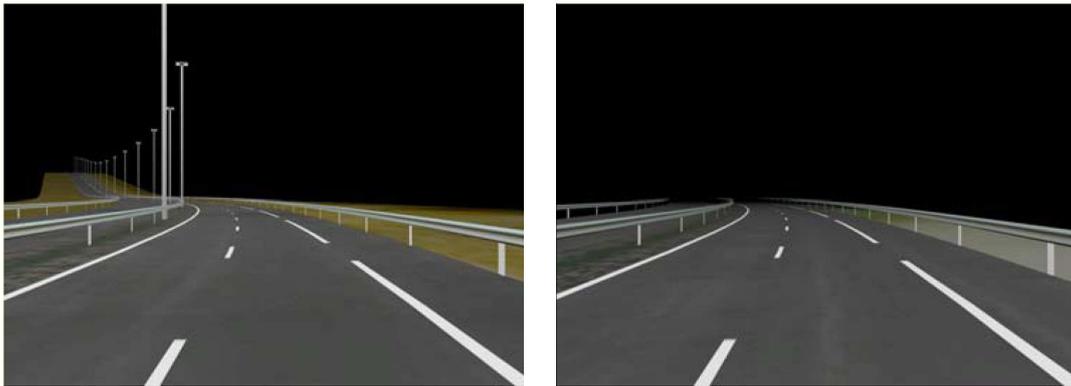


FIGURE 1.3 – Rendus d’une autoroute avec et sans éclairage.

Sans remettre en cause l’expérience qui a été menée, nous pouvons nous demander si un simulateur aux images plus fidèles à la réalité n’aurait pas réduit le biais entre la simulation et la réalité (section 1.2.2), et ainsi améliorer la fiabilité des résultats. En intégrant par exemple un éclairage dynamique de la scène, l’atmosphère nocturne aurait été mieux reproduite (alors qu’ici on se contente d’obscurcir les surfaces et objets éloignés de la voiture), et l’éclairage de la chaussée n’aurait pas été uniforme, puisqu’elle dépend de la position des lampadaires et des phares. Enfin, avec ce type de rendu, les contrastes affichés à l’écran ne correspondent pas à ceux qui sont perçus dans la réalité.

Ces différents points négatifs ont d’ailleurs fait douter l’administration des résultats de cette expérimentation, si bien que ses résultats n’ont pas été exploités par la suite. En 2010, plusieurs autoroutes sont privées d’éclairage pour raison économique, et seules les zones d’urbanisation dense, les successions de tunnels, les zones d’échangeurs et d’intersections resteront éclairées.

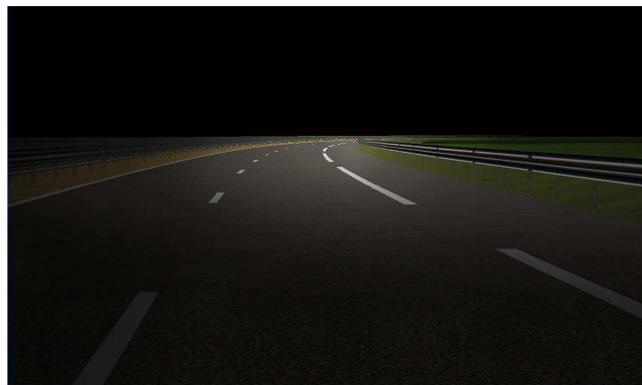


FIGURE 1.4 – Rendu d’une autoroute, éclairée par un phare, de nuit, produite par le simulateur du LEPSiS.

A titre de comparaison, la figure 1.4 présente une images de route nocturne produite par le simulateur du LEPSiS.

1.3 Besoins identifiés

Nous avons montré dans cette introduction qu'un simulateur de conduite constitue un outil technologique irremplaçable, que ce soit pour des recherches scientifiques pour la sécurité routière ou des formations à la conduite. Au LEPSiS, les simulateurs de conduite ont un double intérêt. Ils sont d'abord des sujets de recherche, car leurs utilisateurs sont toujours à la recherche d'un meilleur réalisme, et pas seulement au niveau du rendu visuel. Mais ils sont également des outils pertinents pour participer à l'amélioration de la sécurité routière, car ils permettent d'étudier le comportement d'automobilistes en situation contrôlée, et notamment leurs actions et réactions face à l'environnement routier.

Cependant cet outil reste actuellement limité. En particulier, dans le domaine du rendu visuel, des contraintes techniques limitent le domaine de validité des simulateurs de conduite. Cela a été illustré par l'étude du comportement des conducteurs de nuit avec et sans éclairage. Pour qu'une étude soit valide sur simulateur de conduite, il est nécessaire que la perception des objets routiers (panneaux, aménagements, autres usagers) soit comparable sur simulateur et dans l'environnement de conduite réel.

A l'heure actuelle, les simulateurs de conduite souffrent de limitations, notamment pour rendre les scènes nocturnes ainsi que les sources lumineuses, y compris de jour (feux stop, feux tricolores, etc.). Il y a là une piste de progrès pour avoir des données fiables, par exemple sur l'impact des codes de jour sur la sécurité des deux roues et des piétons, ou de nuit, sur l'impact de solutions d'aménagement et d'éclairage public.

Les problèmes que l'on se pose dans cette thèse sont les suivants : avec des images de synthèse, et avec un écran standard (LDR), comment afficher le plus fidèlement possible ce que l'on voit dans la réalité ? Et une fois un résultat obtenu, comment vérifier si l'observateur perçoit réellement la même chose à l'écran que dans la réalité ?

L'approche proposée au cours de cette thèse est de créer un mode de rendu à grande dynamique de luminance, en anglais *High Dynamic Range* (mais l'acronyme HDR sera utilisé par la suite). Ce rendu permet de générer une image de la scène en luminance, et de la traiter avec un opérateur de reproduction de tons (ou *tone mapping*). Cette approche vise à apporter un maximum de réalisme et de ressemblance à l'écran par rapport à la réalité.

1.4 Objectifs de la thèse

La thèse s'inscrit dans le cadre du développement du rendu visuel d'un simulateur de conduite automobile. Son objectif est d'améliorer la pertinence des différentes études qui sont réalisées sur simulateur, comme l'étude du comportement du conducteur en situation de conduite normale, de ses réactions face à des situations particulières (nuit, météo dégradée, fatigue du conducteur, ...), ou encore l'étude de sa perception de l'environnement.

Pour améliorer la qualité du rendu visuel des simulateurs, ainsi que la qualité des études scientifiques produites, les objectifs suivants ont été listés :

Générer des images de synthèse HDR

Puisque l'approche proposée est de créer un mode de rendu HDR, le premier objectif est bien sûr de pouvoir générer des images HDR, en temps réel. Nous devons pour cela développer un modèle d'éclairage, qui permettra de calculer (ou, tout du moins, d'approximer) la luminance en chaque point de l'image. Ce sont donc bien des images en unités physiques qui seront produites.

Déterminer et implémenter un opérateur de reproduction de tons adapté à la simulation de conduite

Les écrans ont des caractéristiques techniques trop limitées (en termes de luminosité et de contraste, voir le chapitre 2) pour bien reproduire à l'écran la réalité telle que nous la percevons. Des algorithmes de traitement d'images visent néanmoins à compenser ces limitations, en simulant la vision humaine et en adaptant l'image en fonction des caractéristiques de l'écran, pour tenter de reproduire une perception exacte de l'image à l'écran. Ces algorithmes s'appellent des opérateurs de reproduction de tons. A partir des images HDR générées, nous utiliserons donc un de ces algorithmes, en faisant attention aux points suivants :

- nous essayerons de garantir que les contrastes à l'écran soient les mêmes que ceux de la réalité. Dans certaines études comme les études de saillance, il est en effet important qu'un objet visible dans la réalité le soit également à l'écran, et inversement, qu'un objet qui soit pas visible dans la réalité ne le soit pas non plus à l'écran.
- nous simulerons également l'adaptation temporelle de l'œil à la lumière et à l'obscurité. Il existe des cas difficiles pour la conduite, comme par exemple l'en-

trée et la sortie de tunnels, où nos yeux sont soumis à de forts changements de luminosité, et il est important, pour le réalisme, de reproduire cette adaptation temporelle à l'écran.

Utiliser des sources pour la simulation de l'éclairage

Pour des raisons de simplicité, certains simulateurs n'utilisent aucune source de lumière dans leurs simulations. Pour des simulations de jour, ce n'est pas forcément gênant si l'on considère que tous les objets sont éclairés uniformément, mais de nuit, le rendu ne pourra pas être réaliste sans source(s) de lumière. Un objectif est d'utiliser des sources lumineuses (qui peuvent être nombreuses), et de pouvoir régler leur intensité avec des paramètres physiques exacts, que l'on peut par exemple trouver dans des documentations de constructeurs [Philips 99] ou mesurer. La méthode employée devra permettre de simuler des ombres dynamiques, dues par exemple aux mouvements d'une voiture.

Définir une chaîne de production des bases pour un rendu HDR

Nous avons vu que le rendu visuel des simulateurs de conduite nécessitait deux étapes qui doivent être intimement liées pour un résultat optimal : une première étape de modélisation, une seconde de rendu (section 1.2.3). Nous verrons dans le chapitre 5 qu'un rendu HDR nécessite de paramétrer les sources de lumière et les objets de la scène pour que le calcul d'éclairage puisse se faire en unités physiques. Nous décrivons donc la chaîne de production réalisée, pour permettre un gain de temps dans la production et l'utilisation des environnements virtuels dans un mode HDR. La chaîne de production présentera comment doivent être définies les environnements virtuels, et renseignera les infographistes sur les paramètres à définir et les conventions à adopter.

Evaluer le rendu HDR obtenu

Pour qu'une simulation puisse être considérée comme réaliste dans une étude scientifique, il convient de la valider par des expérimentations objectives et/ou subjectives. Il faudra donc proposer une évaluation expérimentale à partir de photos et de vidéos pour valider les résultats obtenus lors du rendu pour la simulation de conduite. Le critère d'évaluation sera la fidélité par rapport à la réalité.

1.5 Organisation du manuscrit

Le chapitre 2, **Lumière et Couleur**, définira les termes photométriques nécessaires à la compréhension du manuscrit. Il introduira les différents dispositifs d'affichage qui seront rencontrés, ainsi que leurs technologies.

Le chapitre 3, **Vision**, présentera les caractéristiques physiologiques et fonctionnelles du système visuel humain. Il se limitera aux aspects de la vision nécessaires à la tâche de conduite.

Le chapitre 4, **Etat de l'art sur le rendu visuel**, sera consacré à un état de l'art sur le rendu visuel. Il présentera d'abord différents rendus HDR de la littérature. Suite logique du rendu HDR, une bibliographie sur les opérateurs de reproduction de tons, sur leurs évaluations, et sur la simulation de l'éblouissement suivra. Enfin, une dernière partie sera consacrée à la saillance d'images et de vidéo.

Le chapitre 5, **Rendu HDR**, décrira la méthode mise au point pour générer un flux d'images de synthèse HDR. Il détaillera d'abord le modèle photométrique employé, puis le choix et la mise en œuvre de l'opérateur de reproduction de tons adéquat pour la simulation de conduite.

Le chapitre 6, **Evaluation**, proposera dans un premier temps des évaluations du rendu d'images HDR et de la simulation de l'éblouissement. Dans un second temps, il abordera l'évaluation de la reproduction de tons, par deux méthodes distinctes. La première, par une métrique, sera une modification de l'algorithme sur la saillance d'Itti [Itti 98] pour pouvoir l'étendre aux images HDR. Dans un second temps, le chapitre abordera les évaluations subjectives effectuées sur la reproduction de tons d'images HDR nocturnes.

Nous concluons dans le chapitre 7, **Conclusion et perspectives**, avec un bilan de nos contributions et les perspectives possibles de nos travaux.

Chapitre 2

Lumière et couleur

Dans cette thèse, nous nous intéressons au rendu et à la reproduction sur écran d'images de synthèse réalistes. Pour développer et utiliser des algorithmes performants, il nous faut connaître les caractéristiques de la lumière, des couleurs et des dispositifs d'affichage. C'est l'objet de ce chapitre, dont les données principales proviennent des sources [Grave 06, Reinhard 05b, Couleur 10, Matković 97].

2.1 Lumière

La lumière se propage en ligne droite, à la vitesse d'environ $c = 3.10^8 m.s^{-1}$ dans le vide. Selon ses comportements ondulatoire et corpusculaire, elle est composée de photons, chaque photon étant associé à une onde électromagnétique.

Chaque onde électromagnétique qui compose la lumière est caractérisée par son énergie et sa longueur d'onde λ (ou sa fréquence ν , telle que $\lambda = c/\nu$). L'ensemble des longueurs d'ondes comprises dans un rayon de lumière composent son spectre de rayonnement. Enfin, l'œil humain n'est capable de percevoir qu'une faible portion de l'ensemble des longueurs d'ondes : c'est le spectre visible, avec des longueurs d'ondes comprises entre 360 nm et 830 nm (figure 2.1).

2.2 Radiométrie

La radiométrie consiste à mesurer le flux d'énergie associé aux rayonnements électromagnétiques. Les principales grandeurs radiométriques utilisées sont les suivantes (figure 2.2) :

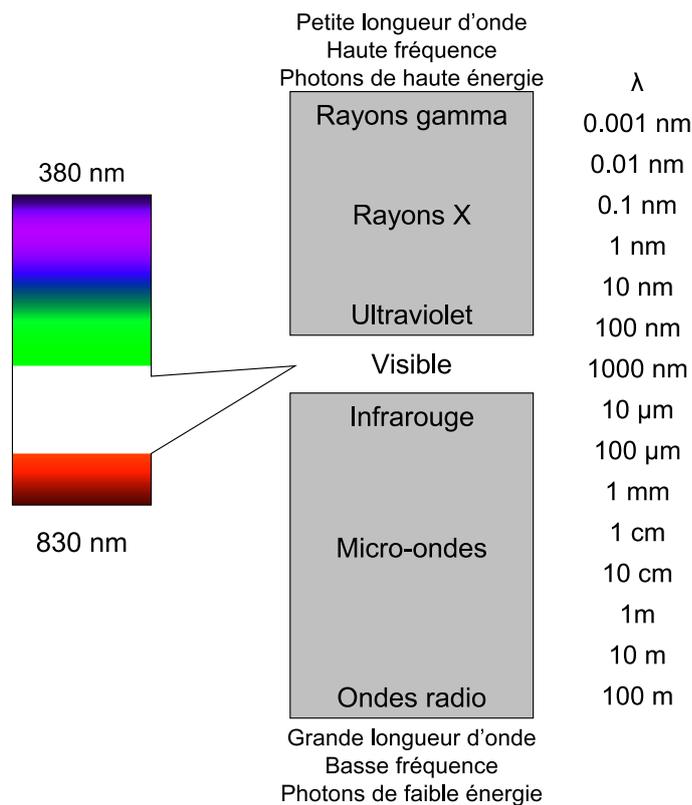


FIGURE 2.1 – Spectre de la lumière blanche.

- le **flux énergétique** ϕ_e , qui est l'énergie rayonnée ou reçue par unité de temps ;
- l'**intensité énergétique** I_e , qui mesure le flux émis par une source ponctuelle, dans une direction donnée et par unité d'angle solide ;
- l'**exitance énergétique** M_e , qui quantifie la puissance rayonnée par une source, par unité de surface et dans un hémisphère ;
- l'**éclairement énergétique** E_e , qui mesure le flux énergétique par unité de surface, quelle que soit la direction d'origine du flux.
- la **luminance énergétique** L_e , ou radiance, qui est le flux énergétique émis par unité de surface de la source et par unité d'angle solide, dans une direction donnée.

Pour les rayonnements polychromatiques (i.e. composés de plusieurs longueurs d'ondes), on considère la répartition spectrale de ces grandeurs par rapport à chaque longueur d'onde λ (par exemple, le flux énergétique spectral sera $\phi_{e,\lambda}$).

- Pour décrire la surface d'un objet, les grandeurs suivantes sont également utilisées :
- la **réflectance**, qui est le rapport de l'exitance énergétique à l'éclairement éner-

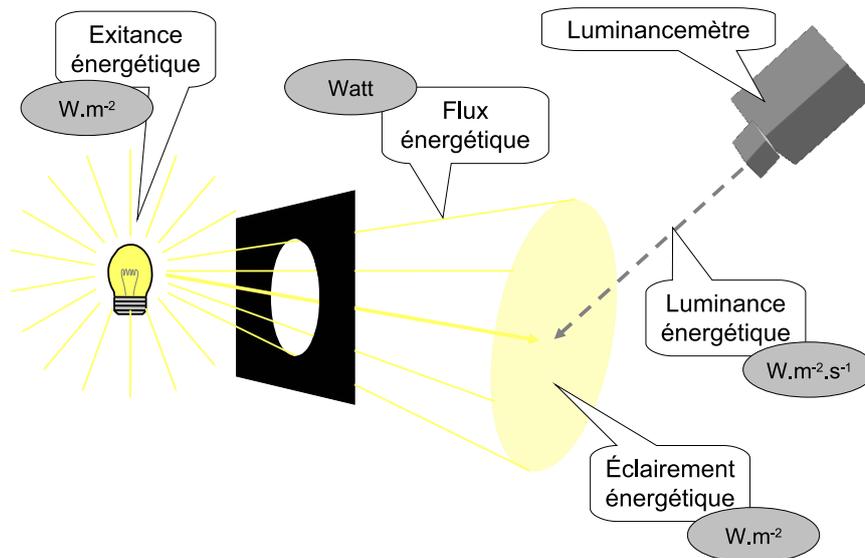


FIGURE 2.2 – Illustration des relations entre les grandeurs radiométriques.

gétique ;

- l'**albédo**, qui est le rapport entre le rayonnement solaire réfléchi et le rayonnement solaire incident. L'albédo est donc équivalent à la réflectance, si la source utilisée est le soleil.

Selon les milieux, la valeur de l'albédo est comprise entre 0 (pour un milieu totalement absorbant dans toutes les longueurs d'ondes) et 1 (pour un milieu parfaitement réfléchissant). Pour les milieux naturels, les valeurs se situent le plus souvent entre 0,3 et 0,7.

2.3 Photométrie

La photométrie est l'évaluation des grandeurs caractéristiques d'un rayonnement, indépendamment de sa longueur d'onde. Quand l'appareil récepteur est le système visuel humain, on parle de photométrie visuelle.

2.3.1 Grandeurs photométriques

La photométrie propose un nouvel ensemble d'unités basées sur la stimulation visuelle, qui viennent compléter les unités radiométriques basées sur l'énergie (figure

2.3). Chaque unité photométrique trouve une correspondance avec une unité radiométrique (tableau 2.1). Les définitions des termes sont donc identiques, si ce n'est qu'elles se rapportent au spectre visible.

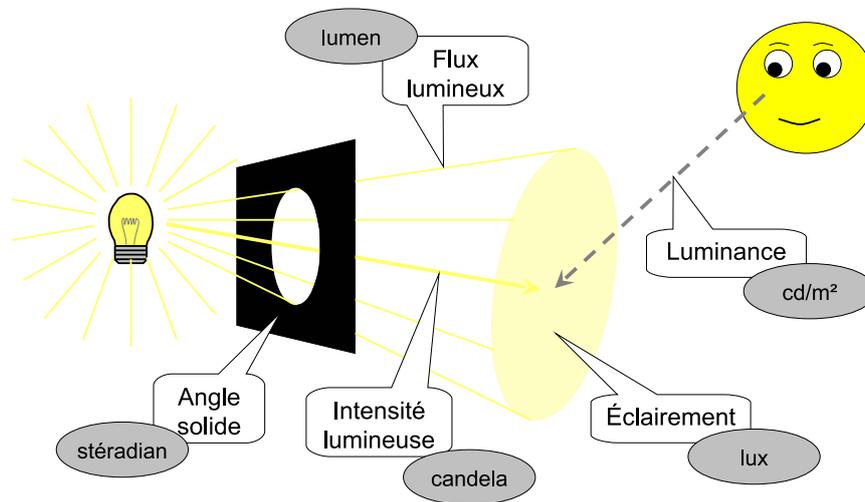


FIGURE 2.3 – Illustration des relations entre les grandeurs photométriques.

Grandeurs radiométriques spectrales	Unités	Grandeurs photométriques	Unités
Flux énergétique	$Watt (W)$	Flux	lumen (lm)
Intensité énergétique	$W.sr^{-1}$	Intensité	candéla (cd)
Exitance énergétique	$W.m^{-2}$	Exitance	$lm.m^{-2}$
Luminance énergétique	$W.sr^{-1}.m^{-2}$	Luminance	$cd.m^{-2}$
Eclairement énergétique	$W.m^{-2}$	Eclairement	lux (lx)

TABLE 2.1 – Correspondance entre grandeurs radiométriques et photométriques

Luminance

Parmi ces grandeurs photométriques, la luminance nous intéresse plus particulièrement. Elle décrit la sensation visuelle qu'on a de la lumière au niveau de l'œil lorsqu'on regarde la surface d'un objet. La gêne ressentie dans une forte pénombre, ou lorsqu'un phare nous éblouit, est due au fait que notre œil n'est pas adapté aux luminances qu'il

voit. Dans certains cas, l'œil peut s'adapter à ces luminances grâce à un processus d'adaptation.

Plusieurs facteurs influencent le niveau de luminance. En premier lieu, c'est bien sûr l'intensité lumineuse mais aussi la couleur de la source de lumière. Ensuite, lorsque la lumière est renvoyée par une surface éclairée, la couleur et la réflectance interviennent également sur le niveau de luminance.

Luminance et imagerie HDR

Quand une photographie est prise, l'obturateur de l'appareil photo s'ouvre pendant quelques instants. La lumière entrant dans l'objectif est concentrée avec une lentille, qui limite le nombre de directions d'où peut provenir la lumière, et est ensuite enregistrée par le capteur de l'appareil photo. Ce capteur est partitionné en pixels de petite taille, chacun enregistrant la lumière qu'il reçoit. Il est ensuite possible de relier la tension extraite du capteur de l'appareil photo à la luminance reçue, et ainsi obtenir une image de la scène en luminance, c'est-à-dire une image High Dynamic Range. [Debevec 97]. La luminance est donc la grandeur photométrique qui intervient pour l'acquisition d'images HDR.

Luminosité

La luminosité est la perception visuelle de la lumière. Elle n'est donc pas une grandeur photométrique mesurable, puisqu'elle correspond à l'interprétation d'un rayonnement lumineux par l'œil humain (contrairement à la luminance, qui est l'intensité lumineuse réelle et mesurable d'une source de lumière). La luminosité est donc une grandeur subjective, fonction des conditions d'observation et de l'état d'adaptation de l'œil de l'observateur. Enfin, la luminosité n'est pas proportionnelle à la luminance, mais à son logarithme.

Contraste

Le contraste est une grandeur sans unité, qui permet la comparaison entre les luminances les plus fortes et les plus faibles. Il répond à trois définitions distinctes :

- contraste d'un caractère sur un fond (CIE 95) : $C_O = L_{blanc}/L_{noir}$;
- contraste psychométrique : $C_P = (L_{blanc} - L_{noir})/L_{noir}$;
- contraste de Michelson : $C_M = (L_{blanc} - L_{noir})/(L_{blanc} + L_{noir})$;

Pour un dispositif d'affichage, le taux de contraste utilise la première définition : c'est le rapport de luminance entre un pixel blanc et un pixel noir. Il s'exprime sous la forme d'une fraction. Par exemple, un taux de contraste de 500 : 1 signifie que le blanc de l'écran est 500 fois plus lumineux que le noir de l'écran (voir section 2.7.2).

2.3.2 Efficacité lumineuse

L'œil humain ne réagit pas de la même manière à chaque longueur d'onde du spectre visible. Sa réponse relative à chaque rayonnement monochromatique est donnée par la fonction d'efficacité lumineuse. La Commission Internationale de l'Eclairage (CIE) a défini, pour un observateur de référence, des courbes de réponse pour les domaines photopique (de jour) et scotopique (de nuit, figure 2.4).

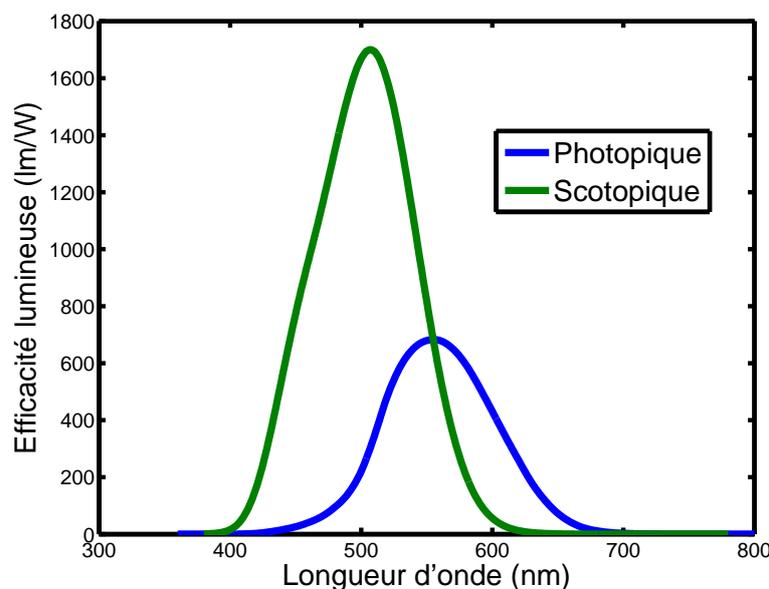


FIGURE 2.4 – Efficacités lumineuses pour un observateur de référence pour la vision photopique et scotopique, en fonction de la longueur d'onde.

Les longueurs d'onde où l'efficacité lumineuse spectrale est maximum sont 507nm pour la vision scotopique, et 555nm pour la vision photopique.

La conséquence de ces courbes est que l'effet de la lumière sur l'œil humain dépend de son spectre. A intensités égales, une lumière jaune (entre 565 et 590nm) sera ainsi toujours perçue comme plus lumineuse qu'une lumière rouge (entre 630 et 780nm).

2.4 Colorimétrie

La colorimétrie est la science de la mesure des couleurs. Bien que chaque personne perçoive les couleurs légèrement différemment, la CIE a également défini un observateur de référence pour la perception des couleurs. Pour cela, elle a procédé à une expérimentation, où les sujets devaient reproduire une couleur donnée sur un écran en réglant sur les intensités des trois couleurs primaires (rouge, vert et bleu). Chaque intensité pouvait être comprise entre -1 et 1 (les intensités négatives correspondaient à l'ajout d'une couleur primaire à la couleur donnée initialement). La figure 2.5 montre les courbes de correspondance obtenues après plusieurs expérimentations sur différents sujets. La longueur d'onde de la couleur donnée initialement est en abscisse, alors que les intensités qu'il faut donner aux trois couleurs primaires pour obtenir une correspondance sont en ordonnée. La couleur donnée initiale C_λ en fonction des intensités $\bar{r}(\lambda)$, $\bar{g}(\lambda)$ et $\bar{b}(\lambda)$ des trois couleurs primaires R (red), G (green) et B (blue) peut donc être obtenue par la formule suivante :

$$C_\lambda = \bar{r}(\lambda)R + \bar{g}(\lambda)G + \bar{b}(\lambda)B \quad (2.1)$$

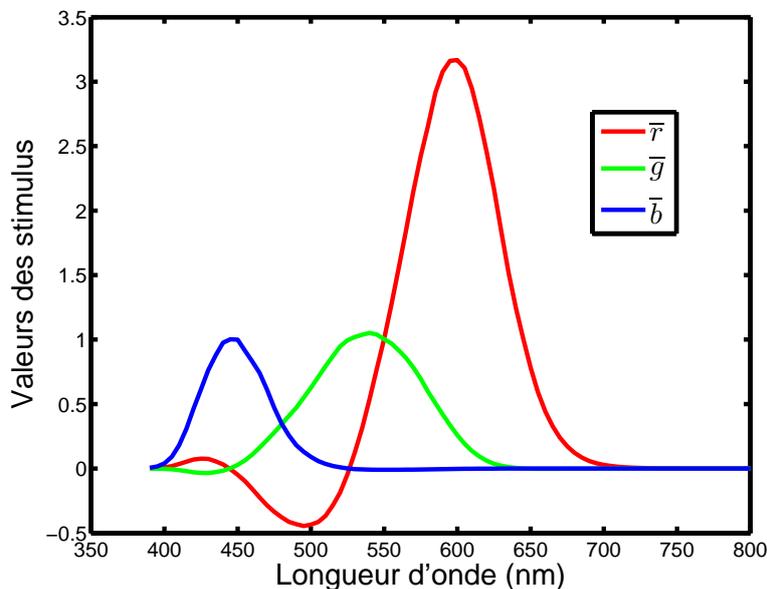


FIGURE 2.5 – Courbes $\bar{r}(\lambda)$, $\bar{g}(\lambda)$ et $\bar{b}(\lambda)$ de correspondance, Stiles and Burch 1959. (Données téléchargées sur <http://www-cvrl.ucsd.edu/>).

De plus, comme il n'est pas pratique d'utiliser un espace de couleur où des valeurs sont parfois négatives, la CIE a réalisé une correspondance où chaque couleur aurait

forcément un coefficient positif. Les fonctions obtenues sont nommées $\bar{x}(\lambda)$, $\bar{y}(\lambda)$ et $\bar{z}(\lambda)$ et sont présentées sur la figure 2.6.

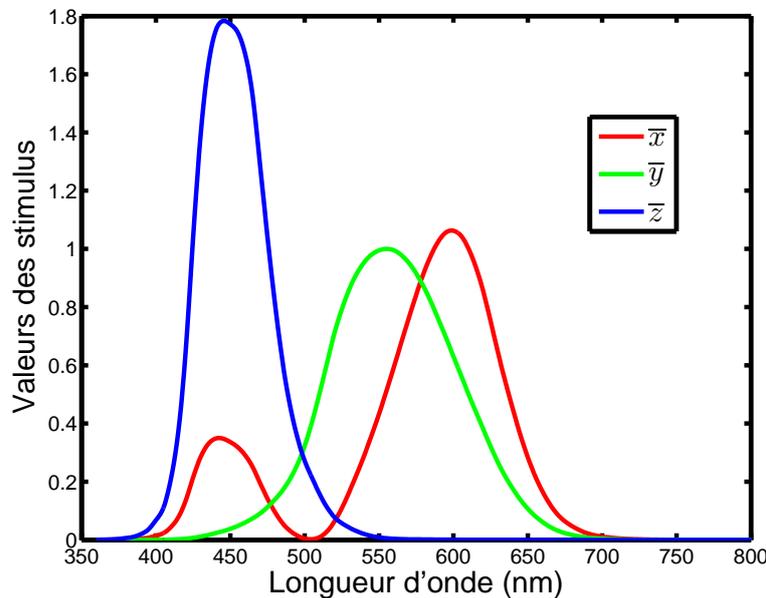


FIGURE 2.6 – Courbes $\bar{x}(\lambda)$, $\bar{y}(\lambda)$ et $\bar{z}(\lambda)$ de correspondance, CIE 1931. (Données téléchargées sur <http://www-cvrl.ucsd.edu/>).

Avec ce nouvel espace, n'importe quelle couleur donnée peut être retrouvée selon la formule suivante, X , Y et Z définissant alors un nouvel espace de couleur :

$$C_\lambda = \bar{x}(\lambda)X + \bar{y}(\lambda)Y + \bar{z}(\lambda)Z \quad (2.2)$$

2.5 Espaces de couleur

Les espaces de couleur sont des repères qui associent des nombres aux couleurs visibles. Ces nombres sont donnés sous la forme de triplets. Les espaces RGB , XYZ vus dans la section précédente sont donc des espaces de couleurs, ce sont même des espaces standards de la CIE. Il existe également de nombreux autres espaces de couleurs (voir ci-après).

Chaque triplet représente une couleur dans un espace, et peut être converti dans un autre espace à l'aide d'une matrice de passage, ou d'une transformation non-linéaire (voir [Lindbloom 09] pour l'ensemble des transformations). Par exemple, la formule suivante permet de passer de l'espace de couleur $sRGB$ (espace représentant les couleurs reproductibles sur les écrans grands publics) à l'espace XYZ :

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} 0.4124 & 0.3576 & 0.1805 \\ 0.2126 & 0.7152 & 0.0722 \\ 0.0193 & 0.1192 & 0.9505 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix} \quad (2.3)$$

En inversant la matrice, on obtient la formule qui permet de faire l'opération inverse :

$$\begin{pmatrix} R \\ G \\ B \end{pmatrix} = \begin{pmatrix} 3.2406 & -1.5372 & -0.4986 \\ -0.9689 & 1.8758 & 0.0415 \\ 0.0557 & -0.2040 & 1.0570 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \quad (2.4)$$

A l'aide de ces matrices de passage, chaque couleur définie dans un espace de couleur donné peut donc être convertie dans n'importe quel autre espace de couleur.

Chaque espace de couleur est généralement défini pour un usage particulier, et possède une gamme de couleurs qui lui est propre. Les espaces suivants sont parmi les plus connus :

- l'espace RVB (rouge-vert-bleu), permettant de décrire les couleurs d'un écran ;
- l'espace CIE RVB, permettant de décrire l'ensemble des couleurs que peut percevoir l'œil humain (chaque couleur étant relié à une longueur d'onde). Certaines couleurs sont décrites par des valeurs négatives pour le rouge ;
- l'espace CIE XYZ, tiré de l'espace CIE RVB, mais qui permet de décrire l'ensemble des couleurs avec des valeurs positives ;
- l'espace xyY, dont Y décrit la luminance alors que le couple x et y définissent la chrominance. Son défaut est que dans le diagramme de chromaticité, les distances entre deux points voisins représentent mal les différences de perception par le cerveau humain. Les espaces suivants corrigent ce défaut.
- le système YUV, utilisé pour la vidéo, avec également un canal pour la luminance et deux pour la chrominance ;
- l'espace CIE L*a*b*, qui se déduit du système xyY par des relations non-linéaires, et qui se rapproche de la perception de l'œil. Ainsi, L est approximativement la racine carrée de la luminance.

D'autre part, les espaces de couleurs suivants sont également proposés par les constructeurs d'écrans et d'appareil photos :

- l'espace sRGB, proposé par Hewlett-Packard et Microsoft pour les écrans d'ordinateurs à tube cathodique ;
- l'espace AdobeRGB, conçu par Adobe Systems Inc. pour améliorer les performances des imprimantes.

2.6 Diagramme de chromaticité

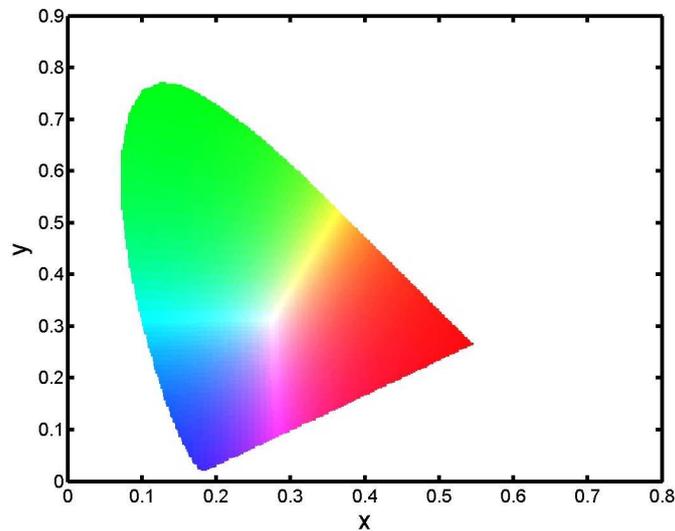


FIGURE 2.7 – Diagramme de chromaticité montrant la gamme de couleurs qu'un observateur standard peut percevoir.

Le diagramme de chromaticité permet une représentation en deux dimensions des couleurs visible par l'œil humain (figure 2.7). La chromaticité d'une couleur peut être calculés à partir des valeurs XYZ , en utilisant les formules suivantes :

$$x = \frac{X}{X + Y + Z} \quad (2.5)$$

$$y = \frac{Y}{X + Y + Z} \quad (2.6)$$

Il existe une composante z mais elle est redondante, puisque la somme des composantes x , y et z est égale à 1. C'est pourquoi on préfère se placer dans l'espace de couleur xyY , où Y représente la luminance, et où les termes x et y représentent la chromaticité (informations de saturation et de teinte), indépendamment de la luminance.

Chaque longueur d'onde du spectre visible se trouve le long de la courbe incurvée (du bleu en bas à gauche au rouge tout à droite), et la droite reliant les deux extrémités de cette courbe s'appelle la droite des pourpres.

Lien avec la reproduction de tons

L'espace de couleur xyY est particulièrement important, car à partir d'une image dans l'espace de couleur RGB , un triplet dans l'espace xyY peut être calculé pour chaque pixel. Pour chaque pixel, la luminance sera découplée des informations sur la couleur. Pour les algorithmes de reproduction de tons, on peut donc se placer dans l'espace de couleur xyY , traiter la luminance indépendamment de la chromaticité, puis revenir dans l'espace de couleur de l'écran pour afficher l'image résultat. Cela permet de traiter la luminance sans modifier les chromaticités (ou tons) d'origines.

2.7 Propriétés des afficheurs

Le terme afficheur regroupe tous les supports sur lesquels les images peuvent être projetées. En ce qui concerne notre travail, ce sont principalement les écrans d'ordinateurs de bureau et les projecteurs. Les propriétés photométriques de ces appareils sont importantes pour la reproduction de tons, car elles sont utilisées pour optimiser la visualisation d'images. Les paragraphes suivants décrivent les grandeurs employées dans les catalogues de vente, et font la part des choses entre les données annoncées et les caractéristiques réelles des afficheurs.

2.7.1 Luminance des afficheurs

La luminance correspond à l'intensité émise par chaque pixel de l'écran, dans un angle solide d'un stéradian. En pratique, dans les catalogues de vente, elle est confondue avec la luminosité, alors qu'elle est pourtant bien donnée en cd/m^2 .

Les luminances affichées varient selon le type d'écran. Les anciens écrans à tube cathodique fonctionnaient avec une luminance maximale de l'ordre de $100 cd/m^2$, tandis que les catalogues de vente actuels indiquent des luminances maximales comprises entre 200 et $300 cd/m^2$ pour des écrans de bureaux, entre 200 et $500 cd/m^2$ pour les téléviseurs LCD, et environ $1500 cd/m^2$ pour un téléviseur plasma. Enfin, à propos des vidéoprojecteurs, on trouve des modèles de 2000 *lumens* jusqu'à 3 300 *lumens* pour le haut de gamme [Darty 10]. A titre indicatif, la luminance d'un écran HDR peut atteindre $4000 cd/m^2$.

Les écrans LCD et les vidéoprojecteurs LCD nous intéressent plus particulièrement, car ils sont beaucoup utilisés pour la simulation de conduite. Ces afficheurs fonc-

tionnent avec un rétro-éclairage permanent. Des cristaux liquides contrôlent l'affichage de l'image, en laissant passer - ou non - la lumière du rétro-éclairage vers l'observateur, et ce pour chaque couleur de chaque pixel.

Le défaut de cette technologie est que les cristaux liquides ne peuvent masquer entièrement la lumière quand ils doivent afficher du noir (ce qui revient à ne pas laisser passer la lumière). La conséquence est que les noirs ne sont jamais très intenses avec cette technologie d'écran : la luminance du noir reste en général assez élevée, notamment pour les écrans bas de gamme (environ 1 cd/m^2). Pour conserver un contraste satisfaisant, les écrans LCD ont donc besoin de luminances élevées pour compenser leur noir bien trop lumineux. Pour les professionnels de l'image, la norme ISO suggère simplement que la luminance du noir soit inférieure à 1 % de la luminance du blanc et que ce même noir ne soit pas supérieur à deux fois la luminance de l'écran éteint.

Enfin, indépendamment de la technologie de l'afficheur, nous pouvons remarquer que la luminosité a une échelle trompeuse, puisqu'elle est proportionnelle au logarithme de la luminance. Lorsqu'on passe d'une luminance de 200 cd/m^2 à 400 cd/m^2 , on ne double pas la luminosité, mais on l'augmente de 13 % seulement. Cependant, les constructeurs mesurent parfois la luminance maximale en mettant le rétro-éclairage au maximum, dans le but d'afficher sur leur catalogue une luminance la plus élevée possible, même si elle ne correspond pas à l'utilisation standard de l'écran. Il faut donc rester prudent quant aux valeurs de luminance annoncées sur les catalogues.

2.7.2 Contraste des afficheurs

Le taux de contraste est le rapport entre la valeur du noir et la valeur du blanc. Un taux de contraste de 500 : 1 par exemple signifie que le blanc de l'écran est 500 fois plus lumineux que le noir de l'écran.

Cependant, il est difficile de savoir quel est le taux de contraste réel d'un écran, car les constructeurs affichent le contraste dynamique (qui s'adapte au cours du temps en fonction des images affichées), et non le taux de contraste réel. Les constructeurs peuvent alors afficher des contrastes très élevés (comme 1 000 000 : 1), forcément plus vendeurs. Mais ces chiffres ne sont pas proportionnels au taux de contraste réel, qu'il est donc impossible de connaître d'après le catalogue, à moins qu'il ne soit indiqué séparément... ou que l'écran n'ait pas de contraste dynamique ! Dans ce cas là, on trouve des valeurs entre 600 : 1 et 1000 : 1 pour les écrans LCD, entre 300 : 1 et 3000 :

1 pour les téléviseurs LCD, entre 10 000 : 1 et 40 000 : 1 pour les téléviseurs plasmas, et entre 500 : 1 et 4000 : 1 pour les vidéoprojecteurs LCD (et jusqu'à 36 000 : 1 pour les DLP ou Digital Light Processing). A titre indicatif, le contraste d'un écran HDR est de 15000 : 1.

A titre d'exemple, certains écrans LCD d'entrée de gamme fournissent un noir qui ressemble à du gris foncé. S'ils affichent une valeur de luminance de 270 cd/m^2 et un taux de contraste de 400 : 1, on peut facilement en déduire la qualité du noir : $270 / 400 = 0,67 \text{ cd/m}^2$. Un écran LCD haut de gamme fournira un noir bien plus profond, à environ $0,1 \text{ cd/m}^2$.

Pour conclure, nous pouvons dire que c'est l'interaction entre les deux grandeurs, la luminosité et le contraste, qui est très importante. Pris séparément, le niveau de luminance ou le taux de contraste n'ont que peu d'intérêt. Mais un écran sera d'autant meilleur qu'il présente des valeurs élevées pour ces deux paramètres.

2.7.3 Gamma des afficheurs

Lors de l'apparition de la télévision au XXe siècle, les ingénieurs de l'époque furent confrontés à une restitution très sombre et contrastée des images sur les postes de télévision.

Le problème venait du fait que la plupart des écrans n'ont pas une relation linéaire entre leur tension d'entrée V et la luminance L qu'ils affichent. On peut modéliser la relation entre ces deux grandeurs avec la formule suivante :

$$L = kV^\gamma \quad (2.7)$$

La puissance γ , ou gamma de l'écran, modélise la non-linéarité de l'écran, et est différente d'une technologie d'écran à l'autre. Le facteur k correspond simplement à la luminance maximale L_{max} que peut afficher l'écran, en faisant l'hypothèse que la tension V est normalisée entre 0 et 1.

En pratique, les moniteurs ont un gamma compris entre 2.2 et 2.8. Le gamma des écrans LCD est souvent proche de 2.2, celui des écrans cathodiques est lui proche de 2.5. Le gamma des projecteurs varie lui sur une plage plus large : entre 1.6 et 3.9.

Après chaque opération sur la reproduction de tons, il faut donc prendre en compte le gamma de l'écran pour ne pas avoir d'image trop sombre. On applique donc les formules suivantes sur les triplets de chaque pixel des images à afficher :

$$R' = R^{\frac{1}{\gamma}} \quad (2.8)$$

$$G' = G^{\frac{1}{\gamma}} \quad (2.9)$$

$$B' = B^{\frac{1}{\gamma}} \quad (2.10)$$

Pour une visualisation optimale des images affichées, il convient donc de relever le gamma de l'écran sur le catalogue de ses caractéristiques techniques, et de mettre à jour sa valeur dans l'algorithme de reproduction de ton utilisé.

Chapitre 3

Le système visuel humain

Dans ce chapitre, nous nous proposons de décrire de manière simplifiée le fonctionnement du système visuel humain, ce qui nous permettra de comprendre les modèles de vision utilisés dans les algorithmes de reproduction de tons. Dans une première partie, nous aborderons l'anatomie de l'œil et le trajet de la lumière jusqu'à son absorption sur la rétine. Dans un second temps, nous détaillerons les caractéristiques de la perception visuelle. Nous y verrons comment l'œil réagit aux contrastes et aux luminances que nous avons décrits dans le chapitre précédent.

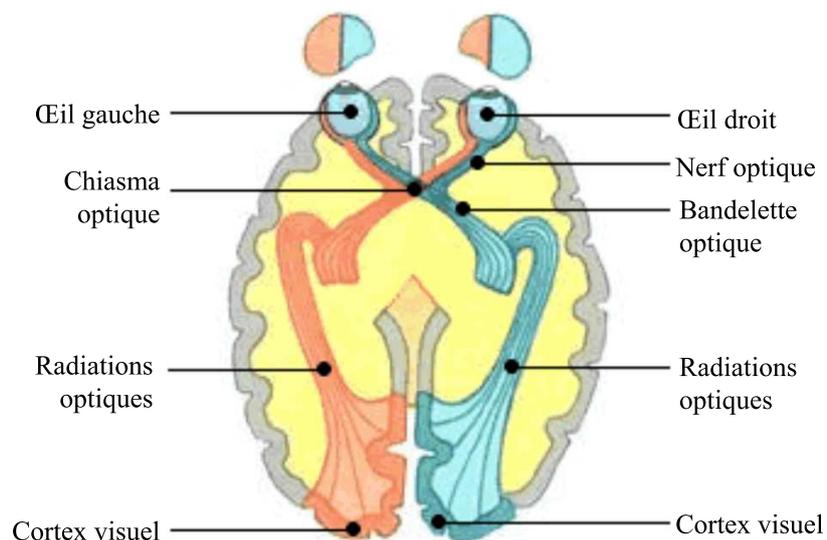


FIGURE 3.1 – Schéma des voies optiques du système visuel humain. (Image provenant de <http://pagesperso-orange.fr/jeanpierre.gadbois/yeux.htm>).

3.1 Anatomie de l'œil

Lorsque l'œil perçoit un signal lumineux, l'image est transmise au cortex visuel, la zone du cerveau traitant les informations visuelles. Des nerfs, bandelettes et radiations optiques permettent d'acheminer le signal visuel des yeux jusqu'au cerveau (figure 3.1). Dans cette partie, seuls les organes au niveau de l'œil, permettant d'acquérir les images transmises au cerveau, seront détaillés.

3.1.1 L'œil humain

La figure 3.2 présente une coupe longitudinale d'un œil humain. Lorsqu'un rayon lumineux entre dans l'œil, il rencontre successivement les organes suivants :

- une membrane muqueuse transparente, la conjonctive, composée d'un tissu fibreux ;
- une seconde membrane transparente, la cornée, principalement composée d'eau ;
- l'humeur aqueuse, également composée d'eau et qui approvisionne la cornée ;
- la pupille, qui est une ouverture de l'iris, et qui régule l'entrée de la lumière dans le globe oculaire ;
- une lentille auxiliaire molle, le cristallin ;
- une cavité postérieure remplie par un fluide albumineux (contenant la principale protéine du sang), le corps vitré ;
- et enfin la rétine, qui est composée de photorécepteurs tapissant la face interne de l'œil. La lumière forme alors une image inversée sur la rétine.

Le globe oculaire mesure environ 2,5 centimètres de diamètre. Les deux yeux sont séparés d'une distance de 6 centimètres environ, mais cette distance peut varier de quelques millimètres d'une personne à l'autre.

3.1.2 La rétine

La rétine est la couche de l'œil qui est sensible à la lumière. Ses photorécepteurs convertissent les photons qu'ils captent en signal neural et le transmettent aux cellules bipolaires. A leur tour, les cellules bipolaires transmettent ce signal aux cellules ganglionnaires et ainsi jusqu'au cerveau (figure 3.3). La rétine est composée d'environ 110 à 130 millions de photorécepteurs. A titre de comparaison, un appareil photo haut de gamme n'est composé que de 15 millions de pixels.

La rétine est composée de deux types de photorécepteurs : les bâtonnets et les

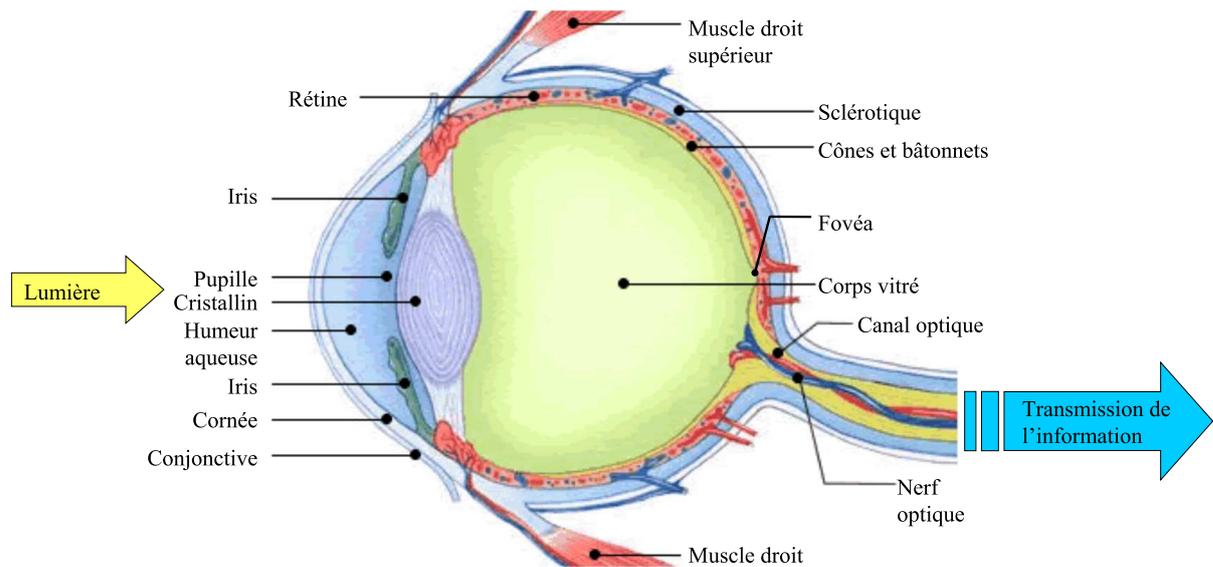


FIGURE 3.2 – Coupe sagittale d'un œil humain. (Image provenant de <http://pagesperso-orange.fr/jeanpierre.gadbois/yeux.htm>).

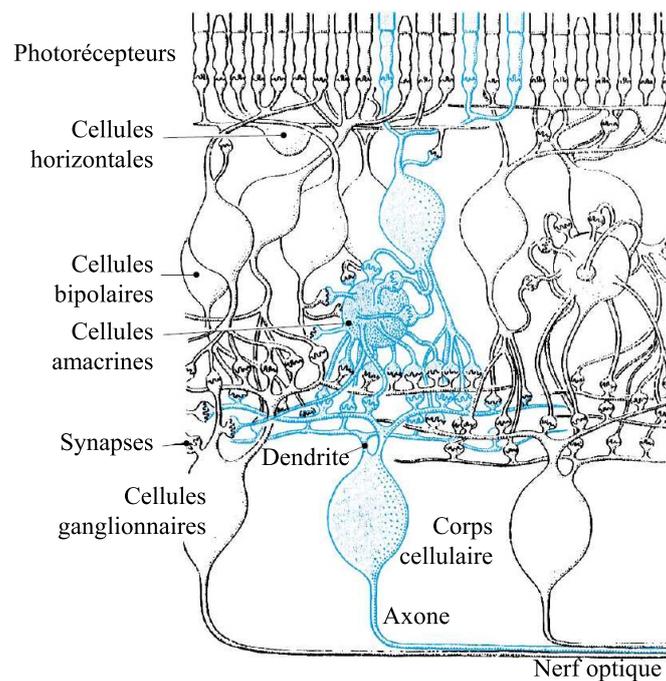


FIGURE 3.3 – Structure de la rétine. (Image provenant de [Grall 97]).

cônes, nommés ainsi à cause de leurs formes respectives. Ces deux photorécepteurs diffèrent non seulement par leur forme, mais aussi par leur domaine de perception et

leur sensibilité. Leur répartition sur la rétine n'est pas régulière (figure 3.4). La plupart des cônes sont concentrés sur la fovéa, permettant ainsi une vision de haute résolution sur le point de fixation de l'œil. À l'inverse, les bâtonnets sont répartis sur le reste de la rétine, avec une concentration quasi-nulle sur la fovéa. Enfin, on peut noter qu'il y a une portion de la rétine qui ne comporte ni cône ni bâtonnet : c'est là que se situe le nerf optique, qui transmet les informations visuelles au cerveau. Cependant, le cerveau est capable d'interpréter les images qui lui arrivent des deux yeux pour reconstruire une image unique et complète, en trois dimensions.

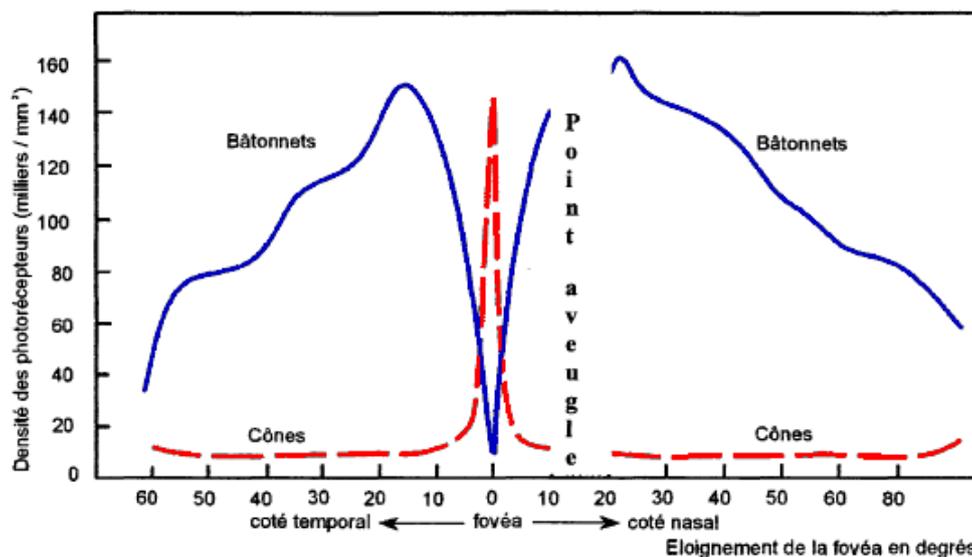


FIGURE 3.4 – Distribution angulaire des cônes et des bâtonnets sur la rétine. (Image provenant de http://artic.ac-besancon.fr/svt/act_ped/svt_lyc/eva_bac/l-bac2007/bac2007-li.htm)

3.1.3 Les bâtonnets

Les bâtonnets composent 95 % des photorécepteurs de la rétine, soit environ 100 à 120 millions d'unités. Ce sont des capteurs très sensibles à la lumière, permettant de voir dans des conditions de faible illumination, du crépuscule jusque dans l'obscurité (vision scotopique). Cette vision n'est qu'en niveaux de gris : ils ne perçoivent pas les couleurs car ils sont tous sensibles de la même manière aux longueurs d'onde comprises entre 400 et 650 nm. Les bâtonnets sont également des cellules très sensibles aux mouvements.

3.1.4 Les cônes

Les cônes ne constituent que 5 % des photorécepteurs de la rétine, soit de 7 à 8 millions d'unités environ. Ce sont des capteurs moins sensibles à la lumière que les bâtonnets, mais qui permettent une vision en couleurs dans des conditions de plus forte illumination (vision photopique).

Chez l'homme, il existe sur la rétine trois types de cônes :

- les cônes L (pour "long"), sensibles aux longueurs d'onde de basse fréquence du spectre visible, avec un pic de sensibilité autour de 498 nm ;
- les cônes M (pour "middle"), sensibles aux longueurs d'onde de moyenne fréquence, avec un pic de sensibilité autour de 534 nm ;
- les cônes S (pour "short"), sensibles aux longueurs de haute fréquence, avec un pic de sensibilité autour de 564 nm.

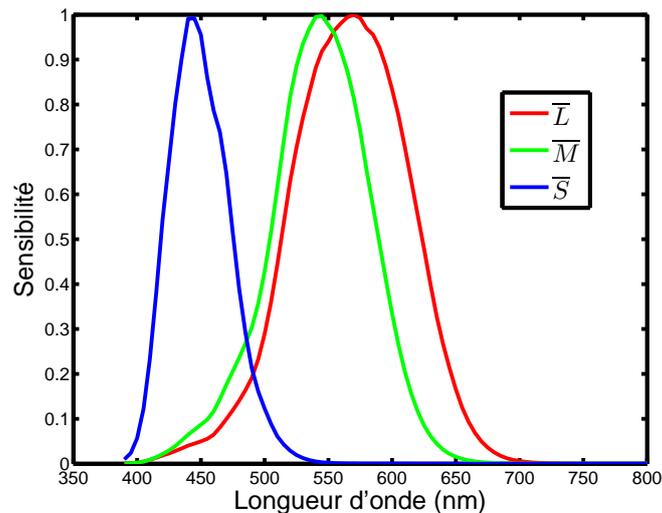


FIGURE 3.5 – Sensibilité spectrale des pigments visuels des cônes. (Données téléchargées sur <http://www-cvrl.ucsd.edu/>).

La figure 3.5 montre la sensibilité spectrale de chaque cône. On constate qu'il y a souvent recouvrement entre les différentes courbes, ce qui signifie qu'un même signal lumineux active généralement les 3 types de cônes.

3.2 Perception visuelle

Après avoir décrit les organes permettant la vision chez l'homme, nous nous intéressons maintenant à la perception visuelle de l'environnement offerte par le système visuel humain.

La dynamique de luminance à laquelle nous pouvons faire face chaque jour est très vaste : la lumière du soleil est environ 10 millions de fois plus intense que la lumière des étoiles (soit jusqu'à 7 unités logarithmiques) [Ferwerda 96].

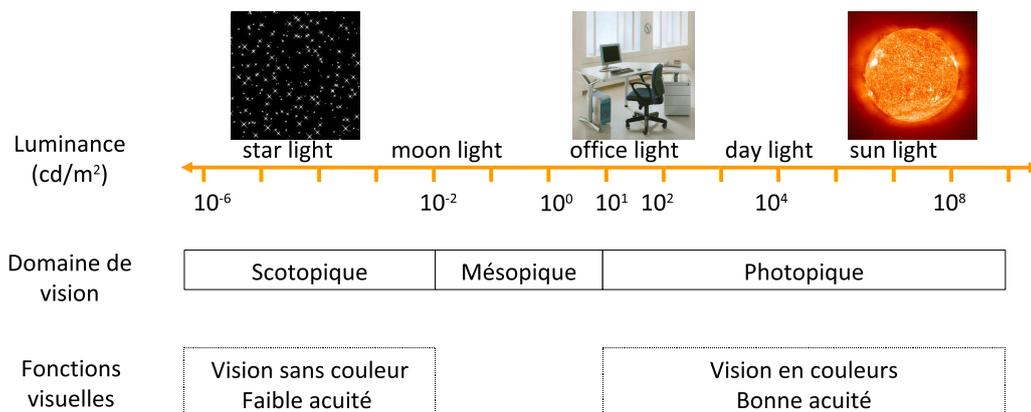


FIGURE 3.6 – Dynamique de luminance que rencontre l'œil humain dans le monde réel. Figure adaptée de [Spillman 90].

L'œil humain n'est cependant pas capable de percevoir toute la dynamique de luminance au même instant. Mais face à cette très grande dynamique, il est capable de s'adapter à la luminosité ambiante, selon trois types de vision :

- la vision **photopique** (diurne), au cours de laquelle seuls les cônes sont utilisés pour la vision, les bâtonnets étant saturés. La vision est alors en couleurs, et l'acuité est la plus importante au niveau de la fovéa. Les cônes étant néanmoins moins sensibles à la luminance que les bâtonnets, les faibles différences de luminance y sont donc difficilement perceptibles. C'est pourquoi il n'est par exemple pas possible de voir les étoiles de jour, avec l'atmosphère ensoleillée. Le domaine photopique correspond aux luminances de 10 à 10⁸ cd/m².
- la vision **mésopique** (crépusculaire), qui utilise à la fois les cônes et les bâtonnets. La perception des couleurs et l'acuité sont atténuées dans ce domaine, par rapport au domaine photopique. Le domaine mésopique correspond aux luminances de 10⁻² à 10 cd/m².

- la vision **scotopique** (nocturne), où seuls les bâtonnets fonctionnent, les cônes n'étant pas sensibles aux faibles luminances. Les bâtonnets ne permettent qu'une vision en niveaux de gris, mais sont très sensibles aux petites différences de luminance. Du fait de la concentration plus faible des bâtonnets sur la rétine (par comparaison à celle des cônes), l'acuité visuelle est plus faible dans le domaine scotopique (c'est pourquoi la lecture est plus difficile dans l'obscurité qu'à la lumière). Cependant, la perception des mouvements en vision périphérique reste bonne. Le domaine scotopique correspond aux luminances de 10^{-6} à 10^{-2} cd/m^2 .

3.2.1 Perception des contrastes

La grandeur à laquelle l'œil est le plus sensible est sans aucun doute le contraste. Lorsque l'œil observe l'environnement qui l'entoure, il apprécie les variations de luminance qui lui permettent de reconnaître les objets et personnes, mais aussi le niveau lumineux ambiant.

Une expérimentation de Baker [Blackwell 46] a permis de déterminer quel était le seuil de détection du système visuel humain. Ce seuil de perception est aussi appelé la différence de contraste juste détectable (*just-noticeable difference* ou JND). Cette expérimentation consiste à afficher sur un écran une image uniforme, de luminance de fond donnée I à laquelle un sujet s'habitue. Puis, un spot lumineux, d'une luminance $I + \Delta I$ légèrement supérieure à la luminance du fond, est flashé à l'écran. Le sujet doit dire s'il détecte ou non ce spot, afin de déterminer la différence de contraste juste détectable pour chaque luminance donnée.

La figure 3.7 illustre les résultats de cette expérience. Les bâtonnets sont actifs dès 10^{-6} cd/m^2 mais saturent à partir de 10 cd/m^2 , alors que les cônes sont actifs à partir de 0.01 cd/m^2 mais ne saturent jamais dans un niveau lumineux stable, quel que soit son intensité ! Seuls de brusques changements lumineux peuvent faire saturer les cônes. Enfin, pour des luminances inférieures à 1 cd/m^2 , les bâtonnets sont plus sensibles aux petites luminances que les cônes : l'amplitude du courant induite par l'absorption d'un photon est 20 fois plus forte chez les bâtonnets que chez les cônes (700 femtoampères contre 33) [Rodieck 03]. Pour des luminances supérieures à 10 cd/m^2 , on retrouve la loi de Weber, selon laquelle le ratio $\Delta I/I$ est constant quelle que soit la luminance.

Enfin, il est à noter que chaque seuil est donné pour une luminance et un état d'adaptation donnés. On ne peut donc pas utiliser directement ces valeurs si l'œil est dans un état d'adaptation différent de la luminance du fond.

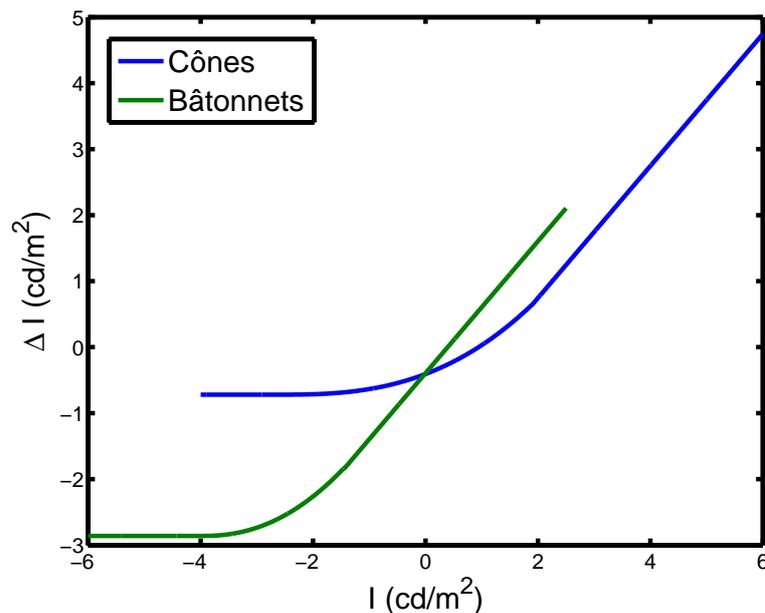


FIGURE 3.7 – Seuils de détection des contrastes en fonction de la luminance et de l'état d'adaptation de l'œil [Rushton 86].

3.2.2 Luminance d'adaptation

L'état d'adaptation de l'œil est un paramètre très important, car il a une grande influence sur la manière dont l'environnement sera perçu. Pour définir cet état, on utilise la luminance d'adaptation L_a , qui correspond à la luminance à laquelle l'œil est adapté à un instant t . L'état d'adaptation est fonction de l'état dans lequel se trouve le système visuel humain à l'instant $t - 1$, mais aussi de la luminance vue par l'œil à l'instant t . Pour être parfaitement adapté à un environnement, l'œil a donc besoin de temps pour s'habituer à la luminance qui lui fait face (ce temps varie de quelques millisecondes à plusieurs minutes suivant l'état d'adaptation dans lequel il se trouve). Le système visuel humain restera ensuite adapté à cet état, si la luminance de l'environnement observé ne change pas.

En traitement d'images, et plus particulièrement dans les algorithmes de reproduction de tons, la luminance d'adaptation est difficile à calculer précisément, les images manipulées étant la plupart du temps très complexes (par opposition à une image de couleur uniforme et de luminance constante). Les premiers opérateurs proposaient d'utiliser une moyenne logarithmique appliquée à l'intensité de l'image entière, en ajoutant une valeur ε très petite (pour éviter d'éventuelles divergences dues à des pixels noirs, d'intensité nulle), et où n est le nombre de pixels de l'image

[Tumblin 93, Ward 94] :

$$L_a = \exp \left(\frac{1}{n} \sum_{i=1}^n \log_{10}(I_i + \epsilon) \right) \quad (3.1)$$

Les moyennes arithmétiques ou géométriques sur l'ensemble des pixels ont également été proposées, mais donnent de moins bons résultats. En l'absence de données ou de connaissances sur l'image observée, cette définition est donc la plus appropriée. Elle permet d'obtenir une approximation de l'adaptation globale (sur l'ensemble de l'image) du sujet.

Dans certains cas, face à des images dotées d'une très grande dynamique de luminance, le résultat obtenu avec une adaptation globale peut être décevant, et on peut alors avoir recours à une adaptation locale. Moon and Spencer [Moon 45] proposent ainsi de calculer la luminance d'adaptation fovéale, prenant en compte la luminance vue par la fovéa (ie. le point de fixation de l'observateur) mais aussi les luminances vues par la parafovéa. La zone fovéale correspond au demi-angle de vision de 0.5 degrés centré sur le point de fixation, la zone parafovéale à un demi-angle de vision compris entre $\theta_f = 0.5$ et $\theta_{pf} = 57$ degrés (cette dernière valeur a néanmoins peu d'influence sur la valeur d'adaptation).

$$L_a = 0.913L_f + \frac{0.0096}{\pi} \int_{\phi=0}^{360} \int_{\theta_f}^{\theta_{pf}} \frac{L(\theta, \phi)}{\theta^2} \cos(\theta) \sin(\phi) d\theta d\phi \quad (3.2)$$

La constante 0.913 résulte de l'intégration de la luminance L_f sur la fovéa. Chaque valeur de luminance y est considérée comme d'égale importance pour l'adaptation. La luminance que voit la fovéa est donc la plus importante pour l'adaptation, puisqu'elle contribue pour plus de 91 % à la luminance vue par l'œil.

3.2.3 Adaptation temporelle

Quand on se déplace d'un environnement ensoleillé vers une pièce sombre, la vision est d'abord difficile, puis l'œil s'habitue à l'obscurité au bout de quelques instants, et la vision redevient normale. Inversement, si l'on est habitué à l'obscurité d'une pièce sombre et que l'on sort vers un extérieur très ensoleillé, on est ébloui par la lumière : on plisse alors les yeux pour se protéger, jusqu'à ce que nos yeux s'accoutument à la forte luminosité ambiante.

Cette adaptation temporelle est un processus essentiel dans la perception visuelle. De plus, sa restitution à l'écran est un réel enjeu pour la simulation de conduite, car

le conducteur peut circuler dans des espaces où la dynamique des luminances peut considérablement varier (les exemples classiques étant le soleil rasant, l'entrée et la sortie de tunnel, ou le passage d'une zone ensoleillée à une zone sombre). Les paragraphes suivants détaillent les différents mécanismes qui entrent en jeu.

Les différents mécanismes de l'adaptation

Il existe au moins quatre mécanismes d'adaptation à la lumière et à l'obscurité, qui permettent à l'œil de modérer les changements brusques de luminosité, et de maintenir une sensibilité sur la grande dynamique des luminances perceptibles.

L'iris contrôle le premier mécanisme de l'adaptation, en régulant la quantité de lumière éclairant la rétine via l'ouverture de la pupille. La nuit, en vision scotopique, la pupille est dilatée à son maximum, à un diamètre d'environ 7 mm , tandis que de jour, la pupille se contracte pour limiter la quantité de lumière entrante, à un diamètre d'environ 2 mm . Ce mécanisme conduit à un changement dans l'adaptation d'une à deux unités logarithmiques, ce qui est encore loin de couvrir les 7 unités auxquelles l'œil peut être confronté. Ce mécanisme est donc négligé par les opérateurs de reproduction de tons.

Néanmoins, la taille de la pupille reste un paramètre déterminant si l'on veut simuler certaines aberrations. Elles sont plus importantes de nuit, lorsque la pupille dilatée à son maximum, que de jour, lorsque la pupille contractée. Pour simuler un halo dû à l'éblouissement face à une source de lumière très lumineuse, la taille de la pupille est donc prise en compte (voir [Spencer 95, Kakimoto 05, Ritschel 09] et le chapitre 5). Pour cela, Spencer propose d'exprimer le diamètre de la pupille D (en mm) en fonction de la luminance d'adaptation L_a (en cd/m^2) par la relation suivante, valable de jour comme de nuit [Spencer 95] :

$$D = 4.9 - 3 \tanh(0.4(\log_{10}(L_a) + 1)) \quad (3.3)$$

A travers un second processus, les photorécepteurs (cônes et bâtonnets) participent également au mécanisme de l'adaptation. Ils contiennent des pigments visuels, appelés rhodopsines, qui captent la lumière sur la rétine. Ces pigments sont constitués de molécules composées de deux parties : une portion protéique et une portion qui absorbe la lumière, appelée chromophore [Rodieck 03]. Lorsque la lumière atteint la rétine, les photons sont absorbés par le chromophore en seulement 3 femtosecondes

(fs). Le chromophore de la rhodopsine change alors de configuration chimique en 200 fs : c'est la photo-isomérisation. Durant la photo-isomérisation, le chromophore passe d'une forme repliée à une forme droite. Tel un interrupteur qui s'ouvre et laisse passer le courant, la rhodopsine diffuse alors des neurotransmetteurs qui acheminent le signal visuel. Le temps d'activation de cet interrupteur est toutefois plus long, de l'ordre de 0.25 à 0.5 *ms*.

Une fois le signal envoyé, le chromophore dans sa position droite est éliminé, et remplacé par un nouveau chromophore en position courbe, prêt à recevoir un nouveau photon. Ce processus de régénération dure environ 110 *ms* pour les cônes et d'environ 400 *ms* pour les bâtonnets. C'est ce temps de réponse qui est pris en compte dans les algorithmes de reproduction de tons doté d'une modélisation du système visuel humain [Pattanaik 00].

Après la transmission du signal, le chromophore se régénère ensuite par réaction chimique en sa forme originelle, et est prêt à recevoir un nouveau photon.

La figure 3.8, tirée de [Ferwerda 96], présente la variation du seuil de sensibilité des cônes puis des bâtonnets au cours de temps, pour une adaptation à l'obscurité. Comme nous l'avons vu précédemment, les cônes s'adaptent plus vite que les bâtonnets. Ce sont donc eux qui permettent dans un premier temps de s'adapter à l'obscurité. Dans un second temps, après quelques minutes, l'adaptation des bâtonnets permet d'améliorer la perception des faibles luminances, sur plus de 2 unités logarithmiques.

Les figures 3.9 et 3.10, tirées de [Ferwerda 96], présentent la variation du seuil de sensibilité des bâtonnets et des cônes au cours de temps, pour une adaptation à la lumière. Nous constatons que le temps d'adaptation des bâtonnets est très rapide, de l'ordre d'une seconde, tandis que le temps d'adaptation des cônes est plus long, de l'ordre de 5 minutes. Nous remarquons également que les bâtonnets s'adaptent progressivement, alors que les cônes s'adaptent en deux temps, en permettant temporairement de détecter un seuil plus faible que le seuil vers lequel ils convergent. Enfin, pour les cônes comme pour les bâtonnets, l'adaptation à la lumière est plus rapide que l'adaptation à l'obscurité.

Il existe également deux mécanismes adaptatifs situés dans le réseau neural. Ils ajustent leur activité à l'aide d'un gain appliqué sur la réponse des photorécepteurs. Un mécanisme multiplicatif ajuste ce gain en fonction de la luminance ambiante. Ce processus est très rapide, il réagit sur le changement de sensibilité dans les premières

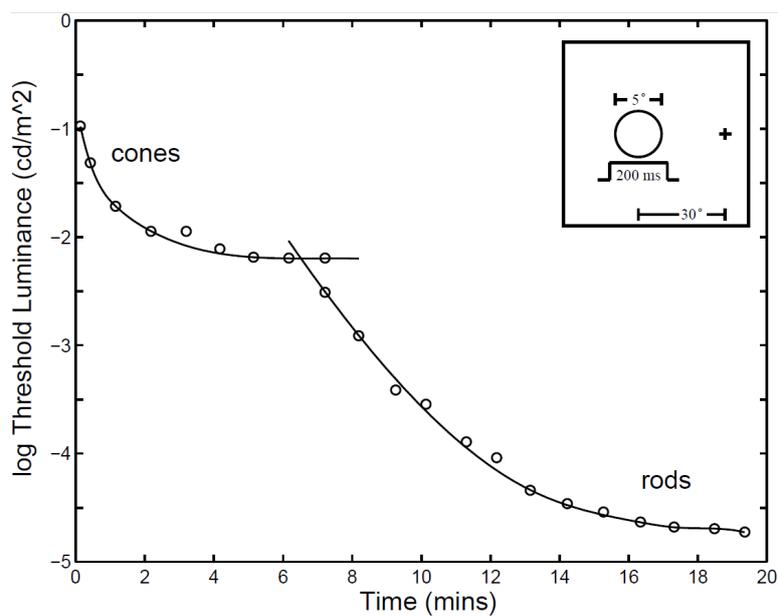


FIGURE 3.8 – Evolution du seuil de détection des contrastes en fonction de l’adaptation temporelle des photorécepteurs, face à une chute brutale de luminance [Ferwerda 96].

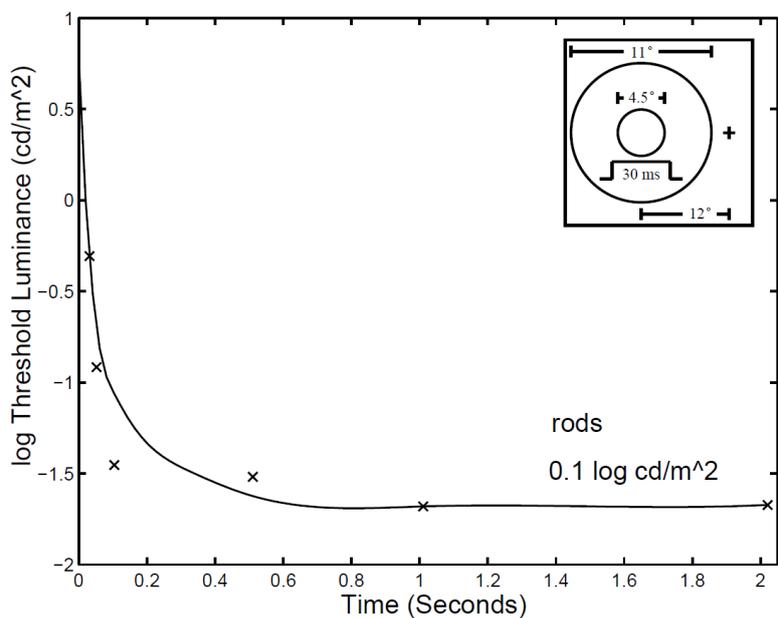


FIGURE 3.9 – Variations du seuil de détection des bâtonnets lors d’une adaptation à la lumière. D’après [Adelson 82].

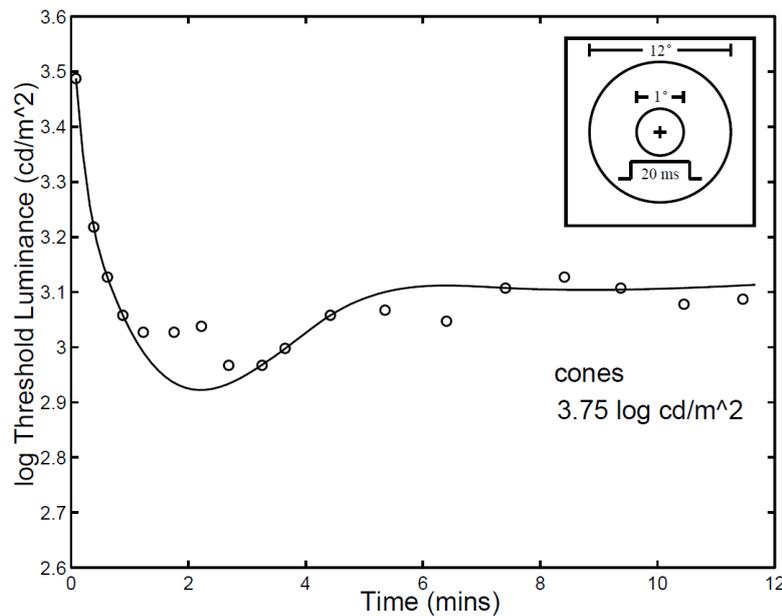


FIGURE 3.10 – Variations du seuil de détection des cônes lors d’une adaptation à la lumière. D’après [Baker 49].

secondes qui suivent le changement de niveau lumineux. Un processus soustractif plus lent réduit le niveau d’activité du système si le niveau lumineux est constant. Ce dernier processus permet une légère amélioration de la sensibilité après plusieurs minutes d’adaptation [Adelson 82, Pattanaik 00]

3.2.4 Éblouissement

Lorsque l’on regarde dans une direction proche d’une source très lumineuse en comparaison du reste de l’environnement, l’œil est ébloui. Le flux de lumière qu’il reçoit est trop élevé pour l’état d’adaptation de l’œil. Si la source lumineuse est très intense (comme le soleil), l’observateur peut même ressentir une gêne, qui l’incitera à cligner des yeux. De plus, même s’il ne ressent pas de douleur, l’éblouissement diminue le confort visuel et conduit à une baisse des performances visuelles, avec les effets suivants :

- la superposition d’un voile lumineux sur la source, qui empêche l’observateur de correctement discerner les objets voisins de la source ;
- si la source est de faible taille angulaire, l’apparition d’une couronne ciliaire : un ensemble de stries radiales, centrées sur la source ;
- de nuit, l’apparition d’un halo coloré autour de la source.

L'ensemble des effets de l'éblouissement est représenté sur la figure 3.11.



FIGURE 3.11 – Simulation des différents effets de l'éblouissement sur l'œil humain pour une source ponctuelle. Au centre la luminance de voile, autour du centre, la couronne ciliaire, et enfin, le halo lenticulaire. D'après [Ritschel 09].

La luminance de voile

Lorsqu'un rayon lumineux traverse l'œil, il est dispersé au niveau de la cornée, du cristallin et de la rétine principalement, selon les proportions indiquées dans le tableau 3.1. Il résulte de cette dispersion de la lumière l'apparition d'un voile lumineux autour de la source qui émet le rayon lumineux.

La luminance de voile peut être une gêne dans des tâches d'observation, puisque les contrastes autour de la source lumineuse sont fortement réduits. La figure 3.12 illustre ce problème, avec un schéma d'œil face à une source lumineuse au point A et un point B . Les points A et B ont leurs points image sur la rétine, respectivement en a et b . Mais le rayon lumineux provenant de la source A se disperse également au point b (flèche en pointillés). Soient L_1 et L_2 deux luminances différentes formant une différence de contraste au point B , et soit s la quantité de lumière provenant de la source A dispersée

Partie de l'œil	Pourcentage de dispersion
Cils	Pourcentage variant
Cornée	25 - 30 %
Humeur aqueuse	0 %
Cristallin	40 %
Iris	≤ 1 %
Humeur vitreuse	10 %
Rétine	20 %

TABLE 3.1 – Contribution à la dispersion de la lumière de chaque partie de l'œil. D'après [Ritschel 09].

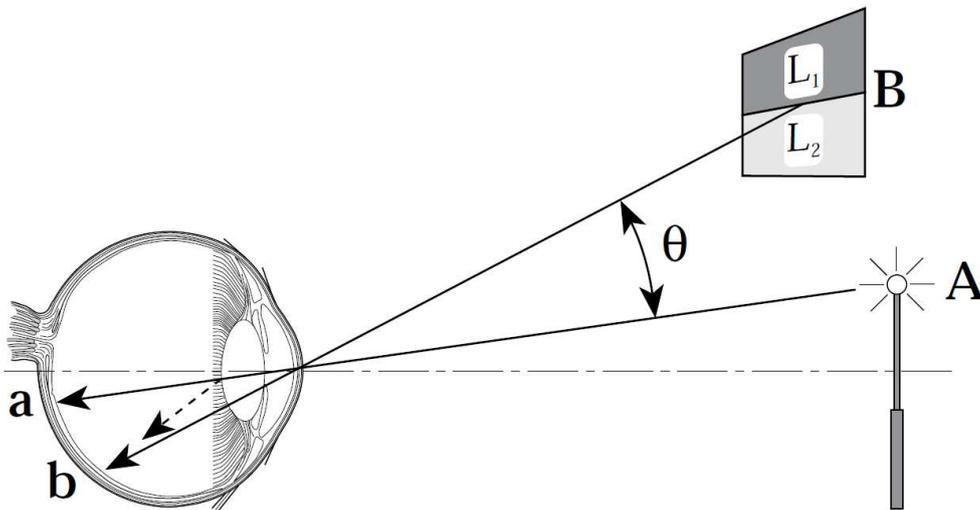


FIGURE 3.12 – Réduction du contraste par un éblouissement. D'après [Spencer 95].

au point b . La différence de contraste L_2/L_1 devient alors $L_2 + s/L_1 + s$, telle que :

$$\frac{L_2 + s}{L_1 + s} < \frac{L_2}{L_1} \quad (3.4)$$

Cette diminution du contraste dépend de la valeur de s . Plus s est élevé, et plus le contraste est diminué. En pratique, la quantité de lumière dispersée s dépend de l'angle de séparation θ entre les points A et B , ainsi que de la luminance et de l'angle solide par rapport à l'œil de la source A .

La manière dont se disperse la lumière dans l'œil a donné lieu à de nombreuses études. Pour la cornée et le cristallin, la lumière est dispersée dans un cône étroit et selon une distribution gaussienne [Campbell 66].

La couronne ciliaire

La couronne ciliaire est un ensemble de stries autour des objets très lumineux. Elles sont dues aux fluctuations d'indices de réfraction des milieux transparents du cristallin. Ces stries sont d'autant plus fines et brillantes que la source est lumineuse, d'un faible angle solide pour l'œil. A l'inverse, si la taille angulaire de la source lumineuse est grande (à partir de 20 minutes d'angle), les stries deviennent moins nettes, et contribuent alors à la luminance de voile.

Le halo lenticulaire

Lorsque l'on regarde un objet très lumineux dans un environnement où le niveau lumineux est faible (visions mésopique et scotopique), comme par exemple lorsque l'on regarde la source d'un lampadaire de nuit, il apparaît une série d'anneaux concentriques colorés, entourant la source de lumière. Ces anneaux forment un halo lenticulaire.

Ce phénomène est dû à la diffraction de la lumière sur un réseau de fibres radiales sur la lentille de l'œil, le cristallin [Spencer 95]. Si l'on considère une petite partie du cristallin, on peut considérer les fibres radiales comme parallèles entre elles. Un rayon lumineux entrant, de longueur d'onde λ , est alors diffracté d'un angle θ perpendiculaire aux fibres selon la formule suivante :

$$\sin(\theta) = \frac{\lambda}{e} \quad (3.5)$$

où e est l'espace entre chaque fibre.

La taille du halo tel qu'il est vu par l'œil est constante et indépendante de la distance entre l'observateur et la source. En revanche, l'intensité du halo décroît avec la distance de la source. Enfin, les couleurs des anneaux varient avec leur distance par rapport au centre : du bleu au rouge, en passant par les couleurs d'un arc-en-ciel, comme l'on peut le voir sur la figure 3.11.

Chapitre 4

Etat de l'art sur le rendu visuel

Les écrans standards affichent des images quantifiées sur 16 millions de couleurs (256^3), avec des luminances généralement comprises entre 0.1 et 300 cd/m^2 , les écrans plasmas seuls pouvant descendre jusqu'à 0.01 cd/m^2 . Cela reste bien peu par rapport à la réalité. Les images HDR ont pour but de palier ce problème d'affichage (en anglais, images *High Dynamic Range*, mais nous utiliserons par la suite le terme HDR). Leur utilisation nécessite deux étapes. D'abord, nous devons créer l'image HDR, qui contient des luminances réelles (et que nous ne pouvons donc pas afficher directement à l'écran). Puis, nous utilisons un opérateur de reproduction de tons (ou en anglais *tone mapping*) pour compresser la dynamique de luminance des images HDR et la faire correspondre à celle des écrans.

Les études pour la sécurité routière peuvent tirer avantage des images HDR, par l'intermédiaire de simulations routières et d'expériences psycho-physiques. Pour ne pas biaiser ces études, il y a cependant un réel besoin d'afficher à l'écran les scènes routières à l'écran telles qu'elles seraient perçues dans la réalité. Les luminances, contrastes, couleurs doivent être fidèles au contenu de l'image HDR, même si les écrans ne permettent pas de le reproduire exactement.

Pour une application en simulation de conduite, comme pour de nombreuses autres applications, le rendu HDR, et surtout l'opérateur de reproduction de tons, doivent être utilisés de manière à ce que le résultat affiché soit fidèle à la réalité. Mais les différents processus nécessaires sont complexes, c'est pourquoi ce chapitre contient un état de l'art sur chacun d'entre eux.

Nous présenterons d'abord un état de l'art sur le rendu HDR. Ensuite, un état de l'art sur les différents opérateurs de reproduction de tons existants montrera les avantages et inconvénients de chaque opérateur, et présentera également quelques protocoles d'évaluations ainsi que leurs résultats. Dans une troisième partie, différentes techniques pour la simulation de l'éblouissement seront décrites. Puisque les écrans standards ne peuvent afficher des zones très lumineuses, il est en effet important de faire comprendre inconsciemment à l'observateur qu'il est face à une zone très lumineuse, en simulant l'éblouissement qu'il ressentirait s'il était face à la même zone lumineuse en situation réelle. Enfin, nous aborderons la saillance d'images et de vidéos. Les algorithmes de calcul de la saillance permettent de prédire les points d'attention d'un observateur lorsqu'il regarde une image ou une vidéo, et nous proposerons dans le chapitre 6 de les utiliser avec des images HDR, pour évaluer l'impact d'un opérateur de reproduction de tons.

4.1 Le rendu HDR

Il existe deux moyens de créer une image HDR : par photographie ou par synthèse d'images. Si la synthèse permet la création de vidéos HDR, la méthode photographique exige encore aujourd'hui plusieurs prises de vue d'une même scène pour reconstruire une image HDR. Ces prises de vue pouvant durer plusieurs secondes, il est donc impossible de faire des vidéos, excepté image par image (*stop motion*), mais cette méthode se révèle vite très fastidieuse. Il existe néanmoins quelques prototypes d'appareils photos et de caméras permettant d'acquérir une image HDR en une seule prise, mais ils ne sont pas encore disponibles pour le grand public (voir les capteurs Viper FilmStream [Grassvalley 10], Pixim [Pixim 10] ou SpheroCam HDR [Spheron 10]).

4.1.1 Imagerie HDR par photographie

Chaque appareil-photo est doté d'un capteur qui mesure la luminance de la scène qui lui fait face (voir la section 2.3.1). Mais ces capteurs ont une dynamique trop faible pour couvrir la gamme réelle de luminances. Leur dynamique est d'environ 4 unités logarithmiques [Reinhard 05b], mais elle peut être réduite si une source très lumineuse éblouit le capteur lors de la prise de vue (voir section 3.2.4 pour la réduction du contraste).

Face à cette limitation technologique, Mann et Picard [Mann 95], puis Debevec et Malik [Debevec 97] ont proposé d'utiliser différentes expositions pour reconstruire

une carte de luminance de la scène photographiée, *ie.* l'image HDR. Grâce à ces différentes expositions, chaque pixel se retrouve correctement exposé sur au moins une des expositions, à l'exception des pixels vraiment trop lumineux (soleil) ou trop sombres (zones très peu éclairées).

Il reste ensuite à sélectionner les pixels correctement exposés de chaque exposition. Plusieurs possibilités existent [Mann 95, Debevec 97, Mitsunaga 99], et consistent à donner plus d'importance aux pixels avec des valeurs comprises entre 50 et 200 qu'aux autres pixels (les pixels ont des valeurs comprises entre 1 et 255). Ces derniers sont en effet considérés comme suspects car proches de la saturation (pour les valeurs au delà de 200) ou bruités (pour les valeurs inférieures à 50).

Enfin, en faisant l'hypothèse que la réponse du capteur est linéaire (capteur CCD), nous pouvons diviser chaque pixel par le temps d'exposition, pour les rendre fonction d'une même grandeur, proportionnelle à un facteur près à la luminance. Ce facteur peut être déterminé soit par la connaissance de la luminance d'un objet de la scène (par exemple une source lumineuse), soit en mesurant la luminance d'un point donné avec un luminancemètre.

L'imagerie HDR par photographie n'aime pas le mouvement

Des difficultés supplémentaires existent lors de l'acquisition et de la reconstruction de l'image HDR.

D'une part, lors de l'acquisition des différentes prises de vue, l'appareil-photo peut bouger, et les pixels des différentes expositions ne correspondent alors plus entre eux. Nous pouvons alors utiliser un pied et/ou une technique d'alignement d'images en cas de besoin. A propos de l'alignement d'images, deux méthodes applicables aux images HDR ont été proposées. Kang et *al.* [Kang 03] estiment le mouvement de l'appareil-photo et d'éventuels objets de la scène en utilisant la réponse de la caméra. Ward [Ward 03] estime quant à lui le mouvement de la caméra sans la réponse de la caméra, mais ne peut en revanche corriger le mouvement d'objets dans la scène sans un algorithme complémentaire. La méthode de Ward est plus rapide, mais se limite à de faibles mouvements de l'appareil-photo.

D'autre part, lors de la reconstruction de l'image HDR, nous pouvons être face à des objets qui se déplacent au cours du temps (personnes, voitures, nuages,...). Des fantômes apparaissent alors sur les images HDR, dus aux déplacements des objets entre les différentes expositions. Pour éviter ce problème, une méthode consiste à ne retenir les valeurs de luminance que d'une seule exposition [Reinhard 05b]. Pour cela,

on calcule la variance de chaque pixel entre les différentes expositions. Les zones de forte variance correspondront aux objets mobiles. Une fois ces zones détectées et segmentées, une interpolation linéaire entre la luminance de points proches de variance faible (points fiables) et la luminance des points proches de variance élevée (points suspects) est utilisée pour déterminer de quelle exposition nous retiendrons les valeurs de luminance.

4.1.2 Imagerie HDR par synthèse d'images

La création d'images de synthèse HDR a d'abord été possible avec les moteurs de rendus par lancer de rayons. Dès 1985, le logiciel *Radiance* de G. Ward [Ward 98] permettait de faire de la simulation d'éclairage, et de sauvegarder le résultat sous un format d'image HDR. Aujourd'hui, la plupart des modeleurs (comme 3DS Max, Maya [Autodesk 10] ou Blender [Blender 10]), permettent de faire des rendus HDR avec un moteur de lancer de rayons.

Le rendu HDR temps réel est lui devenu possible à partir de 2003, lorsque les cartes graphiques ont pu stocker les valeurs exactes de la luminance des images dans la mémoire d'images (*framebuffer*), en valeurs flottantes (*float*), sur 16 puis 32 bits . Auparavant, les valeurs étaient stockées dans un *framebuffer* sur 8 bits, avec les valeurs écrêtées entre 0 et 1 et sur une dynamique de 256 valeurs, ce qui rendait impossible le stockage de valeurs réelles de luminances. Une fois l'image HDR stockée dans un *framebuffer* d'une mémoire suffisante, la reproduction de tons est ensuite réalisée en tant que post-traitement, et éventuellement accompagnée de traitements supplémentaires pour simuler la vision humaine (comme la simulation de l'éblouissement). Dans les applications temps réel, ces post-traitements sont généralement réalisés avec des shaders, qui peuvent directement exploiter les valeurs 16 ou 32 bits contenues dans le *framebuffer*, sans que ces valeurs ne soient écrêtées entre 0 et 1.

Quelques bibliothèques permettant un rendu HDR et les post-traitements en temps réel ont émergé, notamment dans les moteurs de rendu *open-source* OGRE3D [OGRE3D 10] et OpenSceneGraph [OSG 10, OSGPPU 10]. Les deux applications permettent d'afficher un objet quelconque, et de l'illuminer. L'image de l'objet éclairé est stockée dans une texture 32 bits, qui est ensuite post-traitée (figure 4.1).

Le rendu HDR devient également une caractéristique standard dans les jeux vidéo, comme Half Life [Mitchell 06], Far Cry [Crytek 10] et Oblivion [Oblivion 10]. Il est utilisé pour mieux représenter les objets lumineux, et pour simuler l'adaptation visuelle lorsque le joueur passe d'une zone très lumineuse à une zone très sombre, ou

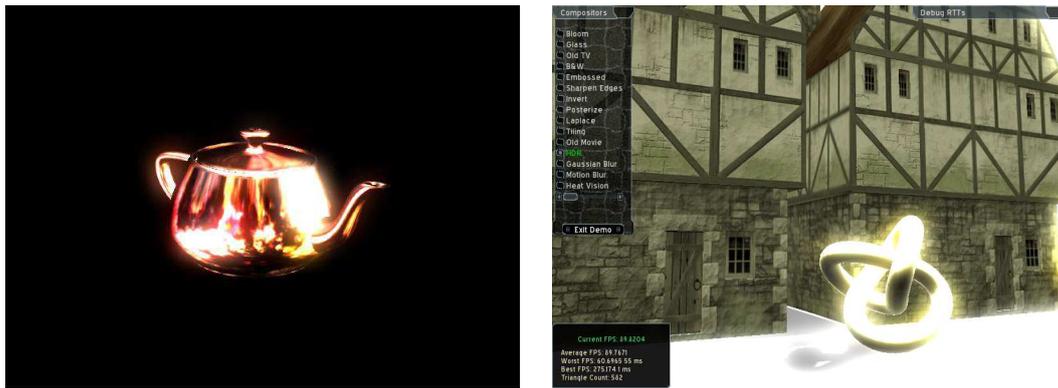


FIGURE 4.1 – Captures d’écran des modes HDR des moteurs de rendu OpenScene-Graph [OSG 10] (à gauche) et Ogre3D [Ogre3D 10] (à droite).

inversement. Ce mode reste cependant une option désactivable, car le temps d’adaptation est parfois vécu comme une gêne pour un joueur qui veut être adapté et prêt à réagir au plus vite, quelles que soient les conditions lumineuses. Pour comparaison, la figure 5.12 présente des images provenant de deux jeux vidéo différents, avec et sans rendu HDR.

4.1.3 Imagerie HDR par photographie pour la synthèse d’images

L’imagerie HDR par photographie est également utilisée pour l’illumination photoréaliste d’objets virtuels. L’image HDR est dans ce cas là omnidirectionnelle (*ie.* un panoramique de 360°), et représente l’environnement entourant les objets à illuminer. Elle est fabriquée à partir de plusieurs prises de vue d’une petite sphère sur laquelle l’environnement se réfléchit, ce qui permet de reconstruire une image panoramique HDR de l’environnement (ou *light probe*). Des objectifs *fisheye* peuvent également être utilisés à la place de la sphère réfléchissante [Lugand 09].

L’image HDR contient donc les mesures de luminance de l’environnement. Ces données sont ensuite utilisées par un moteur de rendu à lancer de rayons pour illuminer les objets virtuels. Cette méthode s’appelle l’*image-based-lighting*. Elle est très utilisée en photographie, au cinéma et dans des publicités.

4.1.4 Manipulation d’images HDR

Les images HDR peuvent stocker des valeurs de luminances en unités physiques, les candelas par mètres carrés (cd/m^2). Il existe plusieurs formats permettant d’enregistrer ces valeurs en *float*, notamment les formats *.hdr*, *.exr* ou *.tiff*.

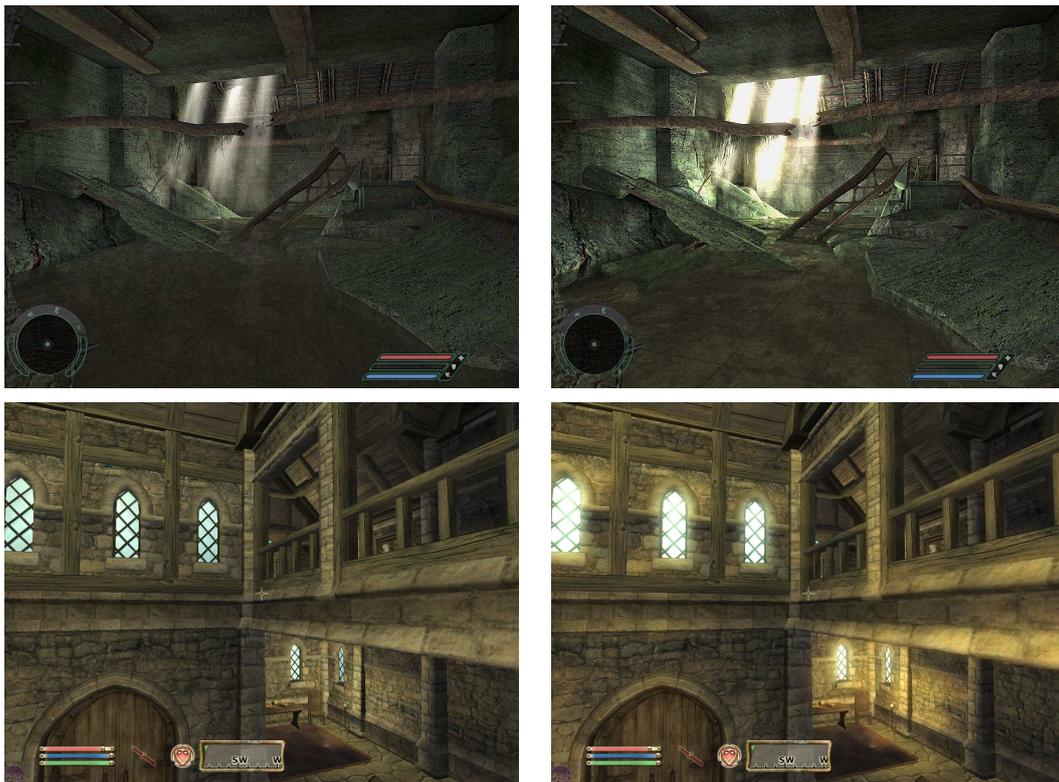


FIGURE 4.2 – Captures d'écran des jeux Far Cry [Crytek 10] et Oblivion [Oblivion 10], sans (à gauche) et avec (à droite) rendu HDR.

Le format `.hdr` en particulier peut contenir dans son *header* une valeur d'exposition, nommée *exposure*. Si elle est présente, elle est à prendre en compte lors de la lecture des données. Les luminances doivent alors être déterminées en fonction des luminances spectrales HDR R , G et B présentes dans le fichier `.hdr`, selon la formule suivante, représentant l'énergie totale d'1 Watt par stéradian par mètre carré dans le spectre visible :

$$(L_R, L_G, L_B) = \frac{179}{exposure}(R, G, B) \quad (4.1)$$

De plus, les algorithmes pour la construction d'images HDR exploitent une partie ou la totalité des informations contenues dans l'entête des photographies, à savoir le temps d'exposition, la sensibilité du capteur, la focale et la balance des blancs. Mais si ces grandeurs ne sont pas toutes exploitées, les valeurs de luminance peuvent être déterminées dans une unité arbitraire, c'est-à-dire à un facteur près par rapport à leurs valeurs réelles en cd/m^2 . Pour les opérateurs de reproduction de tons, cela peut avoir une grande influence sur le résultat affiché à l'écran. Une image de nuit mal calibrée pourrait être lue comme une image avec des valeurs équivalentes à des luminances de

jour... ce qui aura forcément une influence sur le rendu à l'écran de l'image. Cette influence peut être positive si l'on cherche à rendre à l'écran une image avec la meilleure visibilité possible. Mais cette influence peut également être négative si l'on veut rendre à l'écran une image proche de ce que l'on perçoit dans la réalité. Nous devons donc rester vigilants quant aux images à traiter, nous adapter en fonction de la source des images. Le plus prudent reste de vérifier les valeurs manipulées avant et après chaque traitement effectué sur les images HDR.

4.2 Les opérateurs de reproduction de tons

Une fois l'image HDR construite, par photographie ou par synthèse d'images, celle-ci ne peut être directement affichée sur un écran standard. Il est indispensable d'utiliser un opérateur de reproduction de tons pour compresser la dynamique de luminance et pouvoir afficher l'image à l'écran. Mais il est également important de contrôler les modifications effectuées sur l'image HDR, en fonction de l'application visée.

Le but de la reproduction de tons est de reproduire l'apparence visuelle d'une scène sur un écran standard. Ce principe a été décrit par Tumblin et Rushmeier [Tumblin 93], et est illustré sur la figure 4.3. Il existe deux types d'opérateurs de reproduction de tons : les opérateurs globaux et les opérateurs locaux. Un opérateur global applique la même courbe de compression à tous les pixels d'une image, quelle que soit leur position dans l'image. A l'inverse, un opérateur local applique une transformation différente à chaque pixel, qui dépend généralement du voisinage du pixel considéré.

Un autre aspect de la reproduction de tons est le temps. Un opérateur qui dépend du temps est capable de s'adapter aux variations temporelles de luminance. Ce processus est l'adaptation temporelle. Ces opérateurs sont donc *a priori* bien adaptés au traitement d'images dans des flux vidéos, et aux applications interactives.

Nous pouvons également caractériser les différents opérateurs existants par d'autres critères, comme le temps d'exécution, ou leur propriété de pouvoir s'adapter automatiquement au contenu de l'image. La table 4.1 réunit la plupart des opérateurs existants, ainsi que quelques unes de leurs caractéristiques. Le terme « statique » s'oppose à celui de « temporel », et correspond à un opérateur qui n'a pas de modèle de vision dépendant du temps. Le terme « automatique » correspond aux opérateurs qui n'ont pas besoin de réglage spécifique quelle que soit l'image affichée.

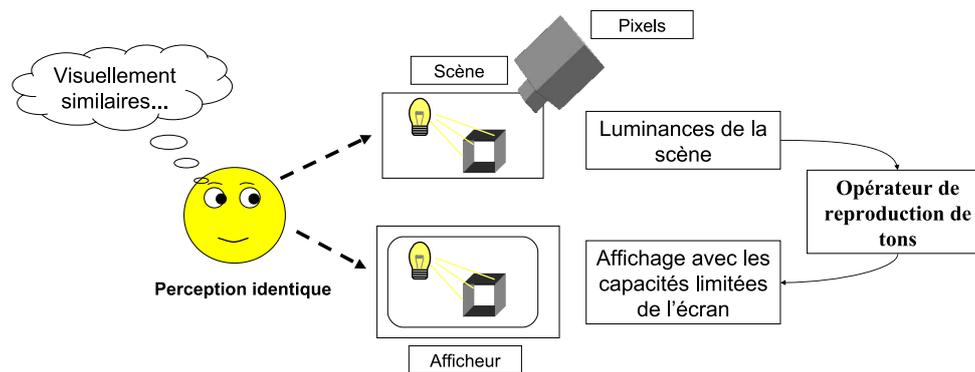


FIGURE 4.3 – L’objectif de la reproduction de tons : avoir une perception identique d’une scène à l’écran et dans la réalité. Adapté d’après [Tumblin 93, Reinhard 05b].

Dans la suite de ce chapitre, nous allons détailler et discuter des performances de différents opérateurs de reproduction de tons de la littérature : d’abord les opérateurs globaux (statiques puis dynamiques), puis les opérateurs locaux (statiques puis dynamiques également). Ensuite, nous présenterons différents protocoles d’évaluations expérimentaux qui permettent d’évaluer et de comparer les performances des opérateurs entre eux.

Il existe plus d’une trentaine d’opérateurs de reproduction de tons référencés dans la littérature. Ne pouvant pas tous les décrire précisément ici, nous nous sommes donc attachés à présenter uniquement ceux utilisés dans les évaluations de la littérature, plus ceux qui nous intéressaient pour le traitement automatique de vidéos HDR, *ie.* les opérateurs temporels.

4.2.1 Symboles et définitions

De manière générale, chaque auteur utilise ses propres notations lorsqu’il propose un nouvel opérateur de reproduction de tons. Pour notre état de l’art, nous avons cependant essayé d’uniformiser les différentes notations selon les symboles suivants :

L_w	luminance de l’image HDR, assimilée à celle de la scène réelle
L_w^{max}	luminance maximale de l’image HDR
L_w^{min}	luminance minimale de l’image HDR
$L_w^{C\hat{o}nes}$	luminance de l’image HDR perçues par les cônes
$L_w^{B\hat{a}tonnets}$	luminance de l’image HDR perçues par les bâtonnets

	Global	Local	Fréquentiel	Statique	Temporel	Automatique
Linéaire	x			x		x
[Chiu 93]		x		x		
[Tumblin 93]	x			x		x
[Schlick 94]	x			x		
[Ward 94]	x			x		x
[Ferwerda 96]	x				x	x
[Rahman 96]		x		x		
[Ward 97]	x			x		x
[Pattanaik 98]		x		x		x
[Durand 00]	x				x	x
[Pattanaik 00]	x				x	
[Durand 02]			x	x		x
[Fairchild 02]		x		x		x
[Fattal 02]			x	x		x
[Reinhard 02]		x		x		x
[Choudhury 03]			x	x		x
[Drago 03]	x			x		x
[Ledda 04]		x			x	x
[Irawan 05]	x				x	x
[Reinhard 05a]	x			x		
[Grave 08]		x		x		x

TABLE 4.1 – Caractéristiques des principaux opérateurs de reproduction de tons

L_{wm}	luminance moyenne de l'image HDR
L_{wa}	luminance d'adaptation de l'observateur devant l'image HDR
$L_{aBâtonnets}$	luminance d'adaptation des bâtonnets de l'observateur devant l'image HDR
$L_{aCônes}$	luminance d'adaptation des cônes de l'observateur devant l'image HDR
B_w	luminosité de l'image HDR
L_d	luminance affichées à l'écran
L_d^{max}	luminance maximale affichable à l'écran
L_d^{min}	luminance minimale affichable à l'écran
L_{da}	luminance d'adaptation de l'observateur devant l'écran
$L_dCônes$	contribution des cônes sur les luminances affichées à l'écran
$L_dBâtonnets$	contribution des bâtonnets sur les luminances affichées à l'écran
B_d	luminosité à l'écran vue par un observateur

$\Delta L(L_a)$ différence de luminance juste perceptible pour un observateur standard adapté à la luminance L_a et regardant un point de luminance L_a

$\Delta L(L_a, L_b)$ différence de luminance juste perceptible pour un observateur standard adapté à la luminance L_a et regardant un point de luminance L_b

4.2.2 Les opérateurs globaux statiques

Les opérateurs globaux appliquent la même courbe de compression à tous les pixels de l'image, à l'inverse des opérateurs locaux. Deux pixels de même luminance L_w seront donc affichés avec la même luminance L_d à l'écran. De plus, le temps d'exécution de ces opérateurs est généralement plus rapide que celui des opérateurs locaux, ce qui les rend intéressants pour des applications temps réel.

De nombreux opérateurs globaux existent dans la littérature. Nous ne pouvons pas tous les présenter entièrement. Cette section détaille donc précisément les opérateurs ayant pour but une visualisation réaliste des images HDR à l'écran, tandis qu'elle ne présente que le fonctionnement des autres opérateurs globaux.

La compression linéaire

La compression linéaire est évidemment l'opérateur de reproduction de tons le plus simple. Elle consiste à compresser uniformément les luminances de l'image, selon la formule suivante :

$$L_d(i, j) = aL_w(i, j) + b \quad (4.2)$$

où

$$a = \frac{L_d^{max} - L_d^{min}}{L_w^{max} - L_w^{min}} \quad (4.3)$$

$$b = \frac{L_d^{min} L_w^{max} - L_d^{max} L_w^{min}}{L_w^{max} - L_w^{min}} \quad (4.4)$$

Chaque contraste étant défini par le rapport entre deux valeurs de luminance, le contraste entre chaque couple de pixels de l'image reste constant au cours de cette compression de la dynamique de luminance. Pourtant, un observateur face à l'écran ne percevra pas l'image de la même manière que dans la réalité, car sa perception des contrastes dépend du niveau lumineux, qui lui varie entre l'écran et la réalité (section 3.2.1). D'autres compressions linéaires sont possibles et obtiennent généralement de meilleurs résultats en ajustant manuellement les valeurs L_w^{min} et L_w^{max} .

L'opérateur de Tumblin et Rushmeier

Tumblin et Rushmeier furent les premiers à poser le problème de la reproduction de tons, comme nous l'avons vu sur la figure 4.3 [Tumblin 93]. Leur objectif est de restituer à l'écran la luminance de la scène observée. Par exemple, une scène de nuit devra produire une image sombre alors qu'une scène de jour produira une image plus lumineuse.

Pour cela, Tumblin et Rushmeier utilisent les travaux de Stevens et Stevens [Stevens 60, Stevens 63] pour trouver une relation entre luminance et luminosité :

$$B = 10^{\beta(L)} L^{\alpha(L)} \quad (4.5)$$

avec

$$\alpha(L) = 0.4 \log_{10}(L) + 2.92 \quad (4.6)$$

$$\beta(L) = -0.4 \log_{10}(L)^2 - 2.584 \log_{10}(L) + 2.0208 \quad (4.7)$$

Tumblin et Rushmeier partent du fait que l'on ne peut pas reproduire à l'écran les véritables valeurs de luminance, la dynamique des images étant généralement supérieure à celle qu'offre un écran standard. Ils proposent en revanche de conserver la luminosité, en ayant une luminosité égale dans la réalité (B_w) et à l'écran (B_d) :

$$B_w = B_d \quad (4.8)$$

d'où, selon l'équation 4.5 :

$$L_d = \frac{(\pi 10^{-4} L_w)^{\frac{\alpha_w}{\alpha_d}}}{(\pi 10^{-4} L_d^{max})^{\frac{\alpha_w}{\alpha_d}}} 10^{\frac{\beta_w - \beta_d}{\alpha_d}}; \quad (4.9)$$

Cet opérateur présente néanmoins certaines limites. La première est que le modèle de Stevens et Stevens n'est valable que pour des scènes simples, et non pour des scènes complexes comme on en rencontre souvent en photographie ou en synthèse d'images. La seconde est que le système visuel humain est capable de s'adapter aux différents niveaux de luminance entre la réalité et l'écran, et est finalement plus sensible aux différences de luminance (donc aux contrastes) qu'à la luminosité. Compresser la dynamique de luminance indépendamment des contrastes n'assure donc pas la similarité des images entre la réalité et l'écran. Enfin, cet opérateur utilise comme unique variable d'entrée la luminance d'adaptation, et y est très sensible. Un brusque changement de cette luminance d'adaptation provoque de grandes changements dans l'apparence de

l'image (notamment son éclaircissement ou son obscurcissement). Le choix de la formule pour calculer cette luminance d'adaptation (voir section 3.2.2) a également une influence. En comparant les images produites avec une moyenne arithmétique et une moyenne logarithmique, nous avons observé que la moyenne logarithmique donnait de meilleurs résultats visuels.

L'opérateur linéaire de Ward

A la différence de l'opérateur précédent, basé sur la conservation de la luminosité, Ward s'intéresse à restituer la visibilité des contrastes dans l'image observée. Pour lui, lors de l'étape de la reproduction de tons, les contrastes visibles dans la scène réelle doivent rester visibles dans l'image affichée, et à l'inverse, les détails invisibles dans l'image réelle ne doivent pas apparaître à l'observateur face à l'écran.

Pour cela, Ward s'appuie sur les travaux menés par Blackwell [Blackwell 46], notamment sur le seuil de détection des contrastes en fonction de la luminance (ou *threshold-versus-intensity function*, voir section 3.2.1). Il propose une compression linéaire de la dynamique de luminance, telle que :

$$L_d = mL_w \quad (4.10)$$

Pour déterminer la constante m , Ward fait deux approximations :

- quelle que soit la luminance réelle L_w , le seuil juste détectable est approximé par $\Delta L(L_{wa})$. Ce terme est effectivement fonction de la luminance d'adaptation réelle L_{wa} , mais il reste une approximation car il devrait également dépendre de la luminance du point fixé.
- quelle que soit la luminance L_d affichée à l'écran, le seuil juste détectable est approximé par $\Delta L(L_d^{max}/2)$. Ici, la luminance d'adaptation devant l'écran est approximée à la moitié de la luminance maximale de l'écran $L_d^{max}/2$. L'approximation est que ce terme devrait normalement dépendre du temps et de l'image affichée, ainsi que de la luminance du point fixé.

Ces approximations permettent de déterminer la constante m , en utilisant la contrainte sur la conservation des contrastes juste détectables. Considérons deux luminances de l'image HDR L_{w1} et L_{w2} , telles que :

$$L_{w2} = L_{w1} + \Delta L(L_{wa}) \quad (4.11)$$

Selon la première approximation, le contraste entre les luminances L_{w1} et L_{w2} est juste détectable. Selon la seconde approximation, ce contraste restera juste détectable

à l'écran si la relation suivante est justifiée :

$$L_{d2} = L_{d1} + \Delta L(L_d^{max}/2) \quad (4.12)$$

Selon l'équation 4.10, la conservation des contrastes est assurée si la constante m est égale au rapport des seuils à l'écran et dans la réalité :

$$m = \frac{\Delta L(L_d^{max}/2)}{\Delta L(L_{wa})} \quad (4.13)$$

Les contrastes juste perceptibles dans l'image HDR sont ainsi également juste perceptibles à l'écran. Tout comme l'opérateur de Tumblin et Rushmeier, cet opérateur est très sensible à la luminance d'adaptation L_{wa} .

L'opérateur de Ward, basé sur l'ajustement d'histogramme

Pour ce nouvel opérateur, Ward part du fait que les écrans standards ont une dynamique de luminance très limitée, bien inférieure à la réalité. Néanmoins, il propose d'exploiter la totalité de la dynamique offerte par ces écrans. L'égalisation d'histogramme est une technique qui permet d'atteindre un tel but, mais qui à le défaut d'augmenter le contraste des images affichées. C'est pourquoi Ward propose plutôt un ajustement d'histogramme, avec un seuil contrôlant les contrastes affichés. Ce seuil, basé sur les valeurs des luminances de l'image HDR et sur la manière dont le système visuel humain les percevrait, permet de préserver les contrastes, évitant ainsi qu'ils ne soient exagérés.

En plus de l'ajustement d'histogramme, Ward propose de simuler avec son opérateur l'éblouissement face aux zones de l'image très lumineuses (voir section 3.2.4), ainsi que la perte de sensibilité aux couleurs et aux contrastes lorsque les luminances sont faibles, tout comme Ferwerda.

La préservation des contrastes est une caractéristique qui nous intéresse tout particulièrement pour la visualisation réaliste des scènes HDR, c'est pourquoi nous présentons en détail cet opérateur.

L'utilisation d'un histogramme a l'avantage de donner une relation directe entre les luminances de l'image HDR L_w et les luminances à afficher L_d . En effet, si nous calculons l'histogramme cumulé P à partir de l'histogramme, et si nous travaillons avec le logarithme de la luminance (pour reproduire la sensibilité de l'œil à la luminosité), l'équation suivante donne une telle relation, en tenant compte des caractéristiques de

l'écran utilisé :

$$\log(L_d) = \log(L_d^{min}) + (\log(L_d^{max}) - \log(L_d^{min})) P(\log(L_w)) \quad (4.14)$$

Il est à noter que l'opérateur de Ward utilise le logarithme de l'image HDR, car l'œil répond à la luminance de manière logarithmique.

Pour déterminer le seuil pour l'ajustement d'histogramme, Ward utilise le principe de la conservation des contrastes. Si l'on considère deux luminances à l'écran L_{d1} et L_{d2} , respectivement deux luminances de l'image HDR L_{w1} et L_{w2} , chaque contraste entre les deux luminances sera juste détectable si :

$$L_{d2} = L_{d1} + \Delta L(L_{da}) \quad (4.15)$$

$$L_{w2} = L_{w1} + \Delta L(L_{wa}) \quad (4.16)$$

Si les luminances L_{d1} et L_{d2} , respectivement L_{w1} et L_{w2} , sont suffisamment proches, alors les contrastes ne seront pas exagérés si :

$$\frac{dL_d}{dL_w} \leq \frac{\Delta L(L_{da})}{\Delta L(L_{wa})} \quad (4.17)$$

Cette inégalité garantit qu'un contraste non visible dans la réalité ne sera pas non plus visible à l'écran (donc qu'il n'y ait pas d'exagération des contrastes). De plus, la conservation des contrastes juste détectables sera assurée s'il y a égalité entre les deux termes. En revanche, cet opérateur ne garantit plus qu'un contraste visible dans la réalité reste visible à l'écran : on pourra éventuellement être confronté à une diminution de la visibilité à l'écran. Ce défaut est dû au choix de la méthode employée, l'ajustement d'histogramme, et est également une conséquence de la volonté initiale d'occuper toute la dynamique de l'écran, même si l'image traitée présente des trous dans son histogramme.

Pour mettre en rapport l'inégalité 4.17 avec les valeurs de l'histogramme de l'image HDR, Ward utilise l'équation 4.14 qui exprime la luminance à afficher L_d en fonction de la luminance de l'image HDR L_w . La dérivée de L_d par rapport à L_w est :

$$\frac{dL_d}{dL_w} = (\log(L_d^{max}) - \log(L_d^{min})) \frac{dP(\log(L_w))}{dL_w} e^{(\log(L_d^{min}) + (\log(L_d^{max}) - \log(L_d^{min})) P(\log(L_w)))} \quad (4.18)$$

l'exponentielle correspondant toujours à la luminance L_d , nous avons donc :

$$\frac{dL_d}{dL_w} = (\log(L_d^{max}) - \log(L_d^{min})) \frac{dP(\log(L_w))}{dL_w} L_d \quad (4.19)$$

Ensuite, pour dériver le terme restant, Ward remarque que l'histogramme d'une image est proportionnel à la dérivée de son histogramme cumulé (voir figure 4.4 pour illustration).

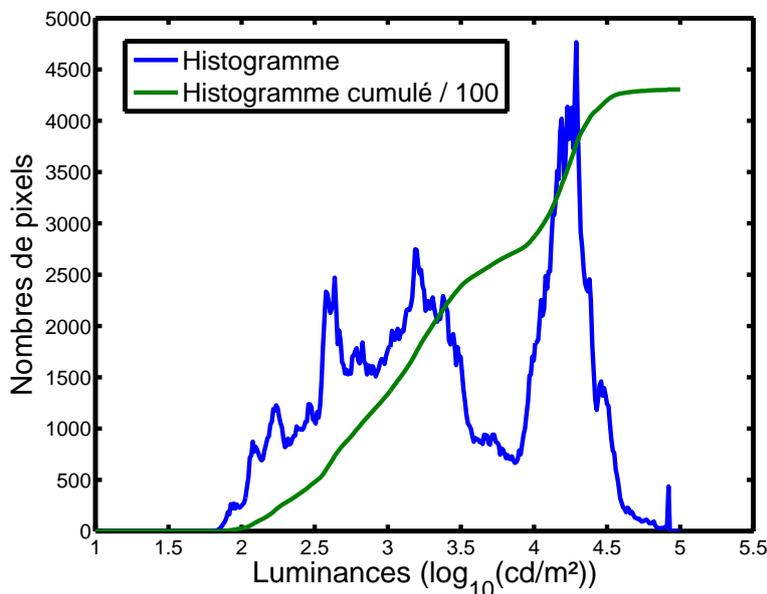


FIGURE 4.4 – Histogramme et histogramme cumulé : l'histogramme est proportionnel à la dérivée de l'histogramme cumulé.

Si nous définissons par $f(\log(L_w))$ le nombre de pixels dans l'intervalle de luminance $L_{w_i} - L_{w_i-1}$, alors pour l'histogramme cumulé normalisé, nous avons $P(\log(L_w))$ qui est la somme des pixels dont la luminance est inférieure à w , divisée par le nombre total de pixels T :

$$P(\log(L_w)) = \frac{1}{T} \sum_{w_i < w} f(\log(L_{w_i})), \quad T = \sum_{w_i} f(\log(L_{w_i})) \quad (4.20)$$

L'expression de l'histogramme comme la dérivée de l'histogramme cumulé normalisé est ensuite :

$$\frac{dP(\log(L_w))}{dL_w} = \frac{1}{T} \frac{f(\log(L_w))}{L_w \Delta b} \quad (4.21)$$

où Δb représente la largeur d'une colonne de l'histogramme :

$$\Delta b = \frac{1}{N} (\log(L_{max}) - \log(L_{min})) \quad (4.22)$$

Les valeurs L_{min} et L_{max} peuvent correspondre aux valeurs minimum et maximum de l'image HDR, mais si l'opérateur est utilisé sur une vidéo HDR, il est préférable

d'utiliser des valeurs constantes, telles les valeurs limites de la perception humaine. Dans ce cas là, nous pourrions prendre $L_{max} = 10^6 cd/m^2$ et $L_{min} = 10^{-4} cd/m^2$.

Avec ce dernier résultat, nous pouvons maintenant exprimer la dérivée de la luminance à l'écran L_d en fonction de la luminance de l'image HDR L_w :

$$\frac{dL_d}{dL_w} = (\log(L_d^{max}) - \log(L_d^{min})) \frac{f(\log(L_w))}{T\Delta b} \frac{L_d}{L_w} \quad (4.23)$$

Et enfin, en utilisant l'inégalité 4.17, nous trouvons la relation qui contraint la distribution de l'histogramme $f(\log(L_w))$:

$$f(\log(L_w)) \leq \frac{\Delta L(L_{da})}{\Delta L(L_{wa})} \frac{T\Delta b}{\log(L_d^{max}) - \log(L_d^{min})} \frac{L_w}{L_d} \quad (4.24)$$

Pour la dernière étape de son opérateur, Ward propose d'ajuster l'histogramme par itérations, en tronquant les pixels qui dépassent le seuil autorisé, puis en recalculant un nouveau seuil en fonction des pixels restants. L'opération est répétée jusqu'à ce que le nombre de pixels dépassant le seuil devienne inférieur à une tolérance donnée.

Les résultats de cet opérateur de reproduction de tons donnent de bonnes performances sur une grande variété d'images, le contraste, la luminosité et les couleurs étant préservées à l'affichage. Il est parfois reproché à cet opérateur de faire apparaître des bandes de niveaux lorsqu'il y a des dégradés de luminance sur l'image. Mais il suffit d'augmenter le nombre d'échantillons de l'histogramme pour résoudre ce problème.

Enfin, cet opérateur est souvent noté comme l'un des meilleurs opérateurs lors des évaluations (voir section 4.3). Cependant, une approximation reste le calcul des contrastes juste détectables. Ward sous-estime ici les seuils juste détectables, car sa modélisation revient à dire que l'œil est adapté à chaque pixel qu'il regarde, ce qui n'est pas vrai dans la réalité, l'œil s'adaptant constamment.

L'opérateur de Schlick

Schlick [Schlick 94] propose un opérateur de forme sigmoïde, défini par l'équation suivante :

$$L_d = \frac{pL_w}{(p-1)L_w + L_w^{max}}, \quad p \in [1, \infty) \quad (4.25)$$

Le paramètre p permet de régler les contrastes ; il est fixé de manière à ce que la plus petite valeur non-noire reste visible après le traitement. Un réglage manuel est généralement nécessaire pour optimiser les résultats de cet opérateur.

L'opérateur de Drago et al.

Drago et al. proposent une compression logarithmique des luminances, pour imiter la réponse de l'œil à la lumière [Drago 03]. La base b du logarithme varie entre 2 et 10, permettant la préservation des contrastes et des détails de l'image HDR :

$$L_d = \frac{\log_b(1 + L_w)}{\log_b(1 + L_w^{max})} \quad (4.26)$$

Enfin, Drago propose une nouvelle correction du gamma des écrans, afin de mieux préserver les contrastes là où les luminances sont faibles.

L'opérateur global de Reinhard et Devlin

Reinhard et Devlin [Reinhard 05a] proposent un opérateur basé sur un modèle de photorécepteur de forme sigmoïde, assez proche du modèle physiologique de Naka et Rushton [Naka 66] :

$$V = V_{max} \frac{L_w}{L_w + \sigma(L_{wa})} \quad (4.27)$$

où V est la réponse du photorécepteur, et σ un terme de demi-saturation, fonction de la luminance moyenne $\sigma(L_{wa})$.

Trois paramètres, définis par l'utilisateur, permettent de contrôler le contraste de l'image : son apparence lumineuse (claire ou sombre), les adaptation chromatique et temporelle. Ces paramètres peuvent être réglés par défaut, mais les meilleurs résultats seront obtenus avec un réglage manuel.

4.2.3 Les opérateurs globaux dynamiques

Les opérateurs dynamiques sont particulièrement intéressants pour le traitement de vidéos, car ils assurent une cohérence temporelle entre les images successives. C'est pourquoi nous les détaillons à part, dans la section suivante.

L'opérateur de Ferwerda et al.

De la même manière que les opérateurs de Ward, l'opérateur de Ferwerda et al. utilise également la conservation des contrastes juste détectables [Ferwerda 96]. Mais alors que les opérateurs linéaires de Tumblin et Ward ne compressent que la dynamique de luminance d'images en noir et blanc, l'opérateur de Ferwerda apporte plusieurs améliorations.

D'abord, il permet de traiter des images en couleurs. Il utilise pour cela l'espace de couleur CIE XYZ, pour calculer la luminance réelle $L_w = Y$ à partir des couleurs de l'image HDR (voir section 2.6).

Mais surtout, Ferwerda intègre un modèle de vision à son opérateur. Il simule la vision des cônes et des bâtonnets, l'acuité visuelle, la perte de sensibilité aux couleurs en vision scotopique, et propose enfin un modèle simple d'adaptation temporelle. Ces simulations se font dans un ordre bien précis. On commence par l'acuité visuelle (avant de transformer l'image de couleurs en luminances), si l'on se situe dans le domaine scotopique. On supprime alors les fréquences $\omega_c(L_{wa})$ supérieures au seuil que peut voir l'observateur, en passant dans l'espace de Fourier f^* .

$$f^*(\omega_c(L_{wa})) = \frac{\Delta L(L_{wa})}{L_{wa}} \quad (4.28)$$

Les luminances des couleurs à afficher sont ensuite calculées selon les équations 4.30 à 4.35. Pour les cônes, m est la même constante que pour l'opérateur de Ward, au choix de la luminance réelle d'adaptation près, alors que la constante b simule le rôle des bâtonnets :

$$m = \frac{\Delta L_{C\hat{c}ones}(L_d^{max}/2)}{\Delta L_{C\hat{c}ones}(L_w^{max}/2)}, \quad b = \frac{\Delta L_{B\hat{a}tonnets}(L_d^{max}/2)}{\Delta L_{B\hat{a}tonnets}(L_w^{max}/2)} \quad (4.29)$$

La constante b permet d'améliorer l'adaptation aux images contenant de faibles luminances, puisque l'opérateur contrôle ainsi la luminosité et le contraste séparément. Quant à elle, la constante k varie de 0 à 1 en fonction des luminances. Elle est nulle dans le domaine photopique (faisant disparaître le terme scotopique), égale à 1 dans le domaine scotopique (les cônes sont alors aveugles, et ne contribuent pas à la vision), et de valeur intermédiaire dans le domaine mésopique, en fonction de la luminance d'adaptation réelle L_{wa} .

$$L_{dR} = mL_{wR} + bL_w \quad (4.30)$$

$$L_{dG} = mL_{wG} + bL_w \quad (4.31)$$

$$L_{dB} = mL_{wB} + bL_w \quad (4.32)$$

$$L_{dC\hat{c}ones} = (L_{dR}, L_{dG}, L_{dB}) \quad (4.33)$$

$$L_{dB\hat{a}tonnets} = (m + b)L_w \quad (4.34)$$

$$L_d = L_{dC\hat{c}ones} + kL_{dB\hat{a}tonnets} \quad (4.35)$$

Cet opérateur est le premier à traiter des images en couleur, à prendre en compte l'acuité visuelle, ainsi que les influences distinctes des cônes et des bâtonnets. Il propose également un modèle simple d'adaptation visuelle, qui sera améliorée par Pattanaik [Pattanaik 00]. Par rapport à la vision des contrastes, Ferwerda donne les valeurs des seuils juste détectables des cônes et des bâtonnets, à utiliser dans la fonction 4.29. Ces valeurs seront réutilisés dans plusieurs autres opérateurs ([Ward 97, Irawan 05]) :

Pour les domaines photopique et mésopique :

$$\log_{10}(\Delta L(L_a)) = \begin{cases} -0.72 & \text{si } \log_{10}(L_a) \leq -2.6 \\ \log_{10}(L_a) - 1.255 & \text{si } \log_{10}(L_a) \geq 1.9 \\ (0.249 \log_{10}(L_a) + 0.65)^{2.7} - 0.72 & \text{sinon} \end{cases} \quad (4.36)$$

Pour les domaines scotopique et mésopique :

$$\log_{10}(\Delta L(L_a)) = \begin{cases} -2.86 & \text{si } \log_{10}(L_a) \leq -3.94 \\ \log_{10}(L_a) - 0.395 & \text{si } \log_{10}(L_a) \geq -1.44 \\ (0.405 \log_{10}(L_a) + 1.6)^{2.18} - 2.86 & \text{sinon} \end{cases} \quad (4.37)$$

Alors que les précédents opérateurs étaient très sensibles à la luminance d'adaptation, l'opérateur de Ferwerda l'est beaucoup moins, grâce à l'ajout de la constante b dans l'opérateur. L'ajout de traitements sur les couleurs, de l'acuité, et de la simulation des bâtonnets et des cônes font de cet opérateur le premier à réellement simuler la vision humaine.

L'opérateur global de Durand et Dorsey

Durand et Dorsey publient un opérateur de reproduction de tons interactif, intégrant des simulations de l'éblouissement, de l'adaptation à la lumière et à l'obscurité, et de l'adaptation chromatique [Durand 00]. Les auteurs réutilisent la compression linéaire proposée par Ward [Ward 94], et améliorent l'opérateur de Ferwerda et *al.* [Ferwerda 96] en proposant de nouveaux paramètres. Durand et Dorsey présentent enfin comment intégrer leur opérateur dans une application temps réel, pour la visualisation interactive d'images HDR.

L'opérateur temporel de Pattanaik et *al.*

Egalement en 2000, Pattanaik et *al.* proposent le premier opérateur automatique, avec une adaptation visuelle dépendante de l'image à afficher, mais aussi de l'état anté-

rieur dans lequel se trouve l'œil [Pattanaik 00]. La description complète de l'opérateur est représentée sur la figure 4.5. Cet opérateur utilise quatre étapes successives :

- un modèle d'adaptation visuelle, qui calcule ce que voit la rétine ;
- un modèle d'apparence, qui calcule ce que perçoit la rétine ;
- un modèle inverse d'apparence, ou modèle inverse d'écran ;
- un inverse d'adaptation, qui permet de simuler l'adaptation de l'œil de l'observateur par rapport à ce qui est affiché à l'écran.

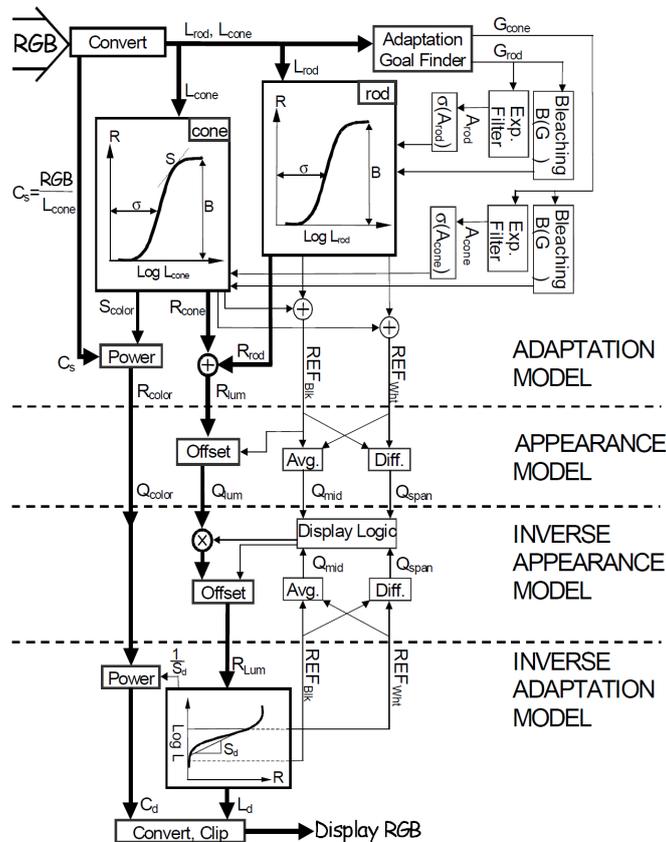


FIGURE 4.5 – Diagramme de l'opérateur de reproduction de tons de Pattanaik et al. [Pattanaik 00].

Le modèle d'adaptation visuelle de cet opérateur est la contribution la plus importante de cet opérateur. Nous pensons que ce modèle est très important pour les traitements de vidéos où de forts changements de luminosité interviennent, c'est pourquoi nous détaillons ce modèle. Pattanaik utilise le modèle d'apparence colorée de Hunt [Hunt 95] pour simuler l'adaptation aux variations de luminances. Les cônes et

les bâtonnets de nos yeux ne peuvent percevoir en un instant qu'une faible domaine de l'ensemble de la gamme des luminances existantes. Mais ce domaine s'adapte continuellement pour s'accoutumer aux luminances qui lui sont présentées. Si σ est la valeur d'adaptation du système visuel humain, alors le modèle de Hunt donne la courbe de réponse neurale en fonction de la luminance L_w de l'image HDR :

$$R(L_w) = B \frac{L_w^n}{L_w^n + \sigma^n}, \quad n = 0.73 \quad (4.38)$$

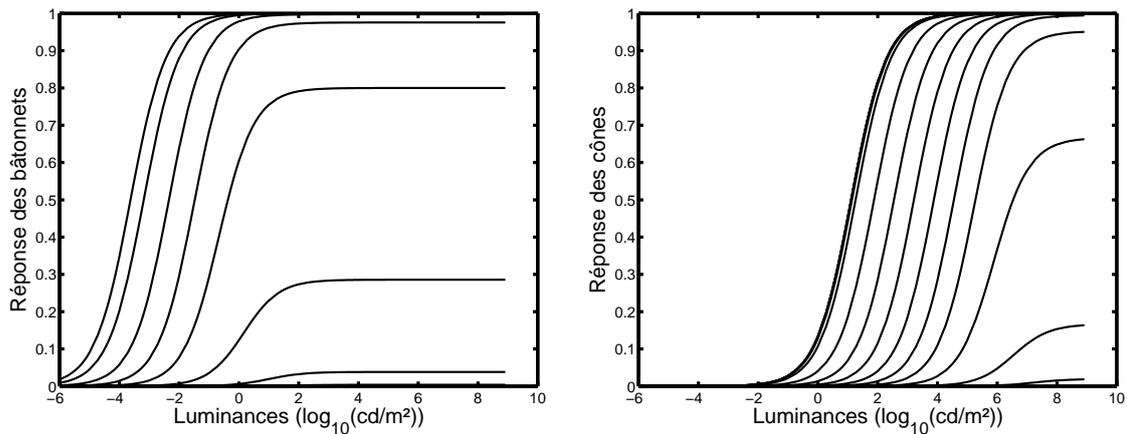


FIGURE 4.6 – Réponse des cônes et des bâtonnets en fonction de la luminance et de leur état d'adaptation, selon le modèle de Pattanaik [Pattanaik 00]. Pour les bâtonnets, les luminances d'adaptation sont $L_a = 10^{-6}, 10^{-5}, \dots, 10^0$ et 10^1 cd/m^2 (de gauche à droite). Pour les cônes, les luminances d'adaptation sont $L_a = 10^{-6}, 10^{-5}, \dots, 10^8 \text{ cd/m}^2$ (également de gauche à droite, mais les courbes de luminances d'adaptation inférieure à 10^{-1} cd/m^2 sont superposées, car les cônes ne sont pas sensibles aux faibles luminances).

La figure 4.6 présente les courbes de réponse issues de ce modèle, pour différentes luminances d'adaptation. Nous y retrouvons plusieurs propriétés décrites dans le chapitre 3. Pour les bâtonnets, la réponse est maximale lorsque les luminances sont faibles (inférieures à 0.5 cd/m^2), mais ils restent actifs jusqu'à 10 cd/m^2 (frontière mésopique - photopique), leurs réponses étant nulles au-delà. Pour les cônes, nous retrouvons une activité à partir de 10^{-2} cd/m^2 (frontière scotopique - mésopique), une sensibilité maximale à partir de 10 cd/m^2 , et une saturation pour les luminances très élevées, comme celle du soleil.

Les bâtonnets et les cônes obéissent indépendamment à l'équation 4.38. Les termes B et σ sont dynamiques, et modélisent l'adaptation visuelle en fonction des luminances

vues à l'instant t , mais aussi de l'état antérieur du système visuel humain. Le terme B modélise l'adaptation des bâtonnets et des cônes, et le terme σ l'adaptation neurale.

Pour simuler les deux processus d'adaptation, Pattanaik propose d'utiliser un filtre exponentiel discret. L'adaptation est discrétisée en différents pas de temps, chaque pas correspondant à l'apparition d'une image par seconde ($T = \frac{1}{60}$ seconde pour une fréquence de 60 Hertz). L'adaptation ne s'exprime plus alors qu'en fonction des temps de réponse de chaque processus.

L'adaptation neurale est un processus temporel, que Pattanaik propose de modéliser à partir des luminances d'adaptation, respectivement $L_{aBât.}$ et $L_{aCônes}$. Ces luminances d'adaptation dépendent des luminances moyennes perçues par les bâtonnets et les cônes à l'instant présent, mais aussi de l'état antérieur dans lequel se trouve son œil. Les luminances moyennes sont donc approximées par la moyenne logarithmique des luminances de l'image HDR à l'instant t , et est respectivement notée par les termes $L_{mBât.}(t)$ et $L_{mCônes}(t)$. Pour comparaison, ce terme est considéré comme le seul terme d'adaptation dans les opérateurs de reproduction de tons où il n'y a pas de modèle d'adaptation visuelle dynamique. Enfin, les termes se référant à l'état antérieur de l'œil, et concrètement à l'état de l'œil à l'image précédente, sont $L_{aBât.}(t-1)$ et $L_{aCônes}(t-1)$.

Pattanaik propose donc les filtres suivants pour tenir compte des deux facteurs d'adaptation, avec les constantes de temps $t_{0Bât.} = 0.15s$ et $t_{0Cônes} = 0.08s$:

$$L_{aBât.}(t) = L_{aBât.}(t-1)e^{\frac{-T}{t_{0Bât.}}} + \left(1 - e^{\frac{-T}{t_{0Bât.}}}\right)L_{mBât.}(t) \quad (4.39)$$

$$L_{aCônes}(t) = L_{aCônes}(t-1)e^{\frac{-T}{t_{0Cônes}}} + \left(1 - e^{\frac{-T}{t_{0Cônes}}}\right)L_{mCônes}(t) \quad (4.40)$$

Suivant le modèle de Hunt [Hunt 95], Pattanaik donne les formules suivantes pour le calcul du terme σ , relatif à l'adaptation neurale :

$$\sigma_{bât.}(t) = \frac{2^{1/0.73}L_{aBât.}(t)}{19000j^2L_{aBât.}(t) + 0.2615(1-j^2)^4L_{aBât.}(t)^{1/6}} \quad (4.41)$$

$$\sigma_{cônes}(t) = \frac{12.9223L_{aCônes}(t)}{k^4L_{aCônes}(t) + 0.171(1-k^4)^2L_{aCônes}(t)^{1/3}} \quad (4.42)$$

Les valeurs j et k étant définies par :

$$j(t) = \frac{1}{5.10^5L_{aBât.}(t) + 1}, \quad k(t) = \frac{1}{5L_{aCônes}(t) + 1} \quad (4.43)$$

L'adaptation due aux bâtonnets et aux cônes est un processus plus lent que l'adaptation neurale. Nous avons vu à la section 3.2.3 que lorsqu'un photon frappe un bâtonnet ou un cône, il est absorbé quasi-instantanément, mais le photorécepteur utilisé ne pourra pas absorber un nouveau photon avant qu'il ne se régénère.

Pattanaik modélise l'évolution du blanchiment et de la régénération des capteurs en fonction des luminances d'adaptation $L_{aBât.}$ et $L_{aCônes}$, données par les équations 4.39 et 4.40. Il utilise les grandeurs intermédiaires suivantes : B modélise la quantité de pigments disponibles pour capter un photon, la fonction $J(\dots)$ décrit le blanchiment des pigments, et la fonction $K(\dots)$ leur régénération. L'ensemble de la modélisation est illustrée sur la figure 4.7.

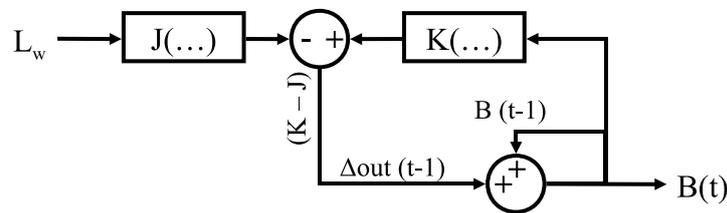


FIGURE 4.7 – Modélisation du blanchiment par la lumière et de la régénération des bâtonnets et des cônes, en fonction du temps.

La concentration de pigments prêts à capter un photon est normalisée entre 0 et 1. Elle est mise à jour en fonction des photorécepteurs blanchis et régénérés, selon les équations suivantes :

$$B_{Bât.}(t) = B_{Bât.}(t-1) + K_{Bât.}(t-1) - J_{Bât.}(t-1) \quad (4.44)$$

$$B_{Cônes}(t) = B_{Cônes}(t-1) + K_{Cônes}(t-1) - J_{Cônes}(t-1) \quad (4.45)$$

La concentration de pigments blanchis par la lumière est déterminée en fonction de la concentration des pigments disponibles et de la luminance incidente :

$$J_{Bât.}(t) = \frac{T}{16} B_{Bât.}(t-1) L_{aBât.}(t) \quad (4.46)$$

$$J_{Cônes}(t) = \frac{T}{2.2 \cdot 10^8} B_{Cônes}(t-1) L_{aCônes}(t) \quad (4.47)$$

Enfin, la concentration de pigments régénérés au cours du temps est donnée par les équations suivantes (les constantes de temps correspondantes sont $\tau_{Bât.} = 400s$ et $\tau_{Cônes} = 110s$, [Hollins 73, Alpern 97]) :

$$K_{Bât.}(t) = \frac{T}{\tau_{Bât.}} (1 - B_{Bât.}(t-1)) \quad (4.48)$$

$$K_{Cônes}(t) = \frac{T}{\tau_{Cônes}} (1 - B_{Cônes}(t-1)) \quad (4.49)$$

Nous pouvons maintenant voir que cette modélisation du blanchiment et de la régénération des capteurs simule une adaptation temporelle. A l'équilibre face à une luminance donnée, le système visuel humain est adapté, la concentration des photorécepteurs actifs est constante, et la différence $K - J$ nulle. Si la luminance L_w augmente, alors le blanchiment des capteurs va s'accélérer (fonction $J(\dots)$), et la concentration des pigments B va diminuer. La réponse des capteurs sera alors plus faible, puisque le terme B contrôle le gain appliqué à la réponse neurale (équation 4.38). Ensuite, le terme B étant compris entre 0 et 1, une chute de la concentration va augmenter le taux de cellules régénérées (fonction $K(\dots)$). Ce taux de régénération va augmenter jusqu'à ce qu'on retrouve un équilibre entre le nombre de capteurs blanchis et régénérés, ce qui correspondra à un nouvel état d'adaptation pour l'observateur.

Dans le modèle d'adaptation de son opérateur de reproduction de tons, Pattanaik simule pour la première fois l'adaptation du système visuel humain au cours du temps, en fonction des luminances affichées et de l'état antérieur de l'œil. C'est pour nous la principale contribution de cet opérateur. La suite du traitement consiste à afficher les réponses des bâtonnets et des cônes générées par l'équation 4.38, en trois étapes successives.

Tout d'abord, un modèle d'apparence définit une référence blanche et une référence noire, pour traduire l'impression visuelle de l'observateur en fonction de la clarté ou de l'obscurité de l'image HDR. Ensuite, un modèle d'apparence inverse utilise les luminances minimales et maximales que peut afficher l'écran d'observation pour les faire correspondre aux références HDR. Enfin, un modèle inverse d'adaptation compresse la dynamique de luminance et traite les couleurs dans un canal séparé, avant d'obtenir les luminances à afficher à l'écran.

L'opérateur d'Irawan et al.

Irawan et *al.* proposent un opérateur basé sur le modèle d'adaptation visuelle de Pattanaik [Pattanaik 00] et sur l'ajustement d'histogramme de Ward [Ward 97]. Irawan utilise la fonction standard des seuils juste détectables en fonction de l'intensité (*threshold-versus-intensity function*) pour proposer de nouvelles valeurs de seuils juste détectables, en fonction de l'intensité et de l'état d'adaptation de l'observateur. L'opérateur d'Irawan nous intéresse tout particulièrement, car il est le seul opérateur à être dynamique, automatique, et à simuler la perception du système visuel humain au niveau des contrastes, et donc de la visibilité. Cet opérateur est donc décrit très en détail

dans cette section. Il peut être décomposé en 3 étapes :

- le calcul de l'état d'adaptation du système visuel humain, en utilisant un modèle de vision dynamique ;
- le calcul des seuils de luminance juste détectables, pour l'état d'adaptation déterminé à l'étape précédente, et pour chaque luminance observée ;
- l'ajustement d'histogramme, effectué avec une contrainte pour ne pas exagérer les contrastes.

Le modèle d'adaptation mis en œuvre est proche de celui proposé par Pattanaik. Irawan réutilise le modèle de Naka-Rushton (équation 4.50, [Naka 66]) pour modéliser la réponse des photorécepteurs en fonction de la luminance L . A la différence de Pattanaik, l'influence du blanchiment et de la régénération des photorécepteurs n'est plus exprimé par un gain, mais est compris dans le terme $\sigma(L_a)$, qui modélise l'état d'adaptation du système visuel humain.

$$R(L, \sigma(L_a)) = \frac{L^n}{L^n + \sigma(L_a)^n}, \quad n = 0.73 \quad (4.50)$$

Le calcul de l'état du système visuel humain prend en compte trois phénomènes d'adaptation : le blanchiment et la régénération des photorécepteurs, et deux adaptations neurales, l'une rapide, l'autre lente. La prise en compte de l'adaptation neurale lente est un ajout par rapport au modèle de Pattanaik.

L'état d'adaptation du système visuel humain $\sigma(L_a)$ est calculé en fonction des trois processus d'adaptation simulés, selon un produit de trois termes. Le terme représentant le blanchiment et la régénération des photorécepteurs sera noté σ_B , le terme correspondant à l'adaptation neurale rapide sera σ_N , et celui correspondant à l'adaptation neurale lente sera σ_C .

Le calcul de ces trois termes nécessite d'abord d'extraire la luminance et la chrominance de l'image HDR, en passant par les espaces de couleur XYZ, puis xyY . La chrominance correspond aux canaux x et y , tandis que la luminance Y sera traitée indépendamment (voir section 2.6). Les images de luminance des bâtonnets et des cônes sont obtenues en utilisant les valeurs X , Y et Z de l'espace de couleur CIE XYZ. [Ward 97] :

$$L_{Bât.} = Y \left(1.33 \left(1 + \frac{Y+Z}{X} \right) - 1.68 \right), \quad L_{Cônes} = Y \quad (4.51)$$

Les luminances moyennes des cônes et des bâtonnets, $L_{mBât.}$ et $L_{mCônes}$, sont déterminées en utilisant une moyenne logarithmique sur ces images de luminance, respectivement $L_{Bât.}$ et $L_{Cônes}$.

Pour simuler le blanchiment et la régénération des photorécepteurs, Irawan définit les fractions moyennes de photorécepteurs prêts à recevoir un photon $P_{mBât.}$ et $P_{mCônes}$, en fonction des luminances incidentes moyennes $L_{mBât.}$ et $L_{mCônes}$, dans le cas où l'œil est adapté à ces luminances [Hood 86] :

$$P_{mBât.}(L_{mBât.}(t)) = \frac{10^4}{10^4 + L_{mBât.}(t)} \quad (4.52)$$

$$P_{mCônes}(L_{mCônes}(t)) = \frac{10^4}{10^4 + L_{mCônes}(t)} \quad (4.53)$$

Dans la réalité, les luminances moyennes $L_{mBât.}$ et $L_{mCônes}$ vues par les photorécepteurs ne sont pas constantes au cours du temps. Pour tenir compte de leurs variations, Irawan définit les fractions de photorécepteurs $P_{Bât.}(t)$ et $P_{Cônes}(t)$, fonction des fractions de l'instant précédent ($P_{Bât.}(t-1)$ et $P_{Cônes}(t-1)$) et des fractions moyennes de photorécepteurs à l'instant présent si l'œil était totalement adapté ($P_{mBât.}(t)$ et $P_{mCônes}(t)$). Les filtres suivants utilisent les constantes de temps $t_{0Bât.} = 110s$ et $t_{0Cônes} = 400s$:

$$P_{Bât.}(t) = e^{\frac{-T}{t_{0Bât.} P_{mBât.}(L_{mBât.}(t))}} P_{Bât.}(t-1) + \left(1 - e^{\frac{-T}{t_{0Bât.} P_{mBât.}(L_{mBât.}(t))}}\right) P_{mBât.}(L_{mBât.}(t)) \quad (4.54)$$

$$P_{Cônes}(t) = e^{\frac{-T}{t_{0Cônes} P_{mCônes}(L_{mCônes}(t))}} P_{Cônes}(t-1) + \left(1 - e^{\frac{-T}{t_{0Cônes} P_{mCônes}(L_{mCônes}(t))}}\right) P_{mCônes}(L_{mCônes}(t)) \quad (4.55)$$

Irawan fait ensuite l'hypothèse que le signal transmis par les photorécepteurs est proportionnel à $L \times P(L)$. De plus, comme il est équivalent de diviser la luminance par un facteur ou de déplacer la courbe sigmoïde vers la droite selon le même facteur, Irawan définit l'état des photorécepteurs en fonction du temps comme l'inverse de la fraction des photorécepteurs disponibles :

$$\sigma_{BBât.}(t) = \left(\frac{1}{P_{Bât.}(t)}\right), \quad \sigma_{BCônes}(t) = \left(\frac{1}{P_{Cônes}(t)}\right) \quad (4.56)$$

Le premier des trois termes de l'adaptation visuelle σ_B étant maintenant connu pour chaque photorécepteur, nous pouvons désormais calculer les deux autres, correspondant aux adaptations neurales rapides et lentes, σ_N et σ_C .

Pour simuler les deux processus d'adaptation neurale, nous avons besoin de déterminer la luminance d'adaptation, calculée en utilisant un filtre exponentiel, de la même manière que l'opérateur de Pattanaik (voir équations 4.39 et 4.40 et [Pattanaik 00]), en fonction de la luminance d'adaptation de l'instant précédent et de la luminance moyenne de l'instant présent. Les constantes de temps des deux processus sont néanmoins différentes : pour l'adaptation neurale rapide, $t_{0Bât.} = 0.15s$ et $t_{0Cônes} = 0.8s$; pour l'adaptation neurale lente, $t_{0Bât.} = 40s$ et $t_{0Cônes} = 10s$.

$$L_{aBât.}(t) = e^{\frac{-T}{t_{0Bât.}}} L_{aBât.}(t-1) + (1 - e^{\frac{-T}{t_{0Bât.}}}) L_{mBât.}(t) \quad (4.57)$$

$$L_{aCônes}(t) = e^{\frac{-T}{t_{0Cônes}}} L_{aCônes}(t-1) + (1 - e^{\frac{-T}{t_{0Cônes}}}) L_{mCônes}(t) \quad (4.58)$$

Chacune des luminances d'adaptation est ensuite utilisée pour calculer l'état de chaque processus d'adaptation neural. Pour le processus rapide, nous avons :

$$\log_{10}(\sigma_{NBât.}(t)) = \frac{2.311L_{aBât.}(t)^{0.3604}}{L_{aBât.}(t)^{0.3604} + 0.008061^{0.3604}} - 2.749 \quad (4.59)$$

$$\log_{10}(\sigma_{NCônes}(t)) = \frac{2.027L_{aCônes}(t)^{0.6406}}{L_{aCônes}(t)^{0.6406} + 5.859^{0.6406}} + 0.01711 \quad (4.60)$$

Pour le processus lent, nous avons :

$$\log_{10}(\sigma_{CBât.}(t)) = \frac{1.735L_{aBât.}(t)^{0.9524}}{L_{aBât.}(t)^{0.9524} + 1.277^{0.9524}} + 0.005684 \quad (4.61)$$

$$\log_{10}(\sigma_{CCônes}(t)) = \frac{1.929L_{aCônes}(t)^{0.8471}}{L_{aCônes}(t)^{0.8471} + 1048^{0.8471}} + 0.01820 \quad (4.62)$$

Maintenant que les trois termes des processus d'adaptation sont calculés, nous pouvons déterminer l'état d'adaptation des bâtonnets et des cônes en fonction du temps, face à l'image HDR entière. Nous notons ces termes $\sigma_{BâtonnetsGlobal}$ et $\sigma_{CônesGlobal}$.

$$\sigma_{BâtonnetsGlobal}(t) = \sigma_{BBât.}(t) \sigma_{CBât.}(t) \sigma_{NBât.}(t) \quad (4.63)$$

$$\sigma_{CônesGlobal}(t) = \sigma_{BCônes}(t) \sigma_{CCônes}(t) \sigma_{NCônes}(t) \quad (4.64)$$

Ces états d'adaptation globaux correspondent à l'adaptation face à l'image entière. Mais en réalité, un observateur regardant une image a son regard qui parcourt successivement plusieurs zones de l'image, par saccades. On estime que chaque zone est fixée approximativement pendant 200 à 300 millisecondes. Lors de chaque fixation, l'œil s'adapte localement à la luminance de la zone de l'image qu'il observe. Si l'on ne veut pas surestimer les seuils juste détectables pour chaque luminance de l'image, il faut

donc tenir compte de cette adaptation locale à partir de l'adaptation globale déterminée précédemment, en réutilisant les formules des processus d'adaptation précédentes. Irawan et *al.* proposent d'utiliser un temps d'adaptation de $0.33s$ pour simuler l'adaptation locale due aux mouvements de saccades, à partir de l'état d'adaptation global. Ce temps étant court, ce sont les processus d'adaptation neuraux qui participent principalement à cette adaptation locale.

Cette dernière étape du modèle de vision dynamique d'Irawan permet d'obtenir les valeurs de $\sigma_{B\hat{a}ttonnetsLocal}(t)$ et $\sigma_{C\hat{o}nesLocal}(t)$, que nous noterons plus simplement $\sigma(L_a)$ par la suite. Les étapes suivantes sont néanmoins à effectuer séparément pour les bâtonnets et les cônes.

L'étape suivante du modèle d'Irawan est le calcul des seuils juste détectables, en fonction des luminances regardées mais aussi de l'état d'adaptation déterminé par le modèle de vision dynamique (*threshold versus intensity and adaptation (TVIA) function*). A partir cet état d'adaptation, nous pouvons déterminer quelle sera la réponse des deux types de photorécepteurs face à une luminance donnée, pour un état d'adaptation donné. Elle est définie par l'équation 4.50. La figure 4.8 illustre la courbe de réponse produite, de forme sigmoïde et centrée sur la valeur de l'état d'adaptation.

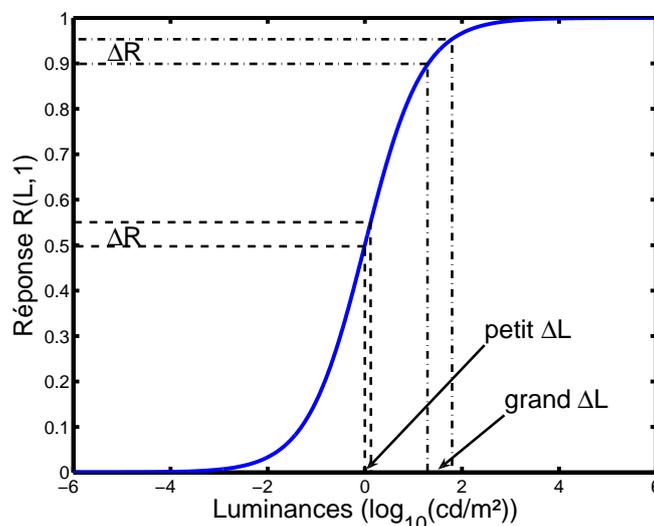


FIGURE 4.8 – Courbe de réponse des photorécepteurs, pour $\sigma = 1$.

Pour une même différence de réponse ΔR en ordonnée, la différence de luminance nécessaire est plus faible si les luminances se trouvent proches de l'état d'adaptation (petit ΔL) que si elles en sont éloignées (grand ΔL , voir figure 4.8). Pour déterminer

une valeur de seuil juste détectable, il est donc très important de connaître la luminance regardée et l'état d'adaptation du système visuel. L'opérateur d'Irawan est le seul, à notre connaissance, à prendre en compte ces deux paramètres. Cependant, il n'existe pas dans la littérature de mesure de seuil de luminance pour l'observation d'une luminance quelconque et pour un niveau d'adaptation donné. Irawan propose d'obtenir ces valeurs de seuil par le calcul.

Soient $\Delta L(L, \sigma(L_a))$ la valeur des seuils juste détectables en fonction de la luminance observée L et de l'état d'adaptation $\sigma(L_a)$, et ΔR la réponse additionnelle produite en ajoutant le seuil juste détectable à la luminance initiale L . Selon l'équation 4.50, ces deux grandeurs sont reliées par la relation suivante :

$$R(L, \sigma(L_a)) + \Delta R = \frac{(L + \Delta L(L, \sigma(L_a)))^n}{(L + \Delta L(L, \sigma(L_a)))^n + \sigma(L_a)^n} \quad (4.65)$$

En réarrangeant l'équation précédente, nous pouvons exprimer les seuils juste détectables de la manière suivante :

$$\Delta L(L, \sigma(L_a)) = \sigma(L_a) \left(\frac{R(L, \sigma(L_a)) + \Delta R}{1 - (R(L, \sigma(L_a)) + \Delta R)} \right)^{\frac{1}{n}} - L \quad (4.66)$$

La difficulté est de calculer le seuil de réponse ΔR , qu'il n'est possible de calculer qu'avec le seuil de luminance juste détectable ΔL que nous cherchons à déterminer. Puisque nous ne pouvons pas calculer ΔR directement, Irawan propose de faire une approximation. Si l'on fait l'hypothèse que l'œil est adapté à chaque luminance qu'il observe ($\sigma(L) = L$), nous pouvons alors utiliser les valeurs de la fonction des seuils juste détectables donnée par Ferwerda, en fonction des luminances uniquement [Ferwerda 96]. La réponse des photorécepteurs sera alors $R(L, L)$, et la réponse générée en augmentant la luminance d'un seuil juste détectable sera $R(L + \Delta L, L)$. La différence entre ces deux grandeurs est la quantité de réponse $\Delta R(L, L)$ nécessaire pour détecter une différence de luminance.

$$\Delta R(L, L) = R(L + \Delta L, L) - R(L, L) \quad (4.67)$$

L'approximation proposée est de ne retenir que la plus petite valeur de $\Delta R(L, L)$, lorsque L varie :

$$\Delta R = \min_L (\Delta R(L, L)) \quad (4.68)$$

Nous noterons enfin que si pour une luminance donnée, la valeur $R(L, \sigma(L_a)) + \Delta R$ dans l'équation 4.66 est supérieure à 1, cela signifie que la luminance observée est trop supérieure de la luminance d'adaptation L_a , que l'œil sature et qu'il ne peut pas, dans

ce cas là, percevoir une différence de luminance avec une luminance encore supérieure. Dans ce cas là, on posera $\Delta L(L, \sigma(L_a)) = \infty$.

L'équation 4.66 nous permet de connaître les seuils de luminance juste détectables $\Delta L(L, \sigma(L_a))$, pour la luminance regardée L et la luminance d'adaptation L_a . Ces valeurs vont maintenant nous permettre de déterminer un seuil à ne pas dépasser lors de l'ajustement d'histogramme.

L'ajustement d'histogramme est une technique publiée par Ward (voir section 4.2.2 et [Ward 97]). Elle permet de compresser la dynamique de luminance de l'image HDR en utilisant toute la dynamique offerte par l'écran, et en assurant que les contrastes ne seront pas exagérés lors de ce traitement. Dans sa publication, Ward faisait l'hypothèse que l'œil est adapté à chaque luminance qu'il regarde, ce qui n'est pas le cas dans la réalité. Dans cette troisième étape de son opérateur, Irawan propose donc d'utiliser les seuils de luminance juste détectables $\Delta L(L, \sigma(L_a))$ pour améliorer l'ajustement d'histogramme. Irawan propose ainsi le seuil suivant, inspiré de celui proposé par Ward (équation 4.24) :

$$f(\log(L_{wi})) \leq \frac{\Delta L(L_{di}, L_{da})}{\Delta L(L_{wi}, L_{wa})} \frac{T \Delta b}{\log(L_d^{max}) - \log(L_d^{min})} \frac{L_{wi}}{L_{di}} \quad (4.69)$$

De la même manière que pour l'opérateur de Ward, les valeurs L_{di} et L_{da} dépendent normalement de l'image produite après l'ajustement d'histogramme. Puisque nous ne pouvons pas les obtenir de cette manière, elle sont approximées par la moitié de la luminance maximale d'un écran standard LCD, 150 cd/m^2 .

D'autre part, la technique d'ajustement d'histogramme réalisée par Irawan est différente dans la manière de traiter les valeurs de l'histogramme qui dépassent le seuil autorisé. Ward tronquait l'histogramme et recalculait le seuil en fonction des valeurs non-tronquées, mais cette méthode peut faire apparaître des sauts d'intensités si l'on traite des vidéos, car un petit changement dans l'histogramme peut grandement modifier son ajustement. Irawan propose une méthode temporellement cohérente pour éviter des discontinuités lors de traitements de vidéos. Plutôt que de tronquer l'histogramme, il redistribue les valeurs qui dépassaient le seuil, selon une méthode itérative, tout en veillant à ne pas dépasser le seuil autorisé (figure 4.9).

4.2.4 Les opérateurs locaux statiques

Les opérateurs globaux appliquent la même courbe de compression à tous les pixels de l'image, à l'inverse des opérateurs locaux. Leur temps d'exécution est donc géné-

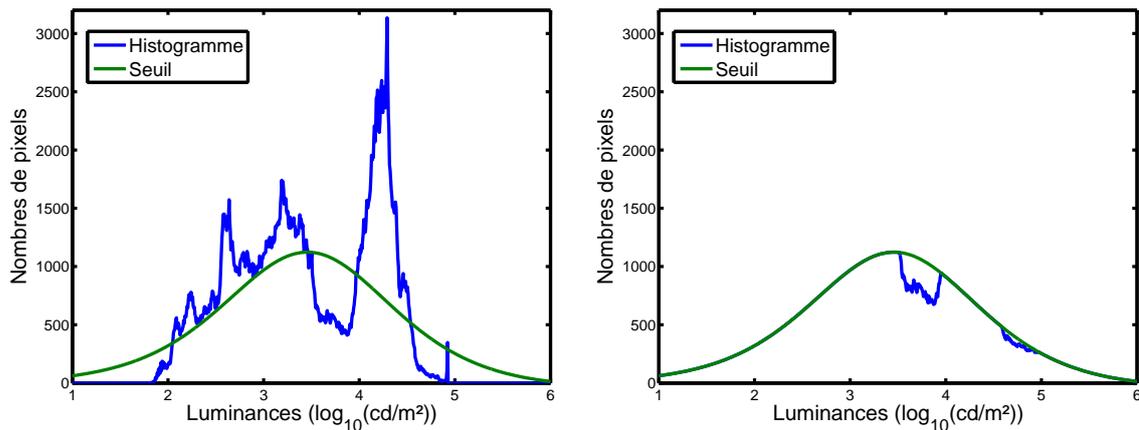


FIGURE 4.9 – Ajustement d’histogramme par la méthode d’Irawan et *al.* L’histogramme initial (gauche) et l’histogramme modifié par ajustement selon le seuil défini.

ralement plus rapide. Leurs fonctions de réponse pour compresser la dynamique de luminance est monotone et croissante, ce qui a l’avantage d’éviter d’introduire des artefacts (notamment des inversions de contraste et/ou des halos). Mais les opérateurs globaux assimilent la luminance d’adaptation à la moyenne logarithmique de l’image entière, alors qu’en réalité, l’œil s’adapte localement aux différentes zones lumineuses de l’image présentée.

Face à la faible dynamique de luminance des écrans standards, il pourrait être plus efficace d’optimiser l’utilisation de la dynamique de luminance offerte par les écrans standards, en utilisant une adaptation *locale*. Le traitement d’un pixel dépend alors de sa luminance initiale, mais également de ses pixels voisins. La conséquence est que deux pixels avec la même luminance initiale peuvent avoir deux luminances différentes à l’écran. La fonction de réponse ne sera plus forcément monotone croissante, ce qui peut avoir l’avantage d’avoir des contrastes locaux plus élevés, mais peut également avoir le défaut d’introduire des artefacts dans l’image traitée.

Les paragraphes suivants présentent quelques uns des principaux opérateurs locaux de reproduction de tons de la littérature.

L’opérateur de Chiu et *al.*

Chiu et *al.* [Chiu 93] ont été les premiers à présenter un opérateur local, en observant que de nombreux artistes utilisent ce procédé pour faire croire à l’observateur qu’il est devant une plus grande dynamique de luminance que celle qui lui est présentée. Chiu utilise un filtre passe-bas, par exemple un filtre gaussien, et réalise l’adaptation

locale suivante :

$$L_d = \frac{L_w}{kL_w^{blur}} \quad (4.70)$$

La constante k est réglée par l'utilisateur en fonction du poids qu'il veut donner à la composante locale.

Les résultats de cet opérateur mettent immédiatement en évidence l'inconvénient des opérateurs locaux, qui est l'apparition d'artefacts autour des zones avec de hautes fréquences spatiales, ici des halos sombres. Ces artefacts peuvent être réduits si l'on utilise un filtre de grande taille (plusieurs dizaines de pixels), mais l'adaptation locale est alors moins marquée.

L'opérateur retinex de Rahman et Johnson

Le modèle Retinex est un modèle de perception de la luminosité et des couleurs [Rahman 96]. Rahman et Johnson proposent un opérateur assez proche de celui de Chiu et *al.*, mais utilisent les logarithmes de luminances de l'image HDR, et peuvent utiliser une décomposition de l'image en plusieurs échelles. Dans ce cas, la formule de la compression de la dynamique est :

$$L_d = \exp \left(C \sum_{n=0}^N \omega_n \left(\log_{10}(L_w) - k \log_{10}(L_w^{blur}) \right) \right) \quad (4.71)$$

Le terme ω_n représente le poids associé à chaque échelle, et chaque image d'échelle n est convoluée avec une filtre gaussien. Le terme k spécifie l'importance donnée à l'image floutée. Plus celle-ci est grande, et plus la compression de la dynamique de luminance sera importante, mais cela créera également d'importants halos. Enfin, le terme C contrôle la saturation des couleurs, et est très important pour les zones où la luminance est faible et où les couleurs sont généralement désaturées.

Les opérateurs iCAM

Fairchild et *al.* proposent plusieurs opérateurs locaux, basés sur des modèles d'apparence colorée [Fairchild 02, Johnson 03, Kuang 07a]. Nous ne développerons ici que le premier des trois opérateurs cités, qui est composé de deux étapes successives : le modèle direct et le modèle inverse (pour l'affichage).

Dans un premier temps, les auteurs utilisent l'espace de couleur XYZ et définissent les luminances d'un point blanc pour réaliser une adaptation chromatique. Un paramètre permet à l'utilisateur de régler l'adaptation en fonction de l'image qu'il veut

utiliser. Ensuite, la compression de la dynamique de luminance est assurée par une fonction exponentielle dans l'espace de couleur *LMS*, dont l'exposant s'adapte localement en fonction des luminances à traiter.

Dans un second temps, un modèle inverse d'apparence colorée est utilisé pour afficher l'image à l'écran. Les auteurs utilisent une fonction exponentielle inverse, puis une adaptation chromatique inverse faisant l'hypothèse que l'observateur est pleinement adapté à l'image qu'il regarde. Avant l'affichage, l'opérateur supprime 1 % des valeurs extrêmes du résultat, et procède à une normalisation des valeurs entre 0 et 1.

Cet algorithme nécessite des valeurs en cd/m^2 en entrée. Un défaut est qu'il n'arrive pas à traiter correctement les images de faible luminance, un voile rouge apparaissant alors sur les images produites.

L'opérateur local de Pattanaik et al.

L'opérateur de Pattanaik et al. [Pattanaik 98] est également basé sur un modèle d'apparence colorée, mais il tient compte en plus de la sensibilité du système visuel humain aux motifs et à la luminosité, ainsi que de l'éblouissement. Cet opérateur local est très complexe.

D'abord, le modèle visuel place l'image HDR dans l'espace de couleur *LMS*. Il réalise une décomposition multi-échelle sur 7 niveaux des motifs de l'image pour extraire des contrastes locaux à différentes fréquences spatiales [Peli 90]. Un gain est appliqué à chacun des contrastes en fonction de la luminance d'adaptation locale.

Ensuite, le modèle d'affichage applique un nouveau gain, qui tient compte de la dynamique de luminance de l'écran utilisé. Les différentes échelles sont ensuite additionnées pour reconstruire l'image à afficher.

L'opérateur local de Reinhard et al.

Reinhard et al. proposent un opérateur basé sur un modèle photographique par zones de luminances [Reinhard 02]. Cet opérateur contient deux algorithmes : le premier est un opérateur global, le second un opérateur local. L'algorithme global est un des opérateurs les plus utilisés, pour sa simplicité et son efficacité, aussi bien au niveau du temps de calcul que du rendu visuel.

L'opérateur global procède en deux temps : d'abord, on effectue une compression linéaire sur la luminance réelle, en utilisant la moyenne logarithmique des luminances de l'image HDR L_m . Ensuite, comme en photographie traditionnelle, les auteurs uti-

lisent une fonction de forme sigmoïde pour obtenir les luminances à afficher :

$$L_d = \frac{L_{lin}}{1 + L_{lin}}, L_{lin} = \frac{0.18}{L_m} L_w \quad (4.72)$$

Cet opérateur global permet également d'introduire un terme supplémentaire pour définir quelle sera la plus petite luminance qui sera affichée en blanc, si le besoin s'en fait sentir. Cela permet de régler manuellement l'éventuelle saturation du capteur face aux hautes luminances.

L'opérateur local modifie le terme du dénominateur $1 + L_{lin}$ en $1 + V_1(s_m)$, où $V_1(s_m)$ représente la moyenne locale autour du pixel considéré, et s_m la taille optimale de la zone sur laquelle la moyenne est calculée. Cette modification locale permet d'associer différents temps d'exposition aux différentes parties de l'image, ce qui rehausse les contrastes. L'inconvénient de cette modification est que le terme s_m doit être réglé par l'utilisateur pour un résultat optimal.

L'opérateur de Grave et Brémond

L'opérateur de Grave et Brémond a été conçu pour une tâche visuelle de détection : il s'applique à restituer la perception des contrastes de scènes routières, *ie.* la visibilité [Grave 08]. Comme l'opérateur de Pattanaik [Pattanaik 98], les auteurs décomposent l'image HDR en 7 bandes de fréquences et les traitent séparément pour restituer l'ensemble des fréquences auquel le système visuel humain est sensible. Un modèle d'écran, prenant en compte la courbe de réponse de l'écran, permet d'adapter l'image à afficher en fonction de l'écran utilisé. Par contre, cet opérateur n'a été testé que sur des images en niveaux de gris, et n'est pas temps réel.

4.2.5 Un opérateur local dynamique

L'opérateur de Ledda et *al.*

Ledda et *al.* proposent un opérateur qui a pour caractéristiques d'être local et de simuler l'adaptation à la lumière ou à l'obscurité au cours du temps [Ledda 04]. A la différence du modèle d'adaptation visuel de Pattanaik [Pattanaik 00], les auteurs proposent ici des temps de réponse différents, suivant que l'adaptation se fasse par rapport à la lumière ou à l'obscurité. Ensuite, pour compresser la dynamique de luminance, les auteurs utilisent la sigmoïde du modèle de Naka et Rushton (voir l'équation 4.50), mais le terme σ simulant la luminance d'adaptation est un terme local. Il est calculé

à partir de la luminance des pixels voisins et à l'aide d'un filtre bilatéral, pour éviter l'apparition de halos.

4.2.6 Les opérateurs fréquentiels

L'opérateur bilatéral de Durand et Dorsey

Durand et Dorsey introduisent un opérateur fréquentiel automatique, qui décompose l'image HDR en deux niveaux avec un filtre bilatéral préservant les coins [Durand 02]. Le premier niveau contient les basses fréquences de l'image HDR (*base layer*), et le second contient les hautes fréquences, *ie.* les détails (*detail layer*). Seules les luminances du niveau contenant les basses fréquences sont compressées. Cet opérateur compresse efficacement la dynamique de luminance et produit des images photoréalistes, mais sans chercher à faire en sorte que l'image soit perçue à l'écran comme dans la réalité, selon le principe énoncé par Tumblin et Rushmeier (voir schéma 4.3 et [Tumblin 93]).

L'opérateur trilatéral de Choudhury et *al.*

Choudhury et *al.* utilisent la même démarche que pour l'opérateur précédent. A l'aide d'un filtre trilatéral, ils extraient les basses fréquences de l'image HDR pour les compresser, avant de réinjecter les hautes fréquences pour afficher l'image finale [Choudhury 03]. L'emploi d'un filtre trilatéral permet une meilleure séparation des basses et des hautes fréquences : il est plus performant dans la préservation des coins, ainsi que dans la suppression du bruit dans les zones de fort gradient et autour des coins. Cet opérateur est donc également très performant pour compresser la dynamique de luminance et afficher de forts contrastes, mais toujours sans chercher à ce que l'image affichée soit proche de celle perçue dans la réalité (voir schéma 4.3, [Tumblin 93]). Cet opérateur demande de plus un temps de calcul élevé.

4.2.7 Un opérateur gradient

L'opérateur gradient de Fattal et *al.*

Fattal et *al.* font le constat que l'œil est plus sensible aux différences de luminances qu'aux luminances elles-mêmes [Fattal 02]. Les auteurs proposent de réduire la dynamique de luminance en compressant les forts gradients de l'image HDR, mais sans supprimer les gradients de faible intensité qui contiennent les détails de l'image HDR.

Les auteurs résolvent ensuite une équation de Poisson pour chaque pixel de l'image pour éviter l'apparition de distorsions lors de la reconstruction de l'image à afficher. Comme les deux opérateurs précédents, cet opérateur est très efficace dans la compression de la dynamique de luminance et présente généralement des images réalistes. Il ne peut toutefois pas être directement utilisé dans des applications de simulations d'éclairage ou d'études de visibilité, car il ne simule pas la perception de l'image HDR par le système visuel humain. Selon les images, les contrastes sont parfois rehaussés à tel point que l'observateur voit que l'image a été traitée, et ne lui semble donc plus naturelle.

4.3 Evaluation des opérateurs de reproduction de tons

Avec l'augmentation du nombre d'opérateurs de reproduction de tons, il est désormais utile de comparer leurs performances respectives pour choisir un opérateur. Il n'est cependant pas simple de comparer les différents opérateurs entre eux, car il n'existe pas de métrique objective pour mesurer les performances de chaque opérateur.

Les évaluations présentées dans cette section sont des évaluations subjectives, avec différents critères de qualité. Certaines de ces études utilisent une *référence*, qui correspond soit à une scène réelle, soit à une image projetée avec un écran HDR. D'autres ne se réfèrent à aucune référence, et laissent le sujet libre dans son évaluation.

L'évaluation de Drago et *al.*

Drago et *al.* ont procédé à une évaluation de 6 opérateurs de reproduction de tons, en utilisant comme critères la similarité et la préférence [Drago 02]. 4 images différentes ont été utilisées, 3 représentant des scènes d'intérieur correctement éclairées, 1 représentant une scène extérieure de jour. Pendant l'expérimentation, les sujets doivent comparer deux à deux et noter sur une échelle de 0 à 100 les images traitées, et cela sans référence. Aucune question précise n'est posée aux sujets. Après une analyse statistique multidimensionnelle, les observations sont rapportées dans un espace à 3 dimensions, chaque dimension correspondant à un critère d'évaluation. Les critères sont la reproduction des contrastes, la reproduction des détails et le naturel apparent. Les algorithmes les mieux notés sont ceux de [Rahman 96] (Retinex), [Reinhard 02] et [Schlick 94].

L'évaluation de Kuang *et al.*

Kuang a proposé trois études distinctes pour évaluer la qualité des opérateurs [Kuang 07b]. Dans la première, les sujets comparaient sans référence 8 algorithmes deux-à-deux, sur 10 images, et devaient choisir l'image qu'ils préféraient. Dans la seconde, ils comparaient sans référence 6 opérateurs deux-à-deux, mais les sujets devaient noter leurs impressions selon 6 critères : les détails dans les zones lumineuses, les détails dans les zones sombres, les contrastes, la finesse, les couleurs, et les artefacts. Dans la troisième, les sujets comparaient 7 algorithmes deux-à-deux avec une référence, la scène réelle. 3 scènes réelles d'intérieur ont été utilisées. Les sujets devaient d'abord choisir l'image qu'ils préféraient parmi les deux présentées, puis noter chaque image selon les critères suivants : contraste dans les zones très lumineuses, contraste dans les zones sombres, couleur dans les zones très lumineuses, couleur dans les zones très sombres, et contraste général.

[Durand 02], puis [Ward 97] et [Fairchild 02] ont été les mieux notés sur l'ensemble des trois expérimentations. Les résultats montrent une forte corrélation entre la préférence et les critères sur les contrastes et les couleurs.

L'évaluation de Yoshida *et al.*

Yoshida *et al.* ont réalisé une évaluation de 7 opérateurs de reproduction de tons en utilisant 5 critères de notation : la luminosité, les contrastes, la reproduction des détails dans les régions sombres, la reproduction des détails dans les régions lumineuses, et le naturel apparent [Yoshida 05]. Les sujets devaient comparer une à une chaque image traitée avec une référence, la scène réelle qui avait été photographiée.

Les résultats de cette étude montrent que les opérateurs globaux sont meilleurs pour la reproduction de la luminosité et des contrastes, et rendent des images qui paraissent plus naturelles. [Ward 97] gagne l'évaluation, bien qu'il ne soit pas le plus performant dans la reproduction des détails dans les zones sombres ou lumineuses.

L'évaluation de Ledda *et al.*

Ledda *et al.* proposent également une évaluation des opérateurs, mais ont la caractéristique d'utiliser un écran HDR comme référence, ce qui leur permet d'utiliser plus d'images (23) que les autres expérimentations [Ledda 05]. Deux expérimentations ont été réalisées. Pour la première, les sujets comparaient les images traitées deux-à-deux avec la référence, et devaient choisir l'image traitée la plus similaire à la référence.

Pour la seconde expérimentation, les sujets procédaient de la même manière, mais jugeaient les images sur la reproduction des détails dans les zones sombres, puis sur la reproduction des détails dans les zones lumineuses.

Sur les 6 opérateurs pris en compte, [Johnson 03] et [Reinhard 02] sont les mieux notés, suivis de [Ward 97] et [Ledda 04].

L'évaluation de Cadík et *al.*

L'évaluation de Cadík et *al.* porte sur la reproduction des attributs perceptuels des images [Cadík 08]. Deux expérimentations ont été menées : la première consistait à noter les images traitées par rapport à une référence (3 scènes réelles ont été utilisées, 1 de jour, 1 au crépuscule, 1 de nuit). La seconde consistait à classer les images, sans référence. Les critères sur lesquels portaient les notations étaient la luminosité, les contrastes, la reproduction des couleurs, la reproduction des détails, et la qualité de l'image dans son ensemble.

Sur les 14 opérateurs utilisés, la compression linéaire (ajustée pour chaque image), [Reinhard 02] et [Tumblin 99] sont les opérateurs les mieux notés et classés. En outre, Cadík propose une mesure de la qualité de l'image dans son ensemble, à partir des performances des opérateurs sur chaque attribut.

L'évaluation de Grave et Brémond

A la différence des évaluations précédentes, l'expérimentation de Grave et Brémond a comme unique critère la reproduction de la détection, c'est-à-dire que l'observateur doit percevoir, du point de vue de ses performances, la scène de manière identique dans la réalité et à l'écran [Grave 08]. Pour mesurer la détection, les auteurs utilisent un anneau de Landolt, dont la position de l'ouverture et le contraste varient d'une image à l'autre.

Au cours de l'expérimentation, les performances visuelles de détection de 10 sujets étaient d'abord mesurés sur une image HDR (la référence), puis sur les images traitées. Sur les 5 opérateurs évalués, l'opérateur de [Grave 08], puis de [Pattanaik 98] sont les meilleurs sur le critère de détection. En revanche, les auteurs insistent sur le fait que les critères de détection et d'apparence ne sont pas corrélés.

L'évaluation de Ashikhmin et *al.*

L'expérience Ashikhmin et Goyal est différente des précédentes, puisque son but n'est pas d'évaluer les opérateurs de reproduction de tons, mais d'étudier l'influence de l'utilisation d'une référence dans ces expérimentations [Ashikhmin 06]. Les auteurs comparent 5 opérateurs sur 4 scènes réelles d'intérieur et 2 scènes affichées sur un écran HDR.

3 expérimentations successives ont été menées. Pour la première expérimentation, les sujets classaient les images traitées en fonction de leur préférence, sans référence. Pour la seconde, les images devaient être classées en fonction du réalisme des images, toujours sans référence. Pour la troisième expérimentation, les sujets classaient les images en fonction de l'image de référence observée, le critère étant d'avoir l'image la plus proche possible de la référence.

Les résultats montrent qu'il n'y a statistiquement pas de différence entre préférence et fidélité s'il n'y a pas de référence. Cependant, les classements diffèrent entre les évaluations du réalisme (sans référence) et de la fidélité (avec référence), d'où la nécessité d'utiliser une référence pour les évaluations de la fidélité, afin d'éviter les biais liés aux jugements des sujets. Enfin, les résultats de la troisième expérimentation mettent en évidence les opérateurs de [Drago 03], [Fattal 02] et [Ward 97].

4.4 Simulation de l'éblouissement

Nous avons vu dans la section 2.7 que les écrans standards ne peuvent afficher de luminances supérieures à 300 cd/m^2 . Cette luminance maximale est trop faible pour que le système visuel humain puisse se sentir ébloui. De plus, l'œil étant sensible jusqu'à des luminances atteignant 10^6 cd/m^2 , il est utile, dans le cadre d'une visualisation réaliste, de suggérer à l'observateur la présence de zones de forte luminosité à l'écran, même si les luminances affichées ne dépassent pas 300 cd/m^2 . Une étude de Yoshida et *al.* met notamment en évidence que la sensation de luminosité perçue peut être augmentée de 20 % en simulant l'éblouissement à l'observateur [Yoshida 08]. Il est de plus important, toujours dans le cadre de la visualisation réaliste, de faire apparaître à l'écran ce que l'observateur voit dans la réalité lorsqu'il est ébloui : la luminance de voile, les stries et le halo lenticulaire (section 3.2.4).

La simulation de l'éblouissement répond donc à deux objectifs :

- faire paraître une zone lumineuse à l'observateur plus lumineuse que ce qu'elle n'est affichée ;

- reproduire à l'écran l'éblouissement tel qu'il serait vu par l'observateur dans la réalité (avec la luminance de voile, les stries et le halo lenticulaire).

L'éblouissement est un phénomène qui intervient lorsque la lumière traverse l'œil et avant qu'elle ne soit captée par les photorécepteurs de la rétine (voir section 3.2.4). La simulation de l'éblouissement est donc un traitement qui s'effectue avant que l'opérateur de reproduction de tons n'intervienne.

Nous présentons dans cette section quatre méthodes de simulation de l'éblouissement de la littérature.

Simulation de l'éblouissement par des filtres gaussiens

L'approche de Kawase est de ne simuler que la luminance de voile autour des zones très lumineuses de l'image HDR [Kawase 04]. D'abord, l'image HDR est seuillée afin de ne traiter que les zones très lumineuses susceptibles d'éblouir l'observateur. Ensuite, la luminance de voile est générée à partir de l'image HDR seuillée, par applications successives d'un filtre gaussien : la première image filtrée est générée à partir de l'image HDR seuillée, la seconde image filtrée est générée à partir de la première, et ainsi de suite. Kawase propose de répéter l'opération 5 fois, et d'additionner chaque image produite pour obtenir la luminance de voile. Selon les applications ou le rendu recherchés, différents poids peuvent être appliqués à chaque image, pour modifier la quantité de lumière diffractée et l'atténuation du voile lumineux en fonction de la distance par rapport à la source.

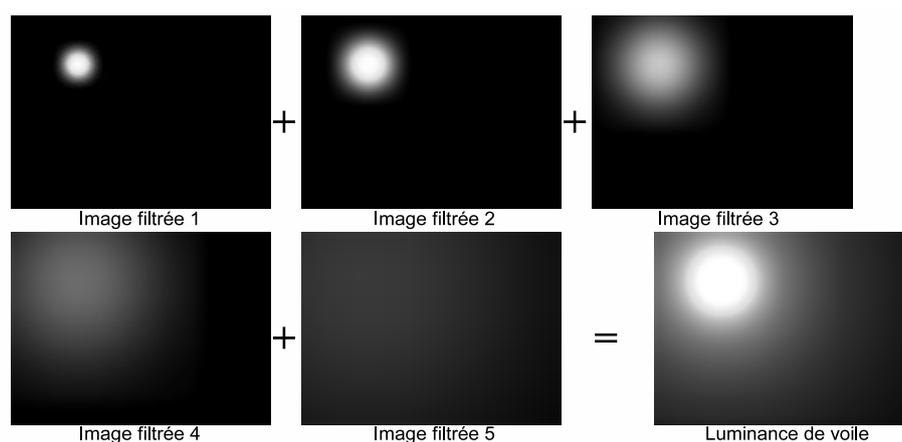


FIGURE 4.10 – La luminance de voile est générée à partir de plusieurs images, traitées successivement par un filtre gaussien. D'après [Kawase 04].

Simulation de l'éblouissement par Spencer et *al.*

Spencer et *al.* proposent pour la première fois de simuler l'éblouissement provoqué par une source ponctuelle (la luminance de voile, la couronne ciliaire et le halo lenticulaire) à partir d'une modélisation physique de la diffraction de la lumière dans l'œil [Spencer 95]. A partir du calcul de la taille de la pupille en fonction de la luminance d'adaptation, qui permet de déterminer si l'observateur se trouve dans le domaine photopique, mésopique ou scotopique, les auteurs proposent différents paramètres pour créer un filtre simulant la diffusion de la lumière dans l'œil, par rapport à un point donné (aussi appelé *point-spread-function*). Ce filtre est normalisé de façon à ce que la somme des composantes du filtre soit égale à 1, afin de respecter la conservation de la lumière dans l'œil. Finalement, ce filtre est convolué avec l'image HDR seuillée, contenant les luminances les plus élevées susceptibles d'éblouir l'observateur.

La luminance de voile et le halo lenticulaire sont calculés en fonction de l'angle par rapport à la source et de la taille de la pupille, à partir de lois empiriques proposés par des physiologues de la vision. La couronne ciliaire est calculée à partir d'une image de stries, simulant les perturbations locales des milieux transparents du cristallin.

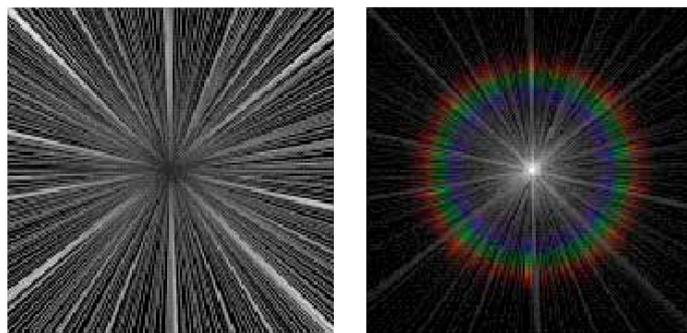


FIGURE 4.11 – Image de stries (gauche) et filtre final utilisé pour simuler l'éblouissement (droite). D'après [Spencer 95]

Simulation de l'éblouissement par Kakimoto et *al.*

Kakimoto et *al.* proposent une nouvelle approche pour la simulation de l'éblouissement, basée sur les ondes optiques [Kakimoto 05]. Les auteurs modélisent l'œil humain comme un système optique avec une ouverture (la pupille), une lentille (le cristallin) et un plan image (la rétine). L'image représentant l'éblouissement est générée en utilisant l'équation de la diffraction de Fraunhofer, qui est équivalente à une transformée

de Fourier en deux dimensions de l'image :

$$I_f(x_f, y_f) = \frac{A^2}{\lambda^2 f^2} |\mathbf{F}[P(x_0, y_0), \lambda f]| \quad (4.73)$$

où A et λ sont l'amplitude et la longueur d'onde de l'onde optique incidente, f est la focale de la lentille, P est la fonction 2D représentant l'ouverture de l'iris et éventuellement la présence d'obstacles pour la lumière entrante (cils, paupières), et \mathbf{F} la transformée de Fourier en deux dimensions.

Simulation de l'éblouissement par Ritschel et al.

De même que Kakimoto et al., Ritschel et al. modélisent la diffraction de la lumière dans l'œil à partir d'ondes optiques [Ritschel 09]. Les auteurs utilisent l'équation de Fresnel pour simuler la diffraction, dont l'équation de Fraunhofer est un cas particulier. La modification correspond à multiplier l'image de la pupille $P(x_0, y_0)$ par un terme complexe $E(x_p, y_p)$, et permet d'obtenir des résultats plus réalistes.

$$I_f(x_f, y_f) = \frac{A^2}{\lambda^2 f^2} |\mathbf{F}[P(x_0, y_0)E(x_p, y_p), \lambda f]| \quad (4.74)$$

De plus, Ritschel et al. modélisent la diffraction de la lumière de manière dynamique, en simulant les variations de l'apparence de l'ouverture au cours du temps. Dans leur modèle dynamique de l'œil humain, les auteurs tiennent ainsi compte des fibres présentes sur la cornée (responsables de 25 à 30 % de la diffraction), de la taille de l'ouverture de la pupille, de la déformation du cristallin, des inhomogénéités de l'humeur vitreuse, des particules présentes sur la rétine, des cils et du clignement des yeux. De même que Spencer et al., ces modélisations servent à créer un filtre pour l'éblouissement qui est convolué aux hautes luminances de l'image HDR.

Ce modèle est le plus complet de la littérature. Il a en revanche un coût d'implémentation et d'exécution, puisqu'il n'est temps réel que pour des filtres de taille 256×256 pixels.

4.5 Saillance d'images et de vidéos

L'attention visuelle est un mécanisme extrêmement difficile à modéliser. Pour l'observateur, deux mécanismes entrent en jeu : le premier, nommé *bottom-up*, est lié aux caractéristiques de l'image qui attirent naturellement le regard de l'observateur. Le second, nommé *top-down*, est lié à la tâche d'observation de l'observateur.

Les modèles *bottom-up* ont connu un grand essor à travers les modèles de saillance [Itti 01]. Ces modèles produisent des images de saillance, qui représentent des zones attirant le regard d'un observateur. Ces modèles ne tiennent toutefois pas compte d'une tâche éventuelle de l'observateur, par exemple une tâche de détection.

Dans le chapitre 6, nous proposerons l'idée d'évaluer les opérateurs de reproduction de tons en utilisant des cartes de saillance. L'évaluation consistera à comparer la carte de saillance d'une image HDR avec celle de son image traitée par un opérateur de reproduction de tons. C'est pourquoi nous présentons dans cette section le plus utilisé des modèles de saillance.

Le modèle de saillance *bottom-up* d'Itti et al.

Itti a proposé plusieurs variantes de son modèle de saillance. Nous présentons ici la seconde version [Itti 00], qui utilise 3 caractéristiques pour identifier les zones qui attireront le regard de l'observateur : les couleurs, l'intensité et les orientations.

Dans un premier temps, Itti traite chaque caractéristique indépendamment des autres. A partir de l'image à traiter, Itti construit d'abord 2 pyramides dyadiques pour les couleurs (chacune contenant 2 couleurs opposées, Rouge-Vert et Bleu-Jaune), 1 pyramide pour l'intensité, et 4 pyramides pour les orientations (4 angles d'orientations distincts). En utilisant plusieurs différences de gaussiennes entre les étages de chaque pyramide, Itti met en relief les contrastes existants dans chaque pyramide de chaque caractéristique. Pour traiter la pyramide des intensités et des orientations, Itti construit une série de 6 cartes, avec des différences d'images de résolutions différentes mais appartenant à la même pyramide (symbole \ominus), et en utilisant les niveaux 3 à 8 de la pyramide :

$$I(c, s) = |I(c) \ominus I(s)|, \quad c \in \{2, 3, 4\}, \quad \delta \in \{3, 4\}, \quad s = c + \delta \quad (4.75)$$

où $I(x)$ représente l'image de niveau x dans la pyramide dyadique des intensités, et :

$$O(c, s, \theta) = |O(c, \theta) \ominus O(s, \theta)|, \quad c \in \{2, 3, 4\}, \quad \delta \in \{3, 4\}, \quad s = c + \delta \quad (4.76)$$

où $O(x)$ représente l'image de niveau x dans la pyramide dyadique des orientations.

Pour les couleurs, Itti procède de plus à une normalisation par la carte d'intensité de la plus petite échelle $I(c)$:

$$C(c, s) = \frac{|C(c) \ominus C(s)|}{I(c)}, \quad c \in \{2, 3, 4\}, \quad \delta \in \{3, 4\}, \quad s = c + \delta \quad (4.77)$$

où $C(x)$ représente l'image de niveau x dans la pyramide dyadique des couleurs.

Nous avons donc 6 cartes pour chaque pyramide de chaque caractéristique. Dans un second temps, Itti fusionne les informations contenues dans chacune d'elles, en utilisant un processus de normalisation pour accentuer l'importance des zones qui peuvent attirer l'œil, et pour minimiser l'importance des autres zones.

Itti fusionne donc les cartes de chaque caractéristique en 3 cartes de « conspécuité », une pour chaque caractéristique, en utilisant le processus de normalisation. La carte de conspécuité des intensités est calculée en normalisant chacune des 6 cartes de la pyramide des intensités, puis en les additionnant entre elles (équation 4.78, où le symbole N représente la normalisation, et où le symbole $\oplus_{c,s}$ représente l'addition de cartes de résolutions différentes, pour tous les niveaux couverts par les variables c et s). La carte de conspécuité des couleurs est calculée en normalisant chaque carte, puis en additionnant toutes les cartes Rouge-Vert et Bleu-Jaune (équation 4.79). Enfin la carte de conspécuité des orientations est obtenue en normalisant chacune des cartes d'orientation, en additionnant les cartes de chaque pyramide d'orientation, puis en normalisant chaque carte intermédiaire de chaque pyramide avant de les additionner ensemble (équation 4.80).

$$\bar{I} = \oplus_{c,s} N(I(c,s)) \quad (4.78)$$

$$\bar{C} = \oplus_{c,s} (N(RG(c,s)) + N(BY(c,s))) \quad (4.79)$$

$$\bar{O} = \sum_{\theta=0^\circ,45^\circ,90^\circ,135^\circ} N(\oplus_{c,s} N(O(c,s,\theta))) \quad (4.80)$$

Finalement, la carte de saillance est obtenue en normalisant les cartes de conspécuité des intensités, couleurs et orientations, et en les additionnant ensemble :

$$Saillance = \frac{1}{3}(N(\bar{I}) + N(\bar{C}) + N(\bar{O})) \quad (4.81)$$

D'autres variantes du modèle de saillance d'Itti existent, proposant de normaliser différemment les images de chaque caractéristique [Itti 98], et d'intégrer d'autres caractéristiques, notamment pour le traitement de vidéos [Itti 03]. D'autres modèles de saillance existent également dans la littérature [Itti 05], notamment celui de LeMeur et *al.*, créé pour compresser la taille de vidéos en fonction de la saillance [LeMeur 06]. En revanche, il n'existe pas à ce jour de modèle de saillance développé pour les images HDR.

4.6 Conclusion

Ces états de l'art sur le rendu HDR, les opérateurs de reproduction de tons, leurs évaluations, la simulation de l'éblouissement et la saillance ont permis de présenter les travaux de la littérature que nous utiliserons par la suite. Dans le cadre de la visualisation réaliste d'images de synthèse HDR, cet ensemble de travaux ne permet pas, à lui seul, de réaliser notre objectif, car plusieurs points restent à développer.

Tout d'abord, sous *OpenGL*, nous n'avons pas trouvé d'exemple de rendu HDR temps réel, permettant d'obtenir des luminances à partir de l'éclairage d'un environnement virtuel. Certes, il existe déjà de nombreuses applications, mais celles-ci ne rendent généralement à l'écran qu'un ou deux objet(s), et il n'y a pas réellement de modèle d'éclairage réaliste qui met en jeu des sources de lumière avec un environnement virtuel éclairé. D'autre part, alors que le rendu HDR permet de générer des images en unités physiques, les applications actuelles n'en tirent généralement pas parti, se contentant de générer des images en unités arbitraires.

Ensuite, nous avons vu qu'il existe de nombreux opérateurs de reproduction de tons. Chaque opérateur crée une image qui lui est propre, et distincte de celles des autres opérateurs. Nous pouvons dire qu'**il y a autant d'images produites que d'opérateurs**, chacune avec des points forts et des points faibles. **Aucun opérateur ne peut donc convenir à toutes les applications possibles.**

L'étape de la compression de la dynamique de luminance dégrade forcément le contenu de l'image HDR, mais nous avons vu, à travers les évaluations d'opérateurs, que certains opérateurs sont plus efficaces que d'autres. L'opérateur de [Ward 97], notamment, a souvent été bien noté dans ces évaluations. De plus, notre critère principal étant le réalisme (*ie.* la fidélité par rapport à la réalité), un opérateur se détache des autres par ses propriétés sur la conservation des contrastes à l'écran, inspirées de l'opérateur de [Ward 97], et son modèle d'adaptation temporelle : **l'opérateur d'Irawan et al.** [Irawan 05]. Nous utiliserons donc cet opérateur par la suite.

D'un point de vue plus général, et à partir des diverses évaluations des opérateurs existants réalisées, nous avons vu qu'il n'y a pas consensus sur un unique opérateur. Si l'opérateur idéal n'existe pas, il existe en revanche des opérateurs plus ou moins bien adaptés à diverses applications. Quel que soit l'opérateur utilisé par la suite, l'évaluation de ses performances est importante pour valider le réalisme des images

produites. De nouvelles évaluations seront donc réalisées sur des images statiques, mais également sur des vidéos, ce qui n'a jamais été réalisé auparavant (section 6.5). Pour une de ces évaluations, nous exploiterons un critère d'évaluation lié à la saillance des images HDR, et chercherons donc à étendre un des algorithmes de saillance déjà existant à ces images.

Enfin, plusieurs techniques simulant l'éblouissement à l'écran ont également été citées. Ces simulations permettant de compenser une partie de la faible puissance lumineuse des écrans standards, notre rendu HDR intégrera donc une de ces simulations.

Chapitre 5

Rendu HDR

Les avancées en matière de synthèse d'images sont telles qu'il est possible de rendre des scènes virtuelles dans une unité physique, en luminance [Ward 98, Dumont 00]. Ce sont des images à grande dynamique de luminance, en anglais *High Dynamic Range*, ou HDR. De plus, le rendu des images HDR est réalisable en temps réel, à condition de ne simuler qu'une partie des trajets des photons de la lumière (par exemple pour *OpenGL*, une seule réflexion entre la source et la caméra pour chaque pixel).

Cependant, les dispositifs d'affichage (hors écrans HDR) sont limités techniquement, notamment par la dynamique de luminance et la quantification sur 256 niveaux. Ils ne sont donc pas capables de restituer le contenu exact des images HDR à l'utilisateur. Ces limites nous contraignent à utiliser un opérateur de reproduction de tons, qui compresse la dynamique de luminance des images HDR, pour pouvoir afficher ces images sur des écrans standards.

Le terme de « rendu HDR » regroupe deux étapes : le rendu de l'image HDR, et son affichage à l'écran via un opérateur de reproduction de tons. Une troisième étape, la simulation de l'éblouissement face aux sources lumineuses, peut également être considérée comme une étape du rendu HDR, ou faire partie de l'opérateur de reproduction de tons. Bien qu'il soit tout à fait réalisable techniquement, le rendu HDR en luminance a encore été peu exploité dans la pratique sur de grands environnements virtuels.

Dans le cadre de cette thèse sur la visualisation réaliste d'environnements virtuels, nous nous proposons de mettre en œuvre un mode de rendu HDR, et de l'appliquer sur des scènes routières, de nuit, via le moteur de rendu *OpenGL* (en pratique, nous

utilisons *OpenSceneGraph*, qui est une surcouche d'*OpenGL*). Dans ce chapitre, nous présentons dans un premier temps les modèles employés pour réaliser le rendu temps réel d'images HDR dans des textures. Dans un second temps, nous choisissons un opérateur de reproduction de tons en fonction de notre critère de visualisation réaliste, et l'implémentons pour traiter en temps réel les images HDR générées.

5.1 Génération d'images de synthèse HDR

L'objet de cette section est d'établir un pipeline de rendu HDR pour la génération d'images de synthèse HDR d'environnements virtuels. Ce pipeline définira les interactions entre les différents éléments de la scène : les sources de lumière et les matériaux qui réfléchissent la lumière. A partir de bases de données virtuelles créées initialement pour un rendu LDR, nous verrons comment exploiter la description des environnements virtuels pour un rendu HDR, en utilisant un modèle physique d'éclairage.

5.1.1 Environnements virtuels

Les environnements virtuels sont généralement créés par des infographistes. La construction de la scène virtuelle se déroule en deux étapes. La première consiste à construire la géométrie de la scène. Si on prend pour exemple une scène routière, cela correspond à définir la forme des routes, bâtiments, voitures, et de tout autre objet de la scène. Ensuite, la seconde étape est d'habiller ces éléments. Cela consiste à leur donner des couleurs. Suivant le logiciel de modélisation utilisé, il existe différents paramètres pour définir les couleurs des éléments.

Dans cette section, nous nous basons sur les conventions adoptées par le moteur de rendu *OpenGL*. Néanmoins, quel que soit le moteur de rendu utilisé, il convient à l'utilisateur de s'assurer que la conversion de l'environnement virtuel du logiciel de modélisation dans le moteur de rendu se fasse sans perte de données, ce qui est généralement un travail délicat.

5.1.2 Description des sources de lumière

OpenGL propose 3 types de sources de lumière, pour compenser le fait qu'il ne simule qu'une seule réflexion de l'émission de la lumière par la source jusqu'à son absorption par la caméra.

Nous avons d'abord les sources **ambiantes**. Même si chaque source ambiante a une position dans l'espace, la lumière émise par ces sources est définie sans origine ni direction, et atteint tous les objets de la même manière, quelle que soit leur orientation par rapport à la source ambiante. Si un coefficient d'atténuation en fonction de la distance est défini, la lumière ambiante ne touchera que les objets proches de la source. Dans le cas inverse, *ie.* en l'absence de coefficient d'atténuation, la lumière ambiante atteindra les objets de la scène entière.

Nous avons ensuite les sources **diffuses**. Pour chaque point de la scène éclairé par une source diffuse, la lumière vient d'une direction donnée (celle de la source), mais est également réfléchi dans toutes les directions. La quantité de lumière réfléchi est plus importante si la source est au-dessus de la surface que si elle est de côté, avec un faible angle incident. Un coefficient d'atténuation en fonction de la distance entre la source et le point éclairé peut également être défini.

Nous avons enfin les sources **spéculaires**. A l'inverse de la lumière diffuse, la lumière spéculaire est réfléchi dans une direction donnée. C'est cette lumière qui permet de simuler les reflets des sources sur les éléments de la scène.

Même s'il ne permet pas de réaliser des simulations d'éclairage exactes, *OpenGL* fournit néanmoins des outils pour s'en approcher. Les paragraphes suivants montrent comment nous pouvons les exploiter pour un rendu aussi physique et exact que possible.

Intensité des sources lumineuses

Chaque source de lumière de l'environnement peut avoir des composantes ambiante, diffuse et spéculaire. Pour simuler un éclairage « physique », les intensités de ces composantes sont décrites en candela. Dans notre modélisation, les intensités des termes diffus et ambiant sont différentes, car elles ne simulent pas le même type d'éclairage. Le terme diffus permet de simuler l'éclairage direct, qui est réfléchi vers la caméra en une seule réflexion. L'intensité diffuse doit donc être égale à l'intensité réelle de la source. Dans notre modèle, le terme ambiant permet lui de simuler approximativement les réflexions secondaires qui ne sont pas calculées par *OpenGL*. La quantité de lumière étant atténuée par les réflexions, l'intensité ambiante doit donc être inférieure au terme diffus. Concrètement, nous fixons généralement le terme ambiant I_A à un dixième du terme diffus I_D :

$$I_A = \frac{I_D}{10} \quad (5.1)$$

De plus, un terme ambiant additionnel peut également être utilisé sur l'ensemble de l'environnement virtuel, pour simuler l'éclairage du ciel (voir l'équation 5.14 ci-après).

Enfin, nous fixons le terme spéculaire I_S égal au terme diffus :

$$I_S = I_D \quad (5.2)$$

Il nous reste maintenant à définir l'intensité diffuse I_D à partir de données réelles. Pour le simulateur de conduite, nous avons utilisé les données des constructeurs de lampadaires [Philips 99] et des mesures de phares automobiles [Schoettle 03]. Nous avons fixé l'intensité maximale des lampadaires à 580 *cd*, et celle des phares à 5000 *cd*. Nous avons de plus défini une couleur légèrement jaune-orangée pour les luminaires (valeurs RGB de 750, 650 et 450 *cd*), tandis que nous avons réglé les phares sur une lumière blanche (valeurs RGB identiques, à 5000 *cd*).

Les intensités que nous venons de définir pour chaque source sont les intensités maximales qu'elles peuvent émettre. Pour plus de réalisme, nous définissons également des formes d'émission pour chacune de ces sources.

Forme d'émission des sources

La forme d'émission des sources diffuses a également été paramétrée. Pour les lampadaires, nous avons relevé la forme d'émission d'un lampadaire réel dans un catalogue [Philips 99], et nous l'avons approximée par une fonction de l'angle vertical θ . L'intensité est obtenue selon la formule suivante :

$$I_D(\theta) = I_D^{max} \left(0.5 + \frac{\cos(2(\theta - 50))}{2} \right) \quad (5.3)$$

En pratique, cette fonction produit un éclairage maximal de $I_D^{max} = 580$ *cd* pour un angle de 50° (figure 5.1).

Pour les phares des voitures, nous avons simplement utilisé une émission en forme de lobe (ou de parabole), réalisée avec une fonction polynômiale du second degré. Des textures réalisées à partir de mesures sur des phares réels peuvent également être utilisées. Pour chaque lobe, l'intensité diffuse est projetée devant le phare sur un angle de $2 \times \theta_s = 52^\circ$. Le coefficient d'atténuation, $k(\theta)$, fonction de l'excentricité, est multiplié par l'intensité diffuse maximale du phare définie dans le paragraphe précédent :

$$k(\theta) = \left(\frac{\cos(\theta) - \cos(\theta_s)}{1 - \cos(\theta_s)} \right)^2, \quad \theta_s = 26^\circ, \quad \theta \leq 26^\circ \quad (5.4)$$

La forme d'émission de l'intensité des phares est représentée sur la figure 5.2.

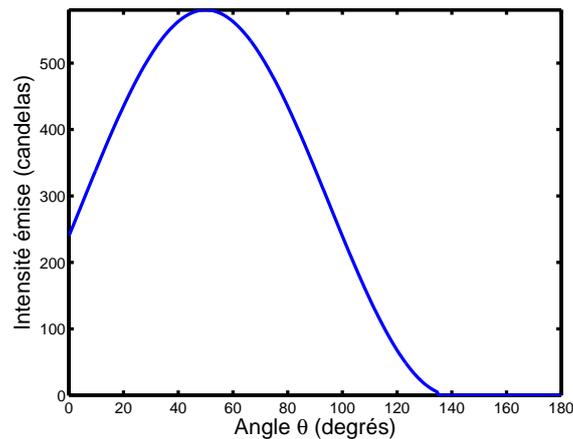


FIGURE 5.1 – Forme d'émission d'un lampadaire, en fonction de l'angle θ (ici vertical).

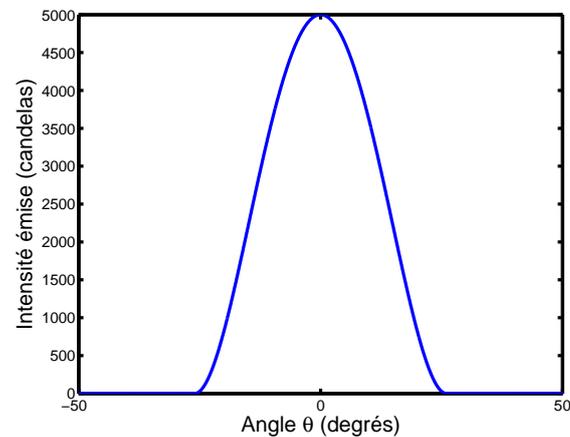


FIGURE 5.2 – Forme d'émission d'un phare, en fonction de l'angle θ (ici horizontal).

Atténuation de la lumière

Dans la réalité, les sources diffuses ont une atténuation proportionnelle à l'inverse du carré de la distance d . Nous affectons donc une atténuation identique aux sources diffuses de l'environnement virtuel. Nous affectons également cette atténuation aux sources ambiantes qui simulent les réflexions secondaires des sources diffuses. A titre d'exemple, pour un lampadaire, nous avons la formule suivante :

$$I_D(d, \theta) = \frac{I_D(\theta)}{d^2} \quad (5.5)$$

Le terme ambiant est atténué de la même manière. En revanche, la lumière spéculaire n'est pas atténuée par la distance.

Représentation visuelle des sources

Dans notre modèle, la représentation visuelle d'une source revient à définir quelle est la luminance de sa texture. Nous disposons de données constructeurs sur les intensités des sources de lumière, mais pas sur leurs luminances : nous devons donc relier ces deux grandeurs en utilisant les propriétés photométriques des sources lumineuses.

Dans le cas particulier où la source est à luminance uniforme (spatialement et angulairement invariante, c'est-à-dire que la luminance ne dépend ni du point de la source observé, ni de la direction d'observation), la luminance L est fonction de l'intensité I et de la surface apparente de la source $S \cdot \cos(\gamma)$, où γ est l'angle entre la normale de la surface et le vecteur source-œil [Mezzone 02] :

$$L = \frac{I}{S \cdot \cos(\gamma)} \quad (5.6)$$

Dans le cas d'une source sphérique diffuse, la surface apparente dans chaque direction est celle du disque de la sphère ($S = \pi R^2$ et $\cos(\gamma) = 1$). Pour un lampadaire dont la source est sphérique et d'intensité uniforme, nous établissons donc la relation suivante :

$$L = \frac{I_D}{\pi r^2} \quad (5.7)$$

L'intensité des lampadaires que nous avons décrite précédemment n'est cependant pas uniforme, puisqu'elle varie en fonction de la direction éclairée (figure 5.1). Cependant, pour garder un aspect visuel globalement uniforme de la source, nous définissons la luminance en fonction de l'intensité maximale de la source, comme si c'était une source d'intensité uniforme. La valeur de luminance placée dans le coefficient émissif est alors :

$$L = \frac{I_D^{max}}{\pi r^2} \quad (5.8)$$

A titre d'exemple, dans notre environnement virtuel, le rayon r de la source apparente du lampadaire est égal à 0.5 mètre. La luminance des lampadaires est donc égale à 738 cd/m^2 .

De plus, si la source n'est pas spatialement uniforme, il faut créer une texture HDR pour modéliser les variations spatiales. Dans le cas d'un phare de voiture, la surface apparente dépend du point de vue de l'observateur, à travers le terme $\cos(\gamma)$, mais aussi et surtout de la forme d'émission de l'intensité du phare. Cette forme d'émission étant difficile à modéliser, nous proposons de simplifier le modèle par une approximation. Cette approximation consiste à garder les luminances de la texture HDR constantes,

quel que soit l'angle d'observation. Nous utilisons donc simplement la texture initiale du phare éteint, que nous multiplions par le terme L_{max} suivant (figure 5.3) :

$$L_{max} = \frac{I_D^{max}}{S}. \quad (5.9)$$

Toujours à titre d'illustration, nous avons approximé la surface apparente d'un phare à celle d'un rectangle, et nous l'avons mesurée égale à $0.3 \times 0.1 = 0.03 \text{ m}^2$. Dans notre environnement virtuel, nous avons donc fixé la luminance maximale des phares à environ $16\,700 \text{ cd/m}^2$ (figure 5.3).



FIGURE 5.3 – Texture initiale de phares éteints (gauche), et image HDR de phares allumés (logarithme décimal de la luminance, droite). Image de la base de données 3D INRETS-MSIS-2007.

Cette modélisation reste une simplification de la réalité, car la luminance d'une source de lumière dépend généralement non seulement du point de la source observé, mais aussi de l'angle selon lequel nous le regardons. C'est clairement une limite de notre modèle, notamment pour le cas du phare de voiture.

5.1.3 Description des objets éclairés

Dans *OpenGL*, un objet à éclairer peut avoir sa ou ses couleur(s) définie(s) de deux manières. Si l'objet a une couleur unie, elle sera décrite par un matériau M , sinon, elle sera décrite par une texture LDR T . Pour simuler la réflexion de la lumière sur ces objets, nous nous servons naturellement des couleurs données par l'infographiste. Nous faisons l'hypothèse que l'infographiste a utilisé un écran pour régler les couleurs des objets, mais que son logiciel de modélisation ne prenait pas en compte le gamma de l'écran dans l'enregistrement des valeurs réelles (section 2.7.3). Pour les matériaux

aussi bien que pour les textures, nous créons des coefficients de réflexion sans dimension à partir des valeurs fixées par l'infographiste : M^γ et T^γ .

Il est également possible de combiner un matériau avec une texture. Mais selon le logiciel de modélisation et la combinaison fixée par l'infographiste, les valeurs du matériau et de la texture peuvent être additionnées, soustraites, multipliées ou moyennées. Il est donc impossible de faire un modèle général, valable dans tous les cas. Pour éviter cette incertitude sur les environnements virtuels destinés à un rendu HDR, nous avons fixé avec l'infographiste les règles suivantes :

- s'il y a un matériau combiné avec une texture, leurs valeurs sont multipliées ;
- s'il y a un matériau mais pas de texture, nous créons par défaut une texture LDR blanche dont les valeurs sont égales à 1 (il n'y a donc aucune influence de la texture créée sur le matériau) ;
- s'il y a une texture mais pas de matériau, nous créons par défaut un matériau LDR blanc dont les valeurs sont égales à 1 (il n'y a donc aucune influence du matériau créée sur la texture) ;
- s'il n'y a ni matériau ni texture, nous créons par défaut un matériau LDR et une texture LDR blancs, dont les valeurs sont égales à 1.

Ces règles nous permettent d'être certains d'avoir à la fois une texture et un matériau défini pour tout objet de la scène. La multiplication est donc l'opération retenue pour le traitement des matériaux et des textures des objets. Il est à noter qu'il existe en fait un coefficient de réflexion par type d'éclairage : ambiant (ρ_A), diffus (ρ_D) et spéculaire (ρ_S). Pour chaque objet éclairé de la scène, ces coefficients de réflexion sont donc définis par les formules suivantes :

$$\rho_A = (M_A T_A)^\gamma, \rho_D = (M_D T_D)^\gamma, \rho_S = M_S^\gamma \quad (5.10)$$

Les textures des termes ambiant et diffus correspondent à une texture de couleur, par exemple une photographie. Par contre, le terme spéculaire est indépendant de la couleur (ou texture) de l'objet, nous ne laissons donc que le matériau pour indiquer quelle quantité de lumière est réfléchiée par l'objet.

5.1.4 Modèle d'éclairage photométrique

Maintenant que les sources et les coefficients de réflexion des objets de la scène ont été définis, nous pouvons proposer un pipeline de rendu HDR. Ce pipeline définit les interactions entre les sources de lumière et les objets réfléchissants. De la même manière que dans le modèle de Phong [Phong 75], la luminance L produite par chaque

source en chaque pixel est égale à la somme des luminances des termes ambiant, diffus et spéculaire :

$$L = L_A + L_D + L_S \quad (5.11)$$

Pour une surface diffusante (lambertienne), la luminance est définie en fonction de l'éclairement E de chaque source et du coefficient de réflexion ρ de l'objet éclairé, selon la formue suivante :

$$L = \frac{E}{\pi} \rho \quad (5.12)$$

Le terme π est une constante due à l'intégration de la quantité de lumière atteignant une unité de surface. L'éclairement E en un point est égal au produit de l'intensité de la source, de son atténuation en fonction de la distance et/ou de l'angle par rapport à la normale de la surface.

La lumière ambiante est une lumière dont le point d'origine n'est pas défini, mais qui peut atteindre chaque pixel de chaque objet. La luminance produite par chaque source ambiante est donc égale au produit de l'intensité ambiante de la source, de l'atténuation de cette intensité en fonction de la distance, et du coefficient de réflexion ambiant :

$$L_A = \frac{I_A}{\pi d^2} (M_A T_A)^\gamma \quad (5.13)$$

Nous pouvons également tenir compte de l'éclairage du ciel dans un second terme ambiant, la luminance due au ciel étant donnée par la relation suivante :

$$L_A = \frac{E}{\pi} (M_A T_A)^\gamma \quad (5.14)$$

E représente l'éclairement produit par le ciel, en *lux*. Les valeurs moyennes d'éclairement pour des conditions d'éclairage données sont les suivantes [Wikipedia 10a, Zacharias 10] : 0.5 *lux* pour une nuit de pleine lune, de 25 000 à 30 000 *lux* pour une scène extérieure sous un ciel couvert (de jour), et 50 000 à 100 000 *lux* pour une scène extérieure en plein soleil.

La lumière diffuse est une lumière dont l'origine et la direction sont bien définies. La luminance produite par chaque source diffuse est égale au produit de l'intensité diffuse de la source, de l'atténuation en fonction de la distance et de l'angle par rapport à la normale de la source, et du coefficient de réflexion diffus :

$$L_D = \frac{I_D \cos(\theta)}{\pi d^2} (M_D T_D)^\gamma \quad (5.15)$$

Le terme $\cos(\theta)$ est l'angle d'incidence entre les vecteurs $-\mathbf{i}$ et \mathbf{n} (voir figure 5.4 pour les notations des vecteurs).

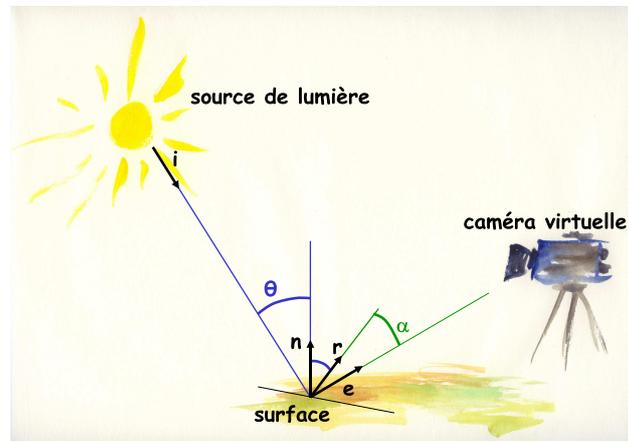


FIGURE 5.4 – Vecteurs normalisés pour les calculs d'éclairage.

Enfin, la lumière spéculaire est une lumière dont l'origine et la direction réfléchi sont bien définies. Nous utilisons la formule du modèle de Phong dans notre modèle. La luminance produite par chaque source spéculaire est égale au produit de l'intensité spéculaire de la source, de l'atténuation en fonction de l'angle par rapport à la position de la caméra, et du coefficient de réflexion diffus :

$$L_S = \frac{I_S \cos(\alpha)^{sh}}{\pi d^2} M_S^\gamma \quad (5.16)$$

où $\cos(\alpha) = \mathbf{r} \cdot \mathbf{e}$, avec \mathbf{r} le vecteur réfléchi, et \mathbf{e} le vecteur pointant vers la caméra (voir figure 5.4). Le terme sh reflète le degré de spécularité de la surface : plus celui-ci est élevé, et plus le reflet spéculaire sera étroit. Il est réglé par défaut à 64.

Nous avons maintenant réuni toutes les composantes pour former le pipeline de rendu HDR, selon la figure 5.5. Les termes ambiant et diffus sont additionnés, avant de les multiplier par la texture de couleurs. Le terme spéculaire est ensuite ajouté. Enfin, ce sont les luminances des sources de lumière qui sont ajoutées, pour obtenir l'image de synthèse HDR.

5.1.5 Optimisation des calculs d'éclairage

Les environnements virtuels routiers nocturnes présentent généralement de nombreuses sources lumineuses, comme des lampadaires ou des phares de voiture. Si nous

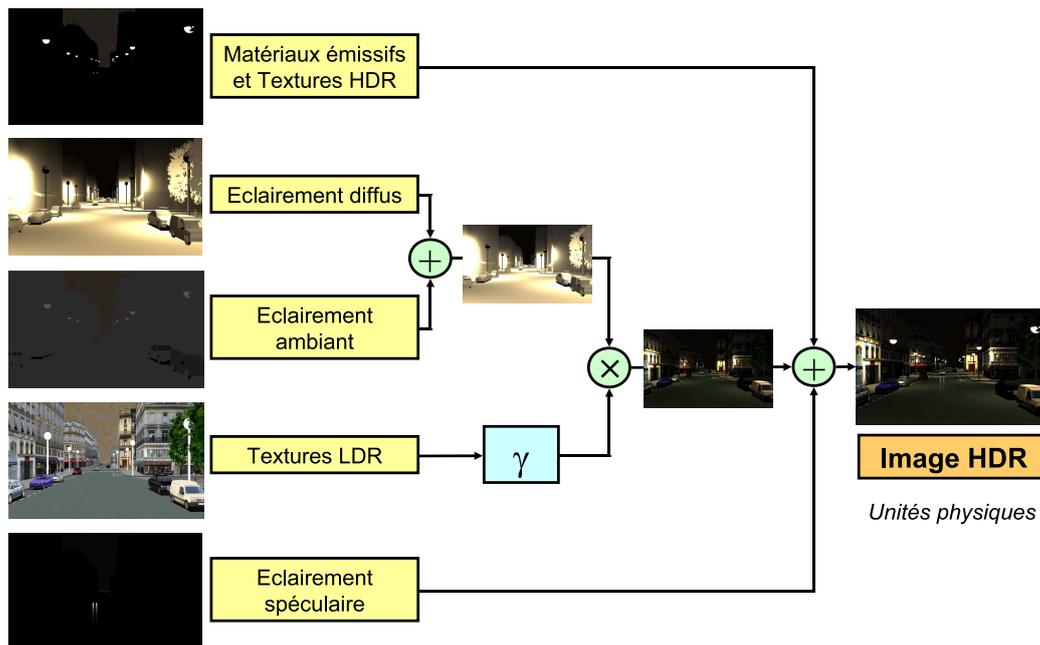


FIGURE 5.5 – Pipeline de rendu HDR pour des environnements virtuels, basé sur un modèle physique d'éclairage. Base de données 3D INRETS-MSIS-2007.

ne faisons pas attention, le temps de calcul dû à l'éclairage augmente très vite avec le nombre de sources, et peut perturber l'exécution du programme. Pour des conditions optimales d'utilisation d'un simulateur de conduite, le rendu doit être au minimum de 30 images par seconde.

L'utilisation d'un nombre élevé de sources ne permet pas d'atteindre de telles performances par défaut. A titre indicatif, *OpenGL* limite le nombre de sources à 8 pour garantir l'exécution en temps réel. Des solutions pour éclairer l'environnement virtuel peuvent être utilisées, comme l'utilisation de cartes de luminances (ou *light maps*), mais ces cartes doivent être précalculées, et ne peuvent pas être modifiées dynamiquement lorsqu'un objet passe entre la source et un point.

Nous devons donc optimiser les calculs d'éclairage pour pouvoir utiliser plusieurs dizaines de sources dans l'environnement virtuel. Si nous considérons une scène avec N sources de lumière ($N > 8$), les N sources participent par défaut à l'éclairage de chaque pixel, quelle que soit la distance entre le pixel éclairé et chacune des N sources. Si une source est éloignée du pixel considéré, son éclairage est faible, et peut être négligé. Une première optimisation consiste donc à négliger la participation d'une source à l'éclairage d'un pixel lorsque cet éclairage est inférieur à un seuil d'éclairage

E_{min} , fixé par l'utilisateur en fonction du niveau d'éclairement moyen de la scène. Une fois ce seuil fixé, l'éclairage d'une source n'est plus calculé que pour les distances inférieures à la distance d , fonction de l'intensité maximale de la source I_D^{max} et du seuil d'éclairement E_{min} :

$$d = \sqrt{\frac{I_D^{max}}{E_{min}}} \quad (5.17)$$

Dans la première version de notre application, toutes les sources pouvaient participer à l'éclairage de chaque pixel. Le temps de calcul de l'éclairage était alors proportionnel à N . Cette optimisation permet de réduire fortement le nombre de sources qui éclairent chaque pixel, puisqu'un pixel n'est éclairé que si son éclairement est significatif. Le temps de calcul est alors bien moindre : il est proportionnel au nombre de sources qui éclairent chaque pixel (si l'on néglige le calcul de chaque distance source-pixel, ce qui n'est en réalité pas tout à fait négligeable, voir section 5.1.7). A titre d'illustration, la limite du temps réel se trouve maintenant à environ 20 sources diffuses par pixel dans notre implémentation. En outre, l'avantage des environnements routiers est que les sources statiques comme les lampadaires sont régulièrement espacées, ce qui permet de rester sous cette limite de 20 sources par pixel, au moins du point de vue des sources statiques. La présence de voitures peut cependant augmenter le nombre de sources par pixel, notamment au niveau des carrefours.

En pseudo-code, l'optimisation proposée consiste à exécuter le schéma suivant :

```
for (int i = 0 ; i < NBRE_SOURCES ; i++)
{
    distance = calculeDistance(positionPixel, positionSource);
    d = sqrt(Idiffuse/Emin);
    if (distance < d)
        calculeTermesAmbiantEtDiffus(...);
    calculeTermeSpeculaire(...);
}
```

Une seconde optimisation porte sur les phares de voiture, qui n'éclairent que vers l'avant : il est donc inutile de calculer la participation des phares sur des pixels se trouvant derrière ces phares. De plus, pour le terme spéculaire, le reflet sur la route se trouvera forcément dans l'axe œil-caméra, il est donc tout aussi inutile de calculer un terme spéculaire lié aux phares en dehors de cet axe.

Si ces optimisations restent insuffisantes, nous pouvons augmenter la valeur du seuil limite E_{min} , au risque de faire apparaître des démarcations sur les zones éclairées

par les sources. Une autre solution consiste à augmenter la distance entre les sources statiques, pour limiter le nombre de sources par pixel, mais au risque de dégrader le réalisme de la scène. Sinon, nous n'avons pas d'autre choix que de supprimer des termes d'éclairage, comme les termes ambiant et/ou spéculaire. Ces suppressions risquent de dégrader l'ensemble du rendu visuel, mais pour ce type d'application, le temps réel prime avant tout.

5.1.6 Résultats

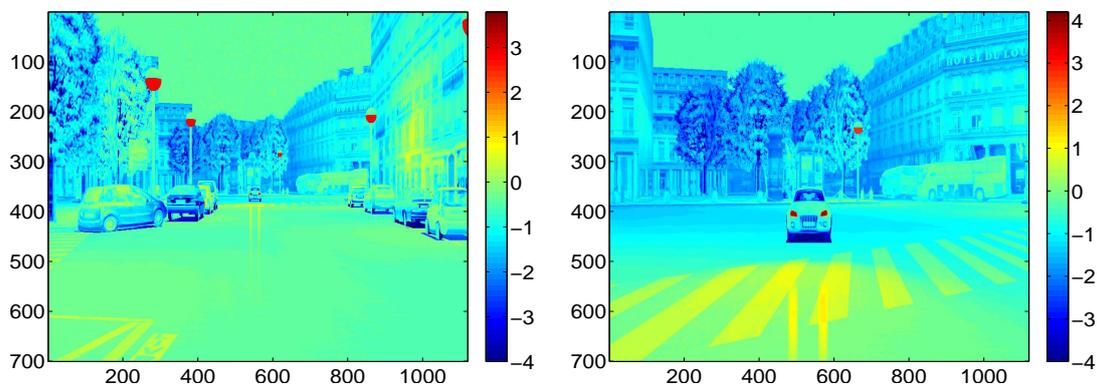


FIGURE 5.6 – Images de synthèse HDR (logarithme décimal de la luminance). Base de données 3D INRETS-MSIS-2007.

La figure 5.6 montre des images HDR en luminances, produites par le pipeline de rendu HDR décrit. Les luminances au sol sont d'environ 1 cd/m^2 , alors que celles du décor sont plus faibles puisqu'il n'est pas éclairé. Certaines zones de la route de l'image de droite ont des luminances plus faibles, car il n'y a pas de lampadaire à proximité. Le terme spéculaire produit enfin des luminances plus élevées que l'éclairage diffus.

Nous pouvons également remarquer que l'*aliasing* influence la valeur des luminances, et notamment la valeur maximale des sources lumineuses, ici contenues dans les phares. C'est important par rapport à la reproduction de tons, car certains opérateurs utilisent la valeur maximale de l'image comme paramètre pour calculer la dynamique de luminance à compresser. Or si celle-ci varie au cours du temps à cause de l'*aliasing*, il pourra apparaître de brusques changements sur la luminosité de l'image produite, car la dynamique de luminance à compresser changera constamment au cours de la simulation. A titre indicatif, la luminance maximale est de $6\,100 \text{ cd/m}^2$ dans l'image de

gauche et de $15\,600\text{ cd/m}^2$ dans l'image de droite.

Enfin, ce modèle ne prend pas en compte des techniques plus sophistiquées pour calculer le terme ambiant (avec de l'*ambient occlusion* par exemple), mais elles peuvent être facilement intégrées en remplaçant le terme ambiant calculé précédemment. Les termes diffus et spéculaire pourraient être également améliorés par l'utilisation de BRDF (*bidirectional reflectance distribution function*). Si la BRDF permet de décomposer le calcul des termes diffus et spéculaire, alors le même modèle pourra être appliqué, sinon, nous n'aurons pas d'autre choix que de traiter les termes diffus et spéculaire ensemble, et donc de multiplier ces deux termes par la texture de couleurs.

Une fois l'image HDR rendue, elle doit être stockée dans une texture 32 bits, pour éviter que les valeurs physiques en luminance ne soient détériorées, ou pire, seuillées entre 0 et 1. Elle sera ensuite traitée par un algorithme de simulation de l'éblouissement et par un opérateur de reproduction de tons, avant d'être affichée à l'écran.

5.1.7 Performances

Les performances des algorithmes sont mesurées via *OpenSceneGraph*, en affichant le nombre d'images rendues par seconde. Pour se synchroniser avec le dispositif d'affichage, *OpenSceneGraph* considère généralement 3 cas :

- si le rendu est très rapide, ou *strictement temps réel* ($\geq 60\text{ Hz}$), *OpenSceneGraph* limite le rendu à 60 images par seconde, et dans ce cas là nous affichons ≥ 60 dans les performances ;
- si le rendu est rapide, ou *juste temps réel* (entre 30 et 60 Hz), *OpenSceneGraph* limite généralement le rendu à 30 images par seconde, mais ce chiffre peut parfois être supérieur. Dans ce cas là, nous affichons dans les performances la valeur minimum relevée ;
- si le rendu est lent, ou *non temps réel*, ($< 30\text{ Hz}$), *OpenSceneGraph* affiche le nombre réel d'images rendues par seconde, généralement variable. Nous reportons également la valeur minimum dans les performances.

Nous étudions ici le temps d'exécution du rendu de l'image HDR dans une texture 32 bits. Ce temps d'exécution dépend fortement du nombre de pixels de l'image HDR à traiter. Une image HD (1680×1050 pixels) sera ainsi plus longue à traiter qu'une image non-HD (1024×768 pixels).

En faisant l'hypothèse que les temps nécessaires pour calculer les composantes diffuse et spéculaire sont approximativement égaux, nous ne nous intéressons ici qu'à l'influence des sources diffuses. Dans le cas où les composantes diffuse et spéculaire d'une source doivent être calculées, cette source devra être comptée deux fois par rapport aux temps de calculs présentés ci-dessous.

Nombre de sources diffuses à l'écran	Nombre d'images par seconde
de 1 à 21	≥ 60
de 22 à 50	≥ 30
de 51 à 80	≥ 20

TABLE 5.1 – Performances du rendu HDR en fonction du nombre de sources diffuses à l'écran.

Le tableau 5.1 illustre les performances pour le cas le plus difficile : celui où toutes les sources sont à l'écran et participent à l'éclairage de la scène observée¹. Quelle que soit la position de l'observateur, le temps réel est strictement maintenu jusqu'à 21 sources. Jusqu'à 50 sources, on reste au-dessus des 30 images par seconde, mais certains ralentissements peuvent être perceptibles et gêner l'observateur dans sa perception de l'environnement virtuel, notamment lors de virages rapides. Au delà de 50 sources, le temps réel n'est clairement plus maintenu.

Nombre total de sources diffuses de la scène, en ayant 15 sources à l'écran	Nombre d'images par seconde
de 15 à 28	≥ 60
de 29 à 80	≥ 30

TABLE 5.2 – Performances du rendu HDR en fonction du nombre total de sources diffuses de la scène, en ayant 15 sources à l'écran.

Le tableau 5.2 permet de montrer les bénéfices de l'optimisation apportée sur l'éclairage, qui permet d'éviter le calcul des composantes des sources éloignées, en tenant compte de la distance de chaque source (section 5.1.5). Nous pouvons voir que si nous nous fixons la limite de 15 sources dans le champ de vision, le nombre total de sources dans la scène peut aller jusqu'à 28 en restant à au moins 60 images par seconde,

1. Attention : c'est bien le calcul de l'éclairage dû aux sources qui est relevé dans le temps d'exécution. Le temps d'exécution nécessaire pour l'affichage des sources elles-mêmes est bien inférieur, puisqu'il n'y a qu'un matériau émissif ou une texture HDR à afficher.

et peut aller jusqu'à 80 sources en restant à au moins 30 images par seconde. On voit donc que le nombre total de sources influence toujours le temps de calcul, mais dans une moindre mesure que si ces deux grandeurs étaient directement proportionnelles.

Résolution de l'image, avec 80 sources diffuses à l'écran	Nombre d'images par seconde
1680 × 1050	≥ 20
1600 × 900	≥ 20
1280 × 1024	≥ 23
1152 × 864	≥ 30
1024 × 768	≥ 30
800 × 600	≥ 60

TABLE 5.3 – Performances du rendu HDR en fonction de la résolution de l'image, avec 80 sources diffuses à l'écran.

Enfin, le tableau 5.3 montre l'influence de la résolution sur le temps d'exécution. Nous pouvons voir que des petites résolutions donnent de bien meilleures performances, ce qui est tout à fait normal puisqu'il y a moins de pixels sur lesquels les calculs doivent être réalisés. Les ordres de grandeur sont respectés : une image de résolution 1680 × 1050 a environ 3 fois plus de pixels qu'une image de résolution 800 × 600, et met bien environ 3 fois plus de temps à être calculée.

5.2 Visualisation réaliste d'images de synthèse HDR

L'image HDR produite contient des valeurs de luminance stockées dans une texture. Nous allons maintenant traiter cette image pour l'afficher à l'écran, en utilisant un algorithme de simulation de l'éblouissement et un opérateur de reproduction de tons. Le but de ces traitements est d'afficher l'image à l'écran telle qu'elle serait perçue dans la réalité, malgré les limites des écrans (cf section 2.7). Des exemples ont déjà été proposés dans la littérature ([Durand 00, Krawczyk 05]), mais nous voulons aller plus loin dans la fidélité des images affichées par rapport au contenu des images HDR de synthèse. Dans cette section, nous choisissons des algorithmes appropriés pour notre application, nous expliquons comment les implémenter en temps réel, et nous présentons les résultats obtenus.

5.2.1 Choix d'un algorithme de simulation de l'éblouissement

La simulation de l'éblouissement répond à deux objectifs, qui sont par ordre d'importance pour la visualisation réaliste :

- reproduire à l'écran l'éblouissement tel qu'il serait vu par l'observateur dans la réalité (avec la luminance de voile, les stries et le halo lenticulaire) ;
- donner l'impression à l'observateur que la zone regardée est plus lumineuse que ce qui est affiché.

Dans la section 4.4, nous avons présenté 4 algorithmes de simulation de l'éblouissement. L'algorithme offrant le meilleur rendu visuel est clairement celui de Ritschel et *al.* [Ritschel 09]. Mais il est aussi le plus difficile à implémenter, et le plus long en terme de temps de calcul. Notre temps pour développer ces algorithmes étant très limité, nous nous sommes tournés vers les deux algorithmes les plus rapides à implémenter : les filtres gaussiens de Kawase [Kawase 04], et la *point-spread-function* de Spencer [Spencer 95].

5.2.2 Implémentation de la simulation de l'éblouissement

Implémentation des filtres gaussiens de Kawase

La simulation de l'éblouissement par des filtres gaussiens a été implémentée de la manière suivante : d'abord, nous éliminons les pixels qui ne participent pas à l'éblouissement. Un simple seuillage ne permet pas de réaliser cette opération, car le traitement est trop brutal et peut faire apparaître des artefacts au cours du temps. Nous avons donc préféré créer la fonction empirique suivante, plus lisse, et permettant d'éviter tout artefact :

$$pixelsLumineux = \begin{cases} luminance & \text{si } luminance > seuil \\ \frac{luminance \times seuil - \frac{seuil^2}{50}}{seuil - \frac{seuil}{50}} & \text{si } luminance > \frac{seuil}{50} \\ 0 & \text{sinon} \end{cases} \quad (5.18)$$

La valeur du seuil peut être réglée soit à une valeur fixe par l'utilisateur, soit en fonction de la luminance d'adaptation L_a du système visuel humain modélisé. La figure 5.7 illustre les valeurs des luminances transmises pour un seuil de 500 cd/m^2 .

L'image produite ne contient alors que les luminances qui peuvent créer une luminance de voile pour l'œil. Nous avons ensuite utilisé 3 filtres gaussiens successifs pour créer la luminance de voile, en utilisant la propriété séparable du filtre pour réduire le temps de calcul. Ce sont donc 6 convolutions qui sont effectuées successivement.

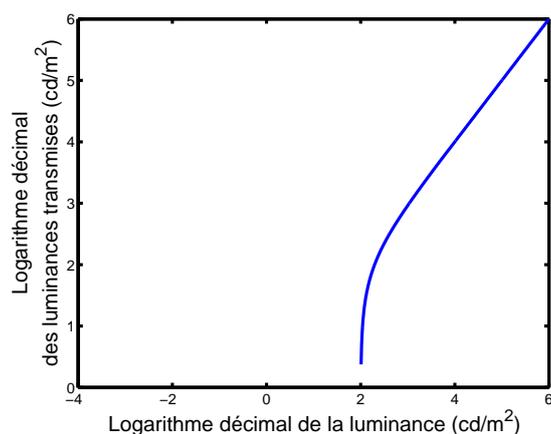


FIGURE 5.7 – Logarithme décimal des luminances retenues pour participer à la simulation de l'éblouissement, en fonction du logarithme décimal des luminances de l'image HDR.

Ensuite, nous additionnons le résultat de chaque convolution, selon les poids suivants :

$$\text{Luminance de voile} = \frac{8}{14}G(\sigma_1) + \frac{2}{14}G(\sigma_2) + \frac{4}{14}G(\sigma_3) \quad (5.19)$$

où $G(\sigma_X)$ représente un filtre gaussien de taille X , avec $\sigma_2 = 2\sigma_1$ et $\sigma_3 = 2\sigma_2$.

L'ensemble de la simulation de l'éblouissement avec des filtres gaussiens est reproduit sur la figure 5.8. Pour accélérer le temps de calcul, nous utilisons la technique du *mipmapping* pour disposer d'une image de résolution inférieure. En pratique, nous utilisons l'image donnée par le premier niveau de la pyramide, où la résolution est divisée par 4 (comme illustré sur la figure 5.8). Nous interpolons ensuite les images finales à l'échelle de l'image HDR. Un étage plus élevé peut être utilisé pour augmenter les performances, mais des artefacts visuels apparaîtront au niveau des halos entourant les sources, dus à l'interpolation finale. Quel que soit l'étage utilisé, nous avons fixé la valeur σ selon la formule 5.20, où n représente l'étage de l'image utilisée sur la pyramide :

$$\sigma = \frac{5}{2^n} \quad (5.20)$$

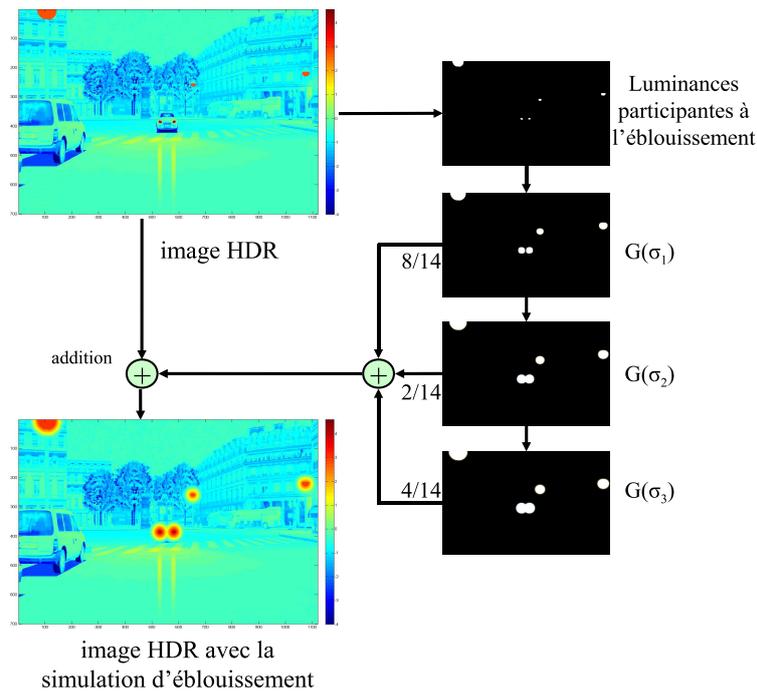


FIGURE 5.8 – Simulation de l'éblouissement avec 3 filtres gaussiens. D'après [Kawase 04]. Base de données 3D INRETS-MSIS-2007.

Implémentation des filtres de Spencer et *al.*

La simulation de l'éblouissement de Spencer est plus difficile à implémenter, et nécessite trois étapes. La première consiste à calculer la taille de la pupille à partir de la luminance moyenne de l'image HDR. La pupille s'adaptant quasi-instantanément à la quantité de lumière entrante, sa taille peut être directement déterminée à partir de l'équation 3.3. La seconde étape consiste à calculer le filtre par lequel les pixels les plus lumineux de l'image HDR seront convolués. Spencer et *al.* proposent 4 fonctions pour simuler les différentes conséquences visuelles de l'éblouissement (la luminance de voile, la couronne ciliaire et le halo lenticulaire) :

$$f_0 = 2.61 \cdot 10^6 e^{(-\frac{\theta}{0.02})^2} \quad (5.21)$$

$$f_1 = \frac{20.91}{(\theta + 0.02)^3} \quad (5.22)$$

$$f_2 = \frac{72.37}{(\theta + 0.02)^2} \quad (5.23)$$

$$f_{3R,V,B} = \frac{248159}{\lambda_{R,V,B}} e^{-(\theta - 3 \frac{\lambda_{R,V,B}}{568})^2} \quad (5.24)$$

Selon le domaine dans lequel se trouve l'œil à l'instant t (scotopique, mésopique ou photopique), Spencer et *al.* proposent d'attribuer différents poids à chaque fonction. Nous avons cependant préféré utiliser une interpolation linéaire de ces poids en fonction de la taille de la pupille, proposée par Durand et Dorsey [Durand 00] :

$$w_1 = 0.478 \quad (5.25)$$

$$w_2 = 0.138 + 0.08 \frac{D_{pupille} - 2}{5} \quad (5.26)$$

$$w_3 = 0.033 \frac{D_{pupille} - 2}{5} \quad (5.27)$$

$$w_0 = 1 - w_1 - w_2 - w_3 \quad (5.28)$$

Le filtre non-normalisé s'obtient finalement par l'opération suivante :

$$filtre_{R,V,B} = w_0 f_0 + g w_1 f_1 + g w_2 f_2 + g w_3 f_{3R,V,B} \quad (5.29)$$

où g représente une image de stries (figure 4.11). Le filtre est ensuite normalisé.

La troisième étape consiste à convoluer ce filtre avec les pixels lumineux de l'image HDR. Pour avoir un rendu visuel satisfaisant, nous avons fixé la taille du filtre à 100×100 pixels. Une convolution directe avec un *pixel shader* ne permet de rendre qu'une image toutes les 5 à 10 secondes, ce qui est insuffisant pour une application temps réel. Nous devons donc utiliser la transformée de Fourier pour obtenir un meilleur temps d'exécution. En effet, la transformée de Fourier d'une convolution de deux images est le produit des transformées de Fourier de ces deux images, et dans le cas où les images ont une résolution élevée, l'utilisation de la transformée de Fourier permet de réduire le temps de calcul nécessaire. Il faut toutefois faire attention lors de son utilisation, car ici les images n'ont pas la même résolution.

L'ensemble de la simulation de l'éblouissement par Spencer et *al.* est reproduit sur la figure 5.9.

Les deux simulations de l'éblouissement présentées dans cette section seront évaluées dans la section 6.2.

5.2.3 Choix d'un opérateur de reproduction de tons

Dans la section 4.2, nous avons présenté 21 opérateurs de reproduction de tons, et nous devons en choisir un pour notre application, la simulation de conduite. En étudiant les différentes évaluations de ces opérateurs (section 4.3), nous avons vu qu'aucun opérateur n'était plus performant que tous les autres, mais que chacun a des avantages et/ou

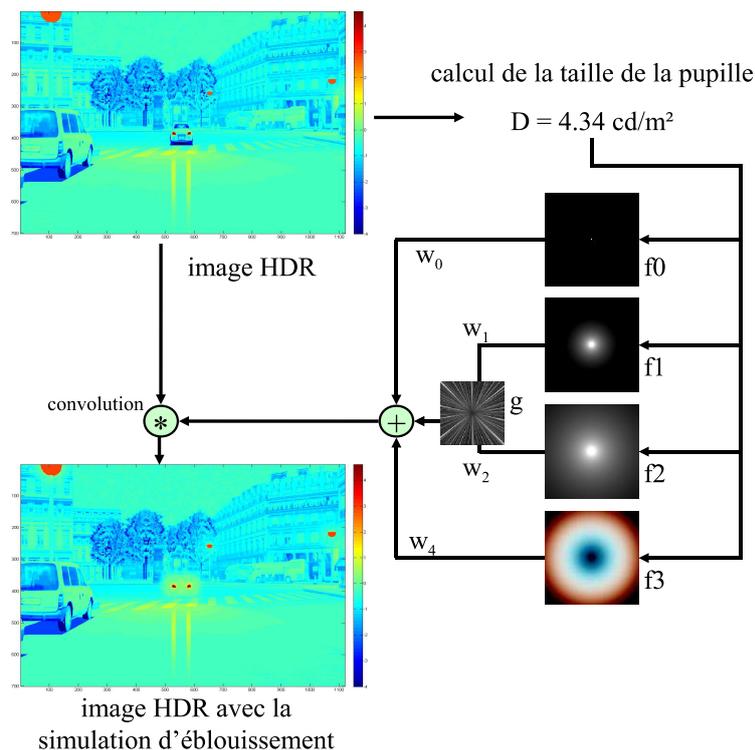


FIGURE 5.9 – Simulation de l'éblouissement, basée sur la *point-spread-function* de Spencer et al. D'après [Spencer 95]. Base de données 3D INRETS-MSIS-2007.

des inconvénients que d'autres non pas. Nous sommes arrivés à la conclusion qu'**un opérateur de reproduction de tons doit être choisi en fonction de son application** [Grave 08]. Pour la simulation de conduite, l'opérateur devra être :

- un opérateur qui préserve les *contrastes* juste perçus à l'écran, pour avoir une détection identique à l'écran et dans la réalité ;
- un opérateur qui essaie de reproduire l'image telle qu'elle est perçue dans la réalité à l'écran (et qui ne se contente donc pas de réduire la dynamique de luminance indépendamment de la perception humaine). Ces opérateurs sont ceux basés sur un *modèle visuel humain* ;
- un opérateur adapté au traitement de vidéos. Il sera doté d'un modèle d'*adaptation temporelle*, pour simuler les changements de luminosité à l'écran comme dans la réalité, et ne traitera donc pas les images indépendamment les unes des autres ;
- un opérateur *automatique*, qui n'a pas besoin de réglage pour traiter les vidéos ;
- un opérateur exécutable en *temps réel*.

	Modèle visuel humain	Contraste	Adaptation temporelle	Automatique	Temps réel
Linéaire		x		x	x
[Chiu 93]					x
[Tumblin 93]	x			x	x
[Schlick 94]					x
[Ward 94]	x	x		x	x
[Ferwerda 96]	x	x		x	x
[Rahman 96]	x				x
[Ward 97]	x	x		x	x
[Pattanaik 98]	x			x	
[Durand 00]	x	x	x	x	x
[Pattanaik 00]	x		x		x
[Durand 02]				x	
[Fairchild 02]	x			x	
[Fattal 02]				x	
[Reinhard 02]				x	x
[Choudhury 03]				x	
[Drago 03]	x			x	x
[Ledda 04]	x		x	x	
[Irawan 05]	x	x	x	x	x
[Reinhard 05a]	x				x
[Grave 08]	x	x		x	

TABLE 5.4 – Autres caractéristiques des principaux opérateurs de reproduction de tons

La table 4.1 résumait les caractéristiques des différents opérateurs présentés. La table 5.4 résume maintenant les caractéristiques qui nous intéressent pour notre application. La colonne « Temps réel » a été remplie d'après les performances annoncées par les auteurs respectifs et/ou d'après la complexité des opérateurs. Les opérateurs utilisant un filtrage bilatéral n'étaient pas temps réel lors de leur publication, cependant il est possible qu'ils le soient devenus aujourd'hui.

Parmi les opérateurs présentés, seuls deux répondent aux critères demandés : l'opérateur global de Durand et Dorsey [Durand 00] et celui d'Irawan et *al.* [Irawan 05]. L'opérateur de Durand et Dorsey est un opérateur linéaire qui exploite directement la fonction des seuils juste détectables en fonction de l'intensité (*threshold versus inten-*

sity function). Mais l'opérateur d'Irawan propose en outre une modélisation plus complète de l'adaptation temporelle, étend la fonction des seuils juste détectables en une fonction qui dépend de l'état d'adaptation et de la luminance observée, et propose une compression non-linéaire de l'ajustement d'histogramme pour optimiser l'utilisation de la gamme de luminances offerte par l'écran.

Le fait de proposer le modèle le plus complet d'adaptation temporelle, et d'utiliser une fonction des seuils juste détectables en fonction de l'état d'adaptation **et** de la luminance observée est particulièrement important pour la simulation de conduite. En effet, lorsqu'il circule, le conducteur s'adapte constamment aux luminances qui lui font face, et n'est donc jamais correctement adapté si celles-ci changent au cours du temps. Enfin, le fait qu'il soit un opérateur perceptuel est un avantage supplémentaire pour notre application, dont le but est clairement de représenter à l'écran la réalité telle qu'elle est perçue.

Lorsqu'il a fallu choisir un opérateur de reproduction de tons, nous avons donc sélectionné l'opérateur d'Irawan et *al.*, car c'est celui qui correspondait le mieux aux critères requis pour notre application. Il ne restait qu'à l'implémenter en temps réel, car les auteurs annonçaient un traitement de 5 à 10 secondes par image.

5.2.4 Implémentation de l'opérateur d'Irawan

La reproduction de tons est une étape de post-traitement dans le rendu HDR. Cette étape est réalisée à partir de l'image HDR, à laquelle nous avons ajouté le résultat de la simulation de l'éblouissement. L'opérateur d'Irawan a été implémenté en utilisant 9 shaders en cascade, chacun utilisant le résultat du précédent. Nous avons vu dans l'état de l'art (section 4.2.2) que l'algorithme d'Irawan pouvait être décomposé en 3 étapes :

- le calcul de l'état d'adaptation du système visuel humain, en utilisant un modèle de vision dynamique ;
- le calcul des seuils de luminance juste détectables, pour l'état d'adaptation déterminé à l'étape précédente, et pour chaque luminance observée ;
- l'ajustement d'histogramme, effectué avec une contrainte pour ne pas exagérer les contrastes.

Nous avons vu que pour produire l'image finale à afficher, l'opérateur d'Irawan calculait deux images, celle perçue par les cônes et celle perçue par les bâtonnets. Dans le cadre de la simulation de conduite de nuit, suivant les conditions d'éclairage, et sous l'hypothèse (raisonnable !) que le conducteur a toujours ses phares allumés,

le domaine de perception du conducteur peut être soit le photopique, soit le haut mésopique. Dans cette implémentation, nous négligeons la contribution des bâtonnets, et nous nous limitons au calcul de l'image perçue par les cônes.

Les 9 shaders implémentant l'opérateur d'Irawan sont organisés de la manière suivante (voir la figure 5.10 pour une vue d'ensemble) :

- Le premier shader convertit l'image HDR en une image de luminance L et deux autres canaux de couleurs (X et Z) de l'espace de couleurs XYZ. Puis, le logarithme décimal de la luminance L est calculé pour anticiper le shader suivant.
- Le second shader calcule la valeur moyenne du logarithme décimal de la luminance, $\log_{10}(L_m)$, qui indique la quantité de lumière que reçoit l'œil à l'instant t . La valeur moyenne est calculée en utilisant le procédé de *mipmapping* sur le logarithme de la luminance. Ce procédé construit une pyramide d'images dont la résolution diminue par 2 à chaque nouvel étage, jusqu'à obtenir une image de taille 1×1 pixel. Le plus haut niveau de la pyramide construite par le *mipmapping* donne donc la valeur moyenne sur toute l'image du logarithme décimal de la luminance.
- Le troisième shader calcule l'état d'adaptation de l'œil, en fonction de la luminance perçue à l'instant t , mais aussi de l'état d'adaptation antérieur de l'œil, dont la valeur d'adaptation à $t - 1$ est stockée dans un buffer. Nous utilisons une première fois les équations du modèle d'adaptation d'Irawan pour calculer l'état d'adaptation face à l'image entière, noté $\sigma_{C\hat{ones}Global}$ dans la description de l'opérateur d'Irawan et *al.* section 4.2.3. Ensuite, nous construisons l'échelle des luminances qui sera utilisée en abscisse pour l'ajustement d'histogramme, sur la gamme des luminances perçue par l'œil humain, de 10^{-4} à 10^6 cd/m^2 . Chaque valeur de l'image de luminances sera donc répartie dans une case de cette échelle de luminance. On utilise une seconde fois les équations du modèle d'adaptation, pour chaque luminance considérée en abscisse. Nous obtenons ainsi les valeurs d'adaptation locale $\sigma_{C\hat{ones}Local}$ face à chaque luminance observée (données par l'échelle de luminance), calculées à partir de la valeur d'adaptation globale de luminance $\sigma_{C\hat{ones}Global}$.
- Les quatrième et cinquième shaders calculent les seuils de luminances juste détectables, pour chaque état d'adaptation local calculé précédemment ($\sigma_{C\hat{ones}Local}$) et pour chaque luminance observée correspondante. Le calcul de ces seuils nécessite deux shaders successifs : le quatrième shader calcule la quan-

tité minimale de réponse additionnelle ΔR pour que l'œil perçoive une différence de luminance ΔL . La quantité minimale ΔR est ensuite utilisée dans le cinquième shader pour calculer les seuils de luminance juste détectables, selon l'équation 4.66.

- Le sixième shader calcule l'histogramme de l'image HDR, à partir d'une image de luminance de faible résolution (de l'ordre de 105×66 pixels) provenant du quatrième étage de la pyramide du *mipmapping* (second shader). Dans ce même shader, nous calculons également les seuils à ne pas dépasser lors de l'ajustement d'histogramme, à partir des valeurs des seuils juste détectables obtenus dans le shader précédent.
- Le septième shader réalise l'ajustement d'histogramme. Nous calculons le nombre de cases où le seuil n'est pas atteint $N_{casesNonPleines}$, ainsi que le nombre total de pixels qui dépassent le seuil autorisé $N_{pixelsARedistribuer}$, afin de les redistribuer dans les cases où le seuil n'est pas atteint. Chaque case non-remplie reçoit ensuite $N_{pixelsRedistribués}$ pixels, tel que :

$$N_{pixelsRedistribués} = \frac{N_{pixelsARedistribuer}}{N_{casesNonPleines}} \quad (5.30)$$

Chaque case non-remplie conserve ensuite les pixels redistribués, à la condition que le nouveau nombre de pixels ne dépasse le seuil autorisé.

- Le huitième shader calcule l'histogramme cumulé, à partir de l'histogramme ajusté obtenu du shader précédent.
- Le neuvième shader réalise finalement la reproduction de tons : les luminances HDR sont compressées en luminances LDR en utilisant la fonction donnée par l'histogramme cumulé, et qui est normalisée entre 0 et 1 (en utilisant la dernière valeur de l'histogramme cumulé). Pour réduire le nombre de cases à utiliser lors de l'ajustement d'histogramme, nous utilisons une interpolation linéaire de la fonction donnée par l'histogramme cumulé normalisé, ce qui permet d'éviter que plusieurs luminances HDR « proches » se retrouvent dans la même case lors de l'ajustement d'histogramme, et avec la même luminance LDR à l'écran. Enfin, les couleurs sont ajoutées à l'image de luminance en utilisant les canaux *X* et *Z* du premier shader, ce qui permet d'afficher l'image à l'écran dans l'espace de couleur *RGB*.

La figure 5.10 résume les différentes étapes de la reproduction de tons par l'opérateur d'Irawan et *al.* Nous pouvons remarquer que les algorithmes pour la simulation

de l'éblouissement utilisent la moyenne des luminances HDR pour calculer la taille de la pupille, alors l'algorithme d'Irawan et *al.* l'utilise dans son modèle de vision. En pratique, pour optimiser les performances, nous avons donc inséré l'algorithme pour la simulation de l'éblouissement dans celui pour la reproduction de tons, entre les shaders 2 et 3.

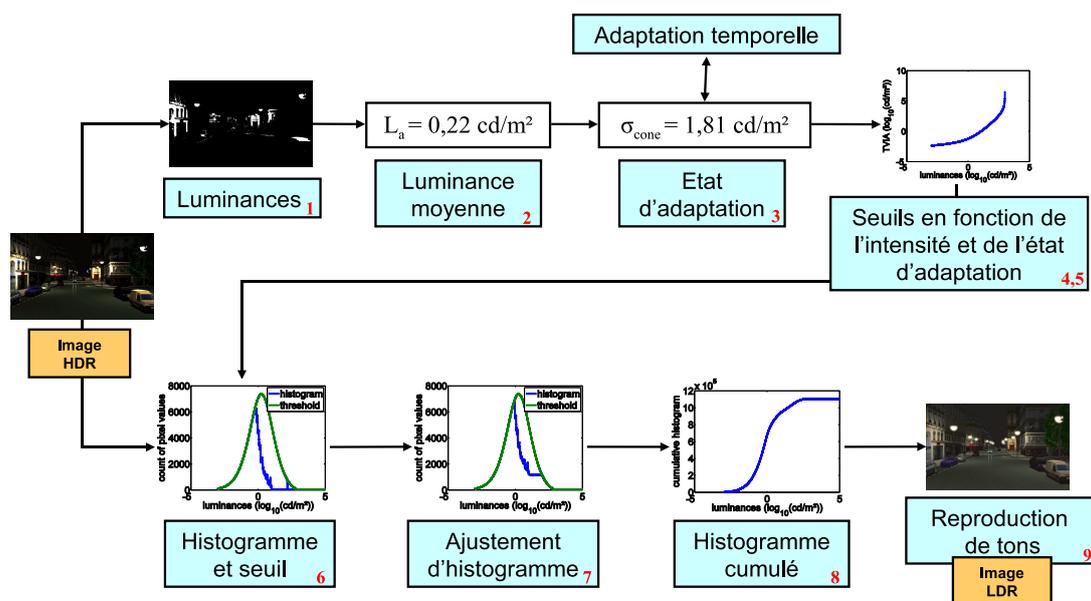


FIGURE 5.10 – Le pipeline de post-traitement proposé implémente l'opérateur d'Irawan et *al.* avec 9 shaders en cascade, pour reproduire sur un écran LDR les tons d'une image HDR. Base de données 3D INRETS-MSIS-2007.

5.2.5 Résultats

Nous présentons dans cette section des images du simulateur de conduite, de nuit, obtenues en utilisant le pipeline de rendu HDR présenté dans la section 5.1.4 et le pipeline de post-traitement de la section précédente. Nous utilisons la simulation de l'éblouissement de Kawase [Kawase 04] et l'opérateur de reproduction de tons d'Irawan et *al.* [Irawan 05]. L'environnement virtuel était initialement une modélisation LDR d'un quartier de Paris, de jour [Giannopulu 08]. Pour utiliser cet environnement de nuit, nous avons uniquement ajouté des lampadaires, puis réglé les intensités des sources de lumière (lampadaires et phares) avec des valeurs physiquement plausibles, en unités physiques. Le rendu HDR, la simulation de l'éblouissement et l'opérateur de reproduction de tons opèrent ensuite automatiquement et sans réglage de l'utilisateur.



FIGURE 5.11 – Captures d'écran du rendu HDR appliqué à la simulation de conduite, sur un environnement virtuel modélisant un quartier de Paris. Base de données 3D INRETS-MSIS-2007.

Un second environnement virtuel a également été utilisé : il modélise un tunnel de l'autoroute A86 modélisé par le LIVIC [Gruyer 10]. Comme précédemment, les seules modifications apportées à l'environnement virtuel concernent l'ajout de sources de lumière et la configuration de leurs intensités.

5.2.6 Performances

Les temps d'exécution de la simulation de l'éblouissement et de la reproduction de tons dépendent de la résolution de l'image à traiter, mais également d'autres paramètres. Dans un premier temps, pour ne pas tenir compte du temps nécessaire pour le rendu de l'image HDR, nous présentons ici les performances mesurées sur le traitement d'une image HDR enregistrée et directement affichée dans une texture HDR.

Pour notre implémentation, le temps d'exécution de la simulation de l'éblouissement ne dépend que de la résolution des images qui sont convoluées, et du nombre et de la taille des filtres gaussiens utilisés. Le temps d'exécution est indépendant du nombre de sources à l'écran, puisque l'image entière est convoluée par chaque filtre gaussien. De plus, pour un rendu visuel correct, la résolution des images convoluées doit être au plus égale à la demi-résolution de l'image HDR.

Le tableau 5.5 donne les performances en fonction du nombre de filtres gaussiens

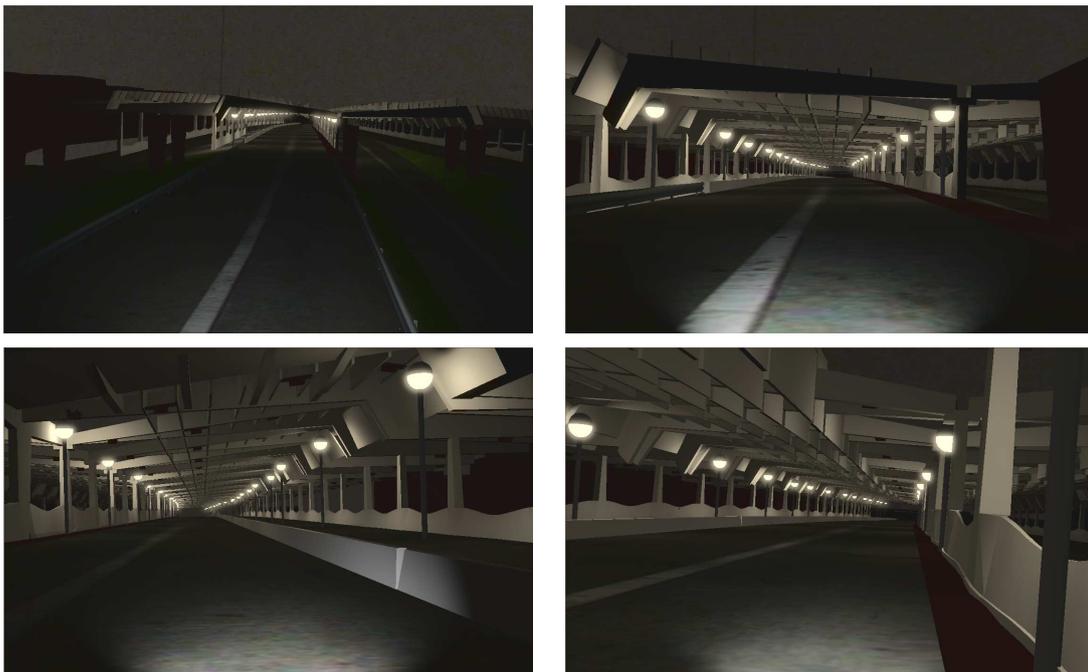


FIGURE 5.12 – Captures d’écran du rendu HDR appliqué à la simulation de conduite, sur un environnement virtuel modélisant un tunnel de l’autoroute A86. Base de données LIVIC-2010.

Nombre de filtres gaussiens	Nombre d’images par seconde
de 1 à 7	≥ 60
8	42
9	20

TABLE 5.5 – Performances de la simulation de l’éblouissement en fonction du nombre de filtres gaussiens utilisés.

Taille de filtres gaussiens, pour 3 filtres successifs	Nombre d’images par seconde
de 1 à 75	≥ 60
≥ 76	30

TABLE 5.6 – Performances de la simulation de l’éblouissement en fonction de la taille des filtres gaussiens utilisés, dans le cas où 3 filtres sont utilisés successivement.

utilisées, pour une image HD (1680×1050 pixels), et pour un filtre de taille 19×19 pixels (taille de filtre proposée par défaut). Si besoin, nous pouvons donc utiliser plus de filtres gaussiens (3 filtres par défaut). A titre d’information, le tableau 5.6 montre que l’on peut utiliser des filtres gaussiens de grande taille si besoin.

Étage (de l'image obtenue par <i>mipmapping</i>) utilisé pour le calcul de l'histogramme	Nombre d'images par seconde
0 (1680 × 1050)	0.72
1 (840 × 525)	2.5
2 (420 × 262)	7.5
3 (210 × 131)	20
4 (105 × 66)	≥ 60

TABLE 5.7 – Performances de l'opérateur d'Irawan et *al.* en fonction de l'étage (de l'image obtenue par *mipmapping*) utilisé pour le calcul de l'histogramme.

Enfin, le temps d'exécution de l'opérateur de reproduction de tons utilisé, celui d'Irawan et *al.* [Irawan 05], dépend essentiellement de deux paramètres : la taille de l'image à traiter, et la résolution de l'image sur laquelle on calcule l'histogramme. Cette dernière résolution doit être de taille suffisante pour être significative, mais il n'y a pas besoin de prendre l'image entière [Irawan 05]. Nous avons donc fixé ce paramètre, en prenant le quatrième niveau de l'image construite par *mipmapping* (voir la section 5.2.4 sur l'implémentation). Dans notre implémentation, le temps d'exécution est par contre indépendant du nombre de cases utilisées dans l'histogramme, chaque case étant traitée en parallèle des autres par les *pixel shaders*. En pratique, l'utilisation de 70 cases s'est montrée suffisante.

Performances de pipeline entier

Nous présentons finalement le temps d'exécution du rendu HDR dans son intégralité. Les matériels utilisés sont un Intel Core 2 Duo 3.16 GHz, avec 3 Go de RAM, et une carte graphique nVidia GeForce GTX 280. Nous réalisons des images de résolution HD (1680 × 1050 pixels). Nous effectuons ces mesures sur une scène ne présentant pas plus de 12 sources de lumière contribuant à l'éclairage de chaque pixel, et comprenant en tout un nombre variable N de sources ($N - 4$ lampadaires, sources diffuses uniquement, et 4 phares de voiture, sources diffuses et spéculaires, voir tableau 5.8). La simulation de l'éblouissement utilise 3 filtres gaussiens. La reproduction de tons est effectuée avec un histogramme calculé sur une image de résolution 110 × 66 pixels.

Les performances présentées dans la table 5.8 sont les performances minimales mesurées lors de la circulation dans l'environnement virtuel de la base modélisant un quartier de Paris (voir section 5.2.5 pour illustration). Avec les optimisations proposées, il vaut donc mieux se restreindre à une douzaine de sources dans la scène si l'on

Nombre total N de sources	Nombre d'images par seconde
de 5 à 12	≥ 39
de 13 à 45	≥ 30
de 46 à 80	≥ 20

TABLE 5.8 – Performances du rendu HDR, en fonction du nombre total de sources présentes dans l'environnement virtuel.

veut conserver quasi-strictement le temps réel et s'assurer d'avoir de bonnes conditions de circulation (sans saccade notamment) en toute circonstance. Si l'on a besoin d'afficher plus de sources, il faudra alors soit chercher à optimiser d'avantage le calcul d'éclairage, soit se résigner à ressentir quelques saccades lors de l'exécution du programme. Néanmoins, le programme restera pleinement interactif tant que l'on ne dépasse pas 45 sources de lumière dans l'environnement virtuel.

5.2.7 Conclusion

Dans ce chapitre, nous avons présenté trois composantes pour obtenir un rendu HDR de la meilleure qualité possible : le rendu physique d'une image HDR représentant la scène, puis la simulation de l'éblouissement face aux zones très lumineuses, et enfin la reproduction des tons de l'image HDR à l'écran. Nous espérons que les résultats présentés améliorent le réalisme des environnements virtuels utilisés, et apporte une meilleure fidélité par rapport à une visualisation réelle ; c'est en tout cas ce que nous évaluerons dans le prochain chapitre.

Ce rendu HDR est temps réel, pour la majorité des cas qui peuvent être rencontrés dans des expérimentations sur simulateur de conduite. Il faudra néanmoins faire attention aux scènes nocturnes où des pixels sont éclairés par de nombreuses sources de lumière (carrefours, embouteillages, ...) ainsi qu'aux scènes de grande taille qui font intervenir de nombreuses sources, car le calcul de l'éclairage peut faire augmenter le temps nécessaire pour rendre l'image HDR, et ainsi pénaliser l'ensemble de la simulation. Le calcul de l'éclairage reste le point le plus coûteux du rendu HDR par rapport au temps de calcul, et des optimisations supplémentaires seront nécessaires si l'on veut visualiser des scènes avec de très nombreuses sources de lumière.

Enfin, les paramètres proposés par défaut pour les filtres gaussiens ou le calcul de l'histogramme, la simulation de l'éblouissement et la reproduction de tons sont strictement temps réel ($> 60 Hz$).

Chapitre 6

Evaluation

Les simulations réalistes ont pour objectif de modéliser et de représenter à l'écran un système, un phénomène ou un environnement virtuel. Ces outils se rapportent à différents domaines d'application, notamment la formation, les études scientifiques, la visualisation, et les jeux vidéo. Deux étapes sont nécessaires pour ces simulations : la modélisation et le rendu. Elles sont intimement liées, car le rendu exploite les données de la modélisation pour l'affichage à l'écran.

En fonction des applications, certaines simulations, déjà réalistes, doivent en plus être visuellement le plus proche possible de la réalité. C'est notamment le cas pour les études scientifiques sur simulateur de conduite, car une étude n'est valide qu'en l'absence de biais sur le comportement du conducteur, entre la réalité et la simulation. La perception de l'environnement par le conducteur jouant un rôle crucial sur son comportement, que ce soit au niveau des informations visuelles, sonores, ou kinesthésiques, il faut donc reproduire le mieux possible l'environnement dans la simulation. Le critère de notre évaluation sera donc la fidélité (ou le réalisme¹) par rapport à la réalité, et sera appliqué sur des environnements virtuels, la nuit.

Le chapitre précédent présentait une méthode de rendu HDR, dont le but est de proposer la visualisation réaliste d'un environnement virtuel à l'écran. Il est maintenant nécessaire d'évaluer la fidélité de ce rendu par rapport à la réalité, afin de savoir dans quelles conditions il peut être utilisé pour des études scientifiques où la

1. Pour le terme « réalisme », il est important de préciser « par rapport à la réalité ». Le terme est alors synonyme de « fidélité ». On est dans le réalisme objectif. A l'inverse, le terme « réalisme » employé seul peut aussi être interprété dans le sens « réaliste aux yeux de l'observateur », et relève plutôt de l'apparence. On serait alors dans le réalisme subjectif.

perception de l'environnement est importante. Notons bien que nous allons procéder à une évaluation du rendu HDR pour une modélisation donnée : ce n'est donc pas le réalisme de l'image produite qui est évalué, mais uniquement le réalisme du rendu HDR, indépendamment de la modélisation.

L'objectif est donc d'évaluer les images et vidéos produites par le rendu HDR. Le rendu HDR proposé présente 3 étapes : le rendu de l'image HDR, la simulation de l'éblouissement, et la reproduction de tons (figure 6.1). L'évaluation de ce rendu est donc l'évaluation des images produites avec ces 3 composantes, toujours indépendamment de la modélisation effectuée. Cependant, l'évaluation globale de ces 3 composantes n'est pas aisée. Notre critère étant la fidélité par rapport à la réalité, l'évaluation la plus significative consisterait à comparer les images virtuelles produites avec des images réelles. Les travaux d'Ashikhmin et Goyal montrent en effet l'importance d'utiliser une référence réelle pour évaluer des images [Ashikhmin 06].

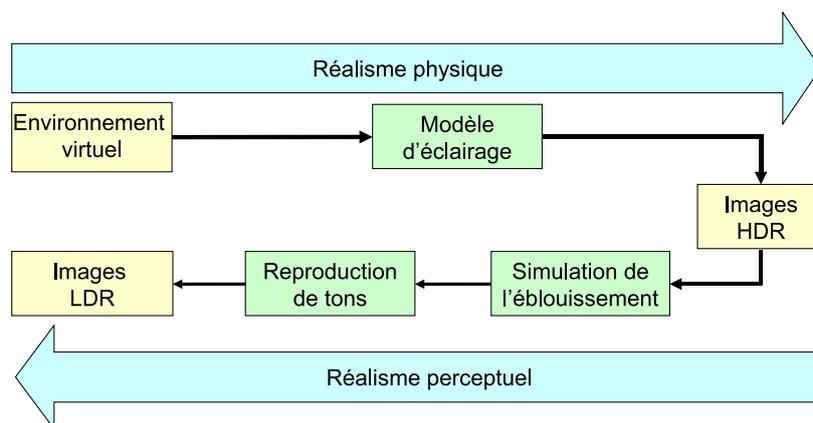


FIGURE 6.1 – Rendu HDR réaliste d'environnements virtuels.

La comparaison virtuel-réel sous-entend que les images produites soient ressemblantes à la réalité, ce qui n'est pas le cas des scènes routières virtuelles utilisées au LEPSiS (cela pourrait néanmoins être envisageable avec les modélisations de certains jeux vidéo, notamment Gran-Turismo 5 [Polyphony 10]). La modélisation exacte de la scène réelle, l'impératif d'un rendu temps réel et le temps de travail limité sont les obstacles principaux qui nous empêchent d'atteindre un rendu suffisamment proche de la réalité pour évaluer directement le virtuel par rapport au réel.

Puisqu'il nous était impossible d'évaluer le rendu HDR dans sa globalité par une comparaison virtuel-réel, nous proposons d'évaluer séparément les 3 composantes du rendu HDR. Chaque évaluation indiquera les qualités et les limites du rendu de chaque

composante, et montrera les améliorations réalisables à l'avenir si un meilleur rendu est souhaité. Nous préciserons en particulier les différences entre le rendu obtenu et la réalité qui peuvent introduire un biais dans la perception de l'observateur.

Dans ce chapitre, nous décrivons successivement les évaluations effectuées sur les 3 composantes : le rendu de l'image HDR, la simulation de l'éblouissement, et la reproduction de tons. La reproduction de tons étant très importante par rapport à l'aspect final de l'image, son évaluation sera particulièrement développée, et présentera l'utilisation d'une métrique et d'expérimentations avec des sujets. A l'inverse, les évaluations du rendu d'images HDR et de la simulation de l'éblouissement seront plus succinctes.

6.1 Evaluation du modèle HDR

La création des images de synthèse HDR consiste à simuler l'éclairage d'un environnement virtuel, et à stocker l'image produite dans une texture 32 bits. Nous avons proposé un pipeline HDR qui utilise une modélisation physique de l'environnement virtuel, notamment au niveau des sources de lumière et des coefficients de réflexion, ainsi qu'une modélisation physique de la propagation de la lumière. Nous avons donc exploité le maximum de données disponibles de l'étape de modélisation pour le rendu HDR, afin d'obtenir une simulation aussi physique que possible et une image HDR avec des valeurs en luminance les plus proches possible de la réalité.

Cependant, il reste des différences entre une image HDR virtuelle et une image HDR réelle capturée par un vidéoluminancemètre. Plusieurs raisons expliquent ces différences. Il y a d'abord une modélisation forcément approximative de la scène (c'est un travail très long et complexe, aussi bien au niveau de la géométrie qu'au niveau de la description des sources de lumière et des coefficients de réflexion des objets), mais il y a également le fait qu'*OpenGL* ne simule qu'une seule réflexion de la lumière entre la source de lumière et la caméra. De plus, il est également reconnu que les termes du modèle de Phong, notamment les termes *ambient* et *spéculaire*, ne reproduisent qu'une approximation grossière de la réalité. Ces raisons, liées à une modélisation et un rendu forcément approximatifs, font que l'on ne peut pas comparer directement les valeurs de deux images HDR, l'une réelle, l'autre virtuelle, par exemple avec une métrique mesurant les écarts entre les différentes valeurs de luminance.

Si la comparaison des valeurs exactes de luminance n'a pas grand sens, une évaluation par comparaison des ordres de grandeur et des contrastes des luminances présentes dans les images HDR réelles et virtuelles reste réalisable. Dans ce but, nous avons réalisé des images HDR réelles à partir de séries de photos de différents temps d'exposition, en utilisant le site internet WebHDR [Jacobs 10]. Un soin tout particulier a été porté à l'acquisition des sources de lumière, en prenant garde à ce qu'elles ne saturent pas le capteur lors de la prise de vue avec le plus faible temps d'exposition. Le site internet utilisé propose une formule pour retrouver les valeurs de luminance à partir des données contenues dans le *header* des images de chaque prise de vue. Si les valeurs de luminances retrouvées lors de la reconstruction sont en général définies à un facteur près, l'auteur du site annonce que ce facteur est ici proche de 1. Les ordres de grandeur des valeurs présentées sont en tout cas crédibles pour une scène de nuit. Les images HDR virtuelles, quant à elles, ont été acquises à partir de la texture 32 bits utilisée pour le rendu. Les figures 6.2 et 6.3 présentent respectivement des images HDR réelles et virtuelles de différentes scènes routières.

	Réel	Virtuel
chaussée	de 1 à 10 cd/m^2	de 10^{-1} à 1 cd/m^2
sources	10^5 cd/m^2	de 10^3 à 10^4 cd/m^2
ciel	de 10^{-1} à 10 cd/m^2	de 10^{-1} à 1 cd/m^2
décor non éclairé	de 10^{-4} à 10^{-2} cd/m^2	de 10^{-4} à 10^{-2} cd/m^2

TABLE 6.1 – Comparaison des ordres de grandeur des luminances présentes dans des images réelles et virtuelles.

La table 6.1 présente les ordres de grandeur des valeurs de luminances que l'on peut relever sur les images présentées. Nous pouvons y constater que les ordres de grandeur sont bien respectés, quel que soit l'élément de la scène considéré. Les plus gros écarts n'excèdent pas 1 ordre de grandeur, excepté pour les lampadaires, mesurés environ à 10^5 cd/m^2 dans le réel et réglés à 7.10^3 cd/m^2 dans le virtuel. Deux explications sont possibles : d'abord, la possibilité que les intensités des sources sont effectivement plus élevées dans le réel que dans le virtuel, ou également, le fait que dans l'environnement virtuel, nous avons placé des luminaires avec une surface apparente plus grande que les luminaires présents dans l'environnement réel. Si l'on fait l'hypothèse que les luminaires ont des intensités égales, il est donc normal que les luminaires réels, avec une plus petite surface apparente, présentent une luminance plus importante (voir équation

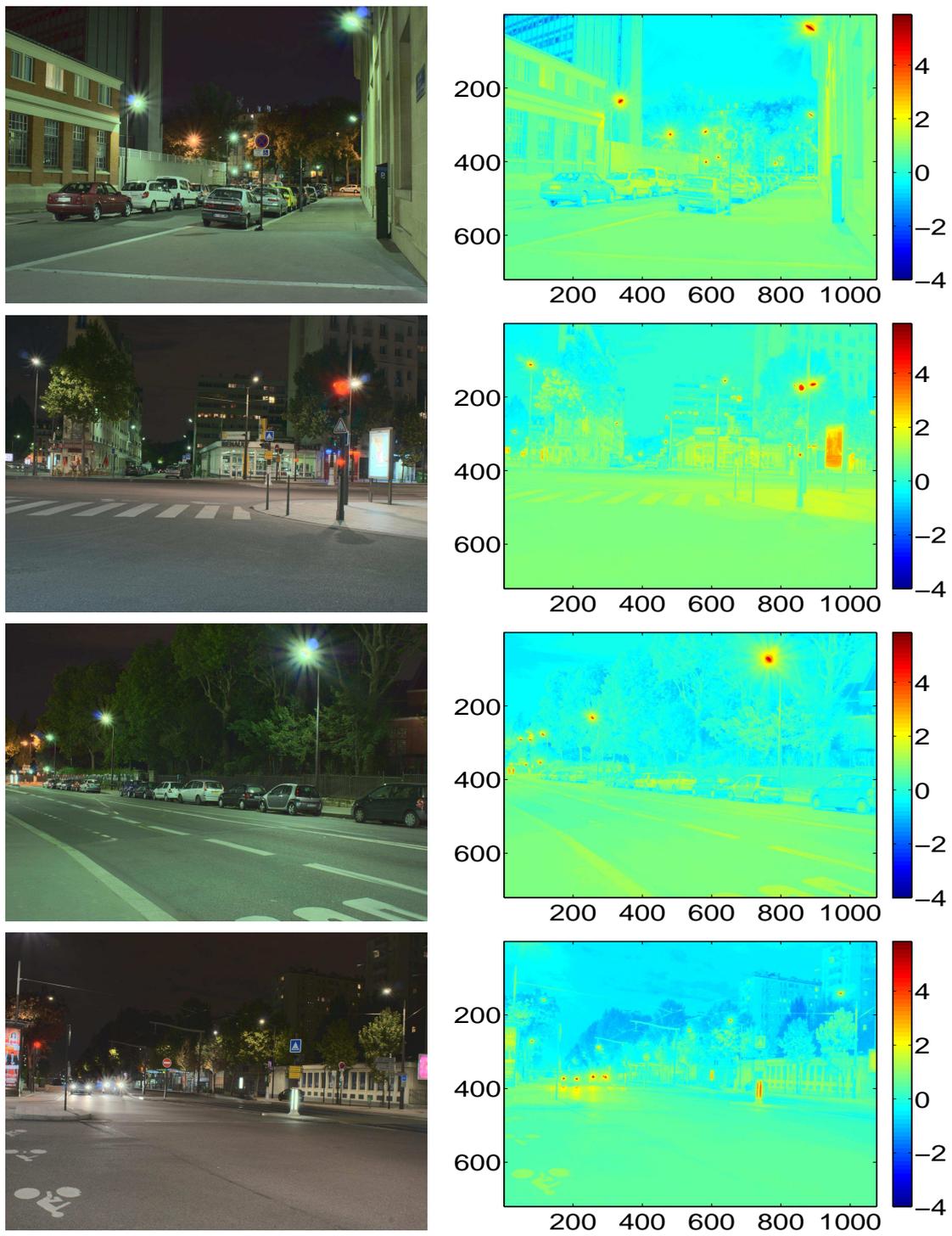


FIGURE 6.2 – Images HDR réelles, traitées par l’opérateur de reproduction de tons d’Irawan *et al.* (à gauche), et le logarithme décimal de leurs luminances respectives (à droite).

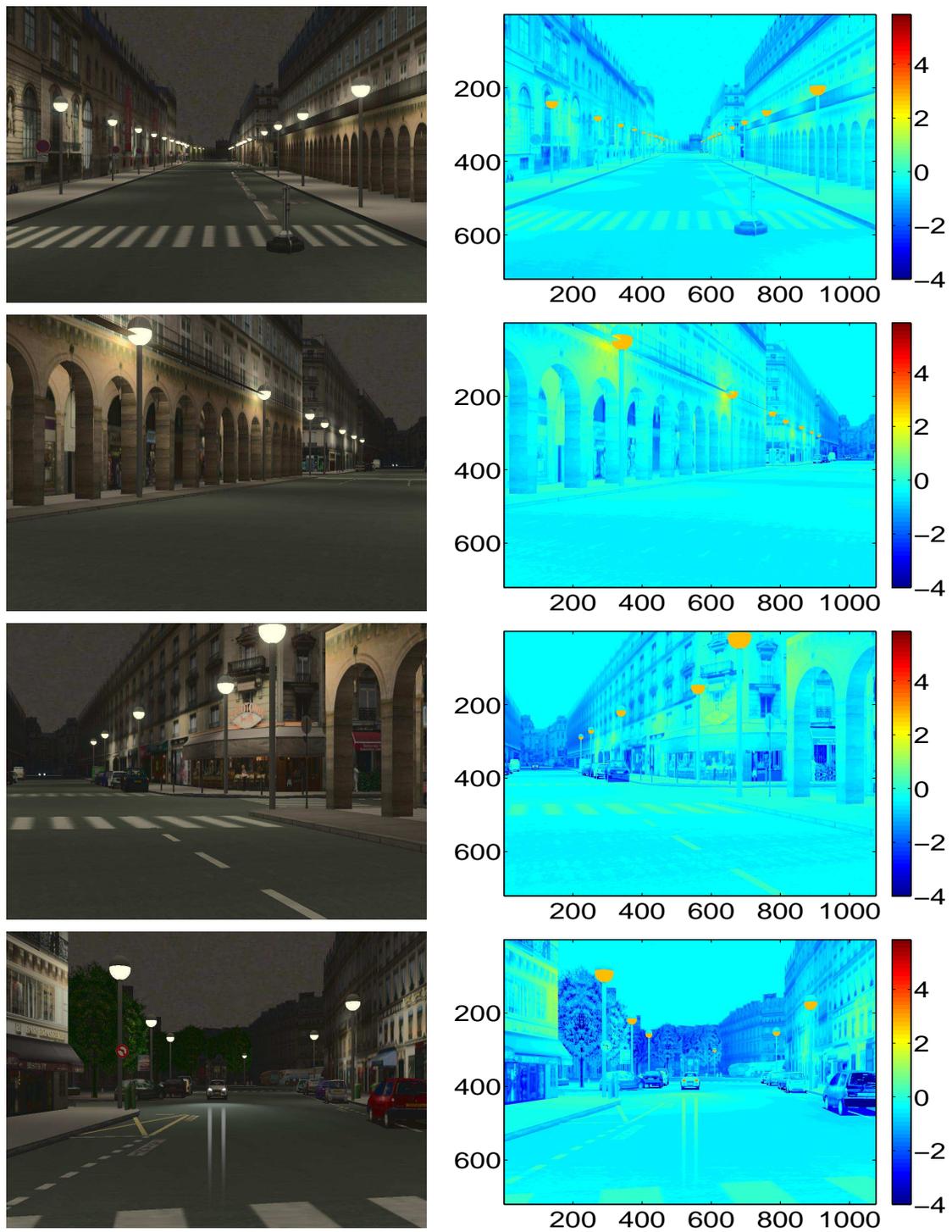


FIGURE 6.3 – Images HDR virtuelles, traitées par l’opérateur de reproduction de tons d’Irawan *et al.* (à gauche), et le logarithme décimal de leurs luminances respectives (à droite). Base de données 3D INRETS-MSIS-2007.

5.6). Par rapport à la chaussée, il semble que les luminances mesurées soient plus élevées dans les images réelles que dans les images virtuelles. Une explication pourrait être que les lampadaires sont placés plus haut dans l'environnement réel : ils ont donc probablement une intensité plus élevée, et éclairent sur une plus grande distance que les luminaires de l'environnement virtuel.

Au niveau des contrastes entre les différents objets de la scène, nous constatons qu'il y a également une très bonne correspondance : le contraste entre les sources et la chaussée est d'environ 10^4 pour les deux environnements (réel et virtuel), et le contraste entre le ciel et la chaussée est d'environ 10, également pour les deux environnements.

Enfin, nous pouvons également voir que la répartition générale des luminances est bien reproduite dans l'environnement virtuel, que ce soit au niveau de la chaussée ou des bâtiments. Sur la chaussée notamment, les luminances s'étalent uniquement sur 1 à 2 ordre(s) de grandeur, et respectent bien le fait que l'éclairage y est globalement uniforme, pour permettre au conducteur de rouler dans de bonnes conditions.

A partir de ces comparaisons, nous pouvons donc avancer que le rendu HDR présenté au chapitre précédent simule correctement les ordres de grandeur des luminances, dans l'environnement virtuel utilisé. Les contrastes entre les différents éléments de la scène sont de plus respectés. Par rapport à la réalité, il reste bien entendu un biais entre les luminances simulées et les luminances réelles de l'environnement. Ce biais est difficile à évaluer, car il n'est pas constant sur toute l'image, et dépend de la modélisation de l'environnement à représenter. Néanmoins, les ordres de grandeur étant proches entre la réalité et le virtuel, ce rendu devrait être suffisant pour la plupart des visualisations réalistes d'environnements virtuels, notamment sur simulateur de conduite.

6.2 Comparaison de simulations de l'éblouissement

La simulation de l'éblouissement permet de compenser le fait que les écrans standards ne peuvent éblouir un observateur, car la luminance maximale affichable est trop faible pour cela. Cette simulation répond à deux objectifs :

- reproduire à l'écran l'éblouissement tel qu'il serait vu par l'observateur dans la réalité (avec la luminance de voile, les stries et le halo lenticulaire) ;
- donner l'impression à l'observateur que la zone regardée est plus lumineuse que ce qui est affiché.

Nous avons présenté dans l'état de l'art (section 4.4) 4 algorithmes pour la simulation de l'éblouissement, et nous cherchons à comparer leurs performances par rapport à la réalité. Nous présentons ici leurs performances pour chaque objectif assigné.

6.2.1 Reproduction de l'éblouissement

Dans le cadre de la visualisation réaliste, nous cherchons à reproduire le plus fidèlement possible l'éblouissement tel que nous le percevons visuellement, à travers sa luminance de voile, ses stries, et son halo lenticulaire.

D'après les différents algorithmes proposés, le classement des modèles physiques, du plus avancé au moins avancé, est le suivant : [Ritschel 09], [Kakimoto 05], [Spencer 95] et [Kawase 04]. L'utilisation d'une modélisation physique ([Ritschel 09], [Kakimoto 05] et [Spencer 95]) donne clairement des résultats visuels plus proches de la réalité, et permet de modéliser les trois composantes de l'éblouissement, tandis que [Kawase 04] ne simule que la luminance de voile. De plus, l'algorithme de Ritschel et *al.* est le seul à simuler dynamiquement les trois composantes de l'éblouissement, ce qui, d'après les auteurs, donne des résultats visuels encore plus réalistes. La comparaison des différents filtres de l'éblouissement (figure 6.4) montre que la simulation ayant le modèle physique le plus proche de la réalité est celui qui donne le meilleur rendu des différentes composantes de l'éblouissement. Pour nous, le meilleur algorithme pour la reproduction des caractéristiques visuelles de l'éblouissement est donc celui de [Ritschel 09] (il n'est cependant pas tout à fait temps réel si l'on utilise les paramètres donnant le rendu le plus réaliste).

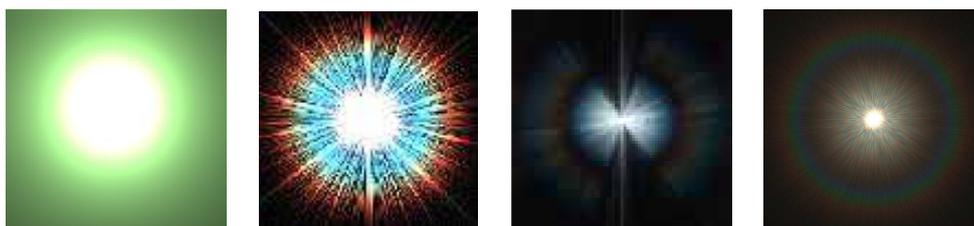


FIGURE 6.4 – Comparaison des filtres proposés par les 4 algorithmes simulant l'éblouissement (de gauche à droite [Kawase 04], [Spencer 95], [Kakimoto 05] et [Ritschel 09]). Base de données 3D INRETS-MSIS-2007.

Pour des raisons de temps, nous n'avons cependant pas pu développer cet algorithme dans notre rendu HDR, et nous nous sommes tournés vers des algorithmes visuellement moins performants mais plus rapides à développer : ceux de Kawase et de

Spencer *et al.* (seul celui de Kawase fonctionne en temps réel à ce jour).

L'algorithme de Kawase ne permet de simuler que la luminance de voile, tandis que celui de Spencer *et al.* simule les trois composantes à partir d'un filtre recalculé à chaque nouvelle image. La figure 6.5 présente des images HDR, toutes traitées avec l'opérateur de reproduction de tons d'Irawan *et al.* [Irawan 05], et compare les rendus sans puis avec les deux algorithmes de simulation de l'éblouissement.

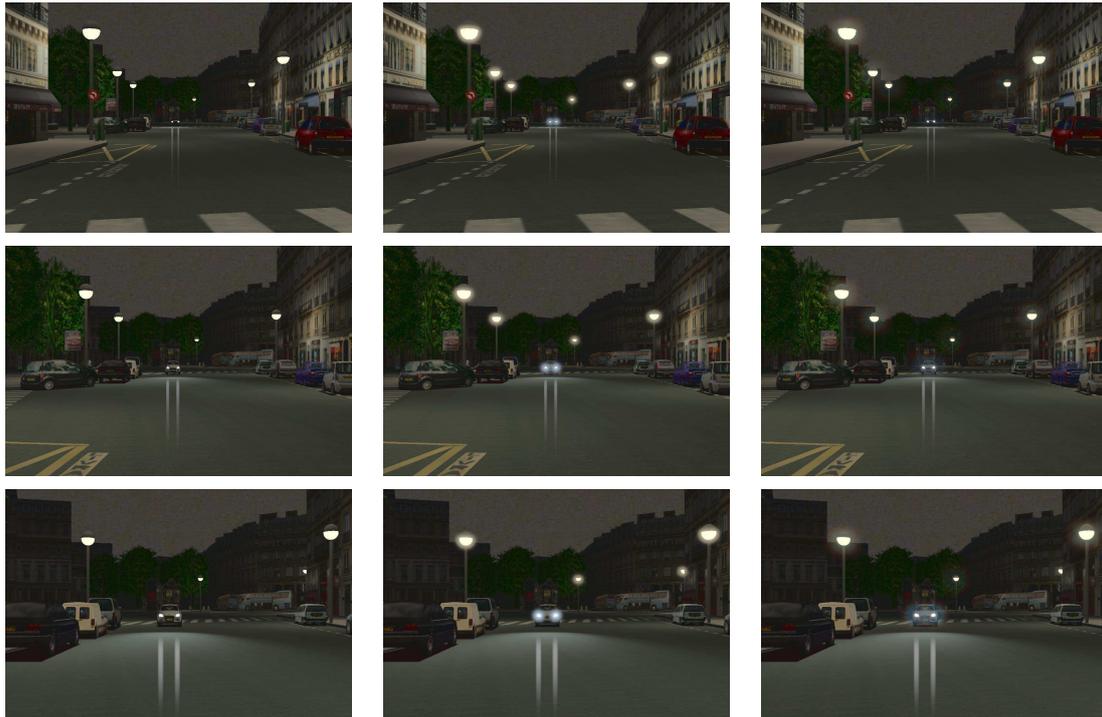


FIGURE 6.5 – Comparaison d'images HDR traitées sans simulation de l'éblouissement (à gauche), puis avec simulation de l'éblouissement de [Kawase 04] (au centre), et [Spencer 95] (à droite). Base de données INRETS-MSIS-2007.

Même si nous ne pouvons pas comparer par rapport la réalité, la figure 6.5 montre que chaque simulation de l'éblouissement apporte plus de réalisme. Pour nous, l'algorithme de Kawase est moins réaliste que celui de Spencer, puisqu'il ne reproduit que grossièrement la luminance de voile. La simulation de l'éblouissement par Spencer est moins marquée, et également plus fidèle à la réalité (notamment lorsque la voiture est proche, sur la dernière ligne de la figure 6.5).

Dans une application où l'on cherche à reproduire exactement les caractéristiques de l'éblouissement, l'utilisation de l'algorithme de Spencer *et al.* semble donc être préférable à celui de Kawase.

6.2.2 Augmentation de la luminosité perçue

La mesure de la luminosité face à une zone lumineuse est une démarche subjective, et est propre à chaque observateur (voir section 2.3.1). L'évaluation de la luminosité devrait donc être réalisée par une expérimentation avec des sujets. Une telle expérimentation a d'ailleurs déjà été réalisée par Yoshida et *al.* [Yoshida 08]. Elle évaluait la luminosité perçue par deux simulations de l'éblouissement distinctes : les filtres gaussiens de [Kawase 04] et la *point-spread-function* de [Spencer 95].

Les résultats montrent que les deux simulations permettent de simuler une luminosité de 20 à 35 % plus élevée. Mais les deux algorithmes n'ont pas exactement les mêmes performances suivant les luminances des zones traitées. Pour des zones entre 200 et 2500 cd/m^2 , la méthode de Spencer simule une luminosité plus élevée et un éblouissement plus réaliste. En revanche, au delà de 2500 cd/m^2 , c'est la méthode de Kawase qui simule la luminosité la plus élevée, car les filtres gaussiens prennent une taille nettement supérieure au filtre de Spencer, ce qui donne aux sujets l'impression d'être face à une source plus lumineuse [Li 99]. Cela a en revanche le défaut de réduire le réalisme de la simulation.

Si nous utilisons à nouveau les images présentées sur la figure 6.5, en comparant notamment les simulations obtenues autour des phares, nous pouvons effectivement constater que les filtres gaussiens donnent l'impression que la luminosité des phares est plus élevée qu'avec la simulation de Spencer et *al.*

6.2.3 Conclusion sur la simulation de l'éblouissement

Nous avons étudié dans les paragraphes précédents deux algorithmes pour la simulation de l'éblouissement : l'un réaliste [Spencer 95], l'autre plus suggestif [Kawase 04]. Nous avons pu constater que si le premier donne de meilleurs résultats sur la reproduction des caractéristiques de l'éblouissement (la luminance de voile, le halo lenticulaire et les stries), le second suggère une luminosité plus importante à l'observateur. Le choix de l'un ou l'autre de ces algorithmes doit donc être fait en fonction de ce qui est le plus important pour l'application.

6.3 Evaluer la reproduction de tons

La reproduction de tons établit une correspondance entre les luminances de l'image HDR et celles proposées par un écran. La dynamique offerte par les écrans standards étant plus petite que celle de la réalité, il y a forcément une dégradation du contenu de l'image. Mais pour une application dont le but est de présenter à l'écran une image visuellement similaire à celle du contenu de l'image HDR, afin que l'observateur ait une perception identique des deux images, l'opérateur de reproduction de tons doit conserver les caractéristiques principales de l'image HDR.

Plus d'une trentaine d'opérateurs existent dans la littérature, et nous avons été amenés dans le chapitre précédent à choisir un opérateur *a priori*, pour son modèle visuel (notamment l'adaptation temporelle) et sa reproduction des contrastes à l'écran. Nous allons maintenant évaluer les performances de cet opérateur, en comparant les images qu'il produit avec celles provenant d'autres opérateurs, et si possible avec l'image HDR d'origine. Le critère d'évaluation sera la **fidélité** : l'image affichée à l'écran doit être la plus fidèle possible au contenu de l'image HDR.

Il existe deux méthodes pour évaluer la qualité d'un opérateur. La première est objective, et fait appel à des métriques qui mesurent l'écart entre l'image proposée et l'image idéale. La seconde est subjective, et utilise des sujets pour effectuer l'évaluation. Les évaluations avec des métriques, en plus d'être objectives, ont également l'avantage de pouvoir être effectuées automatiquement par ordinateur, et peuvent donc être utilisées sur de nombreuses images ou vidéos en un temps raisonnable pour l'évaluateur. **Il n'existe cependant pas de métrique objective validée permettant d'évaluer un opérateur de reproduction de tons.** A l'inverse, les évaluations par des expérimentations permettent d'interagir avec les sujets, et ainsi d'avoir des données plus précises que par des métriques. Il en existe de nombreuses dans la littérature (section 4.3). Les contreparties sont qu'elles demandent généralement beaucoup de temps, et que les résultats peuvent varier d'un sujet à l'autre (des analyses statistiques sont donc effectuées sur les résultats).

Dans la suite de ce chapitre, nous présentons diverses évaluations sur la reproduction de tons, d'abord par une métrique (section 6.4), puis par deux expérimentations avec des sujets (section 6.5).

6.4 Evaluation de la reproduction de tons par une métrique

Nous avons d'abord voulu développer une métrique pour évaluer les images des opérateurs de reproduction de tons. L'intérêt est d'une part d'avoir une fonction qui évalue automatiquement les images ou vidéos traitées, en un temps réduit pour l'évaluateur, et d'autre part d'avoir des résultats objectifs, indépendants de la perception de chacun.

Les précédentes évaluations, réalisées avec des sujets et uniquement sur des images, prenaient en compte différents critères, tels la reproduction de la luminosité, du contraste, des couleurs, des détails dans les zones sombres et lumineuses, la présence d'artefacts et/ou la qualité de l'image dans son ensemble [Ledda 05, Yoshida 07, Cadík 08]. Mais dans le cadre de la simulation de conduite, ces critères ne traduisent pas ce que nous attendons de la reproduction de tons : la fidélité par rapport à la réalité (ou au contenu de l'image HDR). Nous voulons en effet que le conducteur ait la même perception de l'environnement à l'écran et dans la réalité, et les critères utilisés (luminosité, contraste, couleurs, détails) ne permettent pas de savoir si l'observateur a effectivement une perception identique de la scène. De plus, la perception d'images est différente de la perception de vidéos : sur une image, l'observateur peut disposer de temps pour détailler le contenu de l'image, alors que sur une vidéo, l'observateur ne perçoit pas forcément le contenu entier de chaque image. C'est pourquoi nous proposons d'introduire un nouveau critère d'évaluation. Ce critère permettra l'évaluation d'images mais aussi de vidéos.

6.4.1 La saillance comme critère d'évaluation

La saillance est la propriété de capter l'attention. Notre hypothèse est qu'un opérateur sera performant, du point de vue de la saillance, et par extension de l'attention, s'il conserve les zones saillantes de l'image ou de la vidéo HDR. Selon nous, le regard de l'observateur doit être attiré par les mêmes zones. Nous avons donc suivi cette hypothèse et étudié la conservation de la saillance par la reproduction de tons à travers une évaluation.

Dans un premier temps, nous ne nous intéressons qu'à une étude statique, sur des images. L'évaluation est réalisée en utilisant des cartes de saillance, qui associent une valeur de saillance à chaque pixel. La carte de saillance de l'image HDR sert de référence, et est comparée avec les cartes des images traitées par différents opérateurs de

reproduction de tons ([Tumblin 93, Ward 94, Pattanaik 00, Reinhard 02, Reinhard 05a, Mantiuk 08]). Ainsi, plus la différence est faible entre les paires de cartes de saillance, et meilleur sera l'opérateur de reproduction de tons. Le schéma de l'évaluation proposée est reproduit sur la figure 6.6.

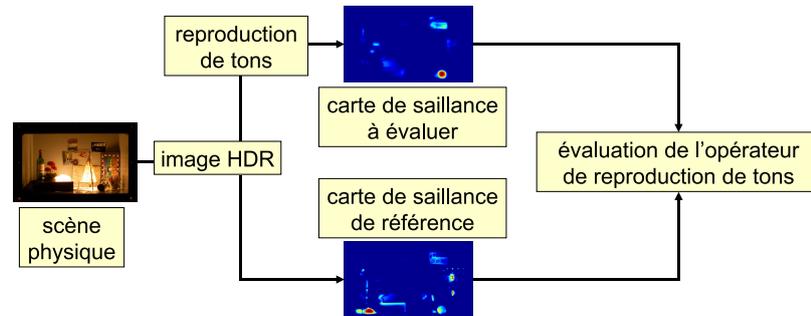


FIGURE 6.6 – Evaluation d'opérateurs de reproduction de tons par la saillance.

Plusieurs algorithmes existent pour calculer des cartes de saillance, tels [Itti 98, Itti 00, Itti 03, Itti 04, LeMeur 06], les trois derniers cités permettant également de traiter des vidéos. Pour notre évaluation, nous avons utilisé l'algorithme d'Itti et Koch [Itti 00], pour deux raisons : il a été évalué et validé sur de nombreuses images, et l'algorithme est disponible en ligne [Walther 06].

Cet algorithme a néanmoins été réalisé pour des images LDR, et pas pour des images HDR. Or il n'existe pas, dans la littérature, d'algorithme de calcul de cartes de saillance spécifiquement développé et évalué pour des images HDR. Dans la suite de cette section, nous allons vérifier si l'algorithme d'Itti et Koch peut être utilisé avec des images HDR.

6.4.2 Saillance d'images HDR par [Itti 00]

L'évaluation sur la reproduction de la saillance que nous proposons nécessite de calculer la carte de saillance de chaque image HDR. Nous allons donc étudier les performances de l'algorithme de calcul de la saillance d'Itti et Koch sur des images HDR.

La table 6.2 présente différentes images HDR, dont nous avons affiché les images LDR, traitées par l'opérateur [Reinhard 02] (voir (a) page 150). Nous présentons également les cartes de saillance obtenues avec l'algorithme d'Itti et Koch sur les images LDR de la première colonne (b). Ensuite, les cartes de saillance des images HDR ont été calculées avec les paramètres par défaut de l'algorithme d'Itti et Koch (c).

Au vu des cartes de saillance produites directement à partir des images HDR (c), nous remarquons que les points les plus saillants sont ceux qui présentent une forte intensité dans l'image HDR, notamment les sources de lumière. Comparé à (b), le nombre de zones saillantes est généralement réduit (systématiquement lorsqu'il y a une source de lumière dans l'image) : des zones saillantes ne sont donc plus détectées.

En regardant les cartes de conspécuité d'une image HDR traitée avec [Itti 00] (figure 6.7, seconde ligne), nous pouvons de plus remarquer que les contributions des intensités et des orientations se rapportent uniquement à la zone la plus lumineuse de l'image². Or, si l'on peut imaginer qu'un observateur sera plus sensible aux intensités d'une image HDR qu'à celles d'une image LDR, il reste néanmoins sensible aux couleurs et aux orientations présentes dans la scène entière.

A partir de ces constatations, nous avons conclu que l'algorithme d'Itti et Koch n'était par défaut pas adapté aux images HDR, et nous allons maintenant proposer d'adapter leur modèle au traitement d'images HDR.

6.4.3 Modification de l'algorithme d'Itti et Koch [Itti 00]

Le contraste plutôt que la différence

En regardant attentivement les 3 cartes de « conspécuité » (de l'intensité, des couleurs et des orientations, figure 6.7), nous avons observé que celle des couleurs contient plus d'information que les deux autres. La raison est la suivante : lors des différences inter-échelles, les couleurs sont normalisées en chaque pixel par l'intensité, alors que les cartes des intensités et des orientations sont calculées uniquement à partir des différences inter-échelles (voir équations 4.75 et 4.77). En sachant que l'œil est plus sensible au contraste qu'à la différence de luminance, nous proposons de modifier le calcul des cartes de conspécuité des intensités, selon la formule suivante :

$$I'(c,s) = \frac{|I(c) \ominus I(s)|}{I(c)}, \quad c \in \{2,3,4\}, \quad \delta \in \{3,4\}, \quad s = c + \delta \quad (6.1)$$

où $I(x)$ représente l'image de niveau x dans la pyramide dyadique des intensités.

2. C'est en fait le cas pour de nombreuses images HDR, dès qu'une source de forte intensité est présente dans la scène.

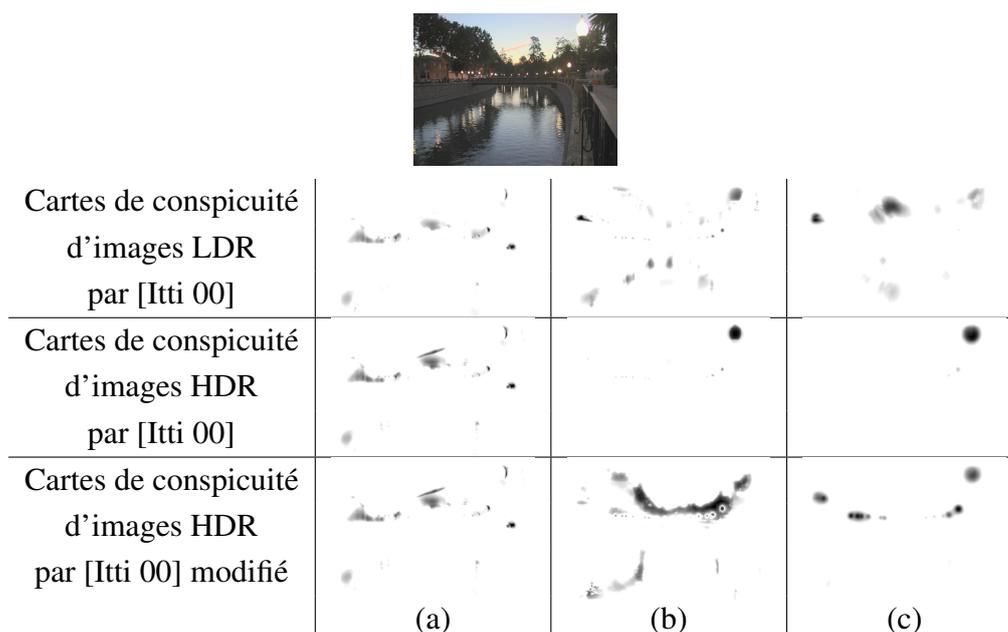


FIGURE 6.7 – Comparaison des images de conspécuité de l'algorithme d'Itti et Koch [Itti 00], en utilisant en entrée une image LDR (1^{ère} ligne) ou une image HDR (2^{nde} ligne). Pour comparaison, nous présentons les cartes de conspécuité obtenues avec l'algorithme d'[Itti 00] modifié sur des images HDR (3^{ème} ligne). Les colonnes montrent les couleurs (a), les intensités (b) et les orientations (c). Image provenant de [Reinhard 05b].

Un filtre de Gabor plutôt qu'une différence de filtres

D'autre part, pour détecter les orientations, Itti et Koch proposent d'utiliser une différence de filtres de Gabor, à des échelles différentes. Mais cela conduit à détecter les zones autour des bords, et non les bords eux-mêmes (figure 6.8). C'est pourquoi nous proposons de supprimer cette différence multi-échelle pour calculer la carte des orientations :

$$O'(c, \theta) = \frac{O(c, \theta)}{I(c)}, \quad c \in \{2, 3, 4\} \quad (6.2)$$

où $O(x)$ représente l'image de niveau x dans la pyramide dyadique des couleurs.

Quelques exemples de cartes de saillance obtenues avec les modifications proposées de l'algorithme d'Itti et Koch pour traiter des images HDR sont présentées sur la table 6.2 (d). Ces nouvelles cartes présentent un nombre plus important de zones saillantes, et selon nous d'une pertinence plus élevée que celles proposées par l'algorithme d'Itti et Koch par défaut (c).

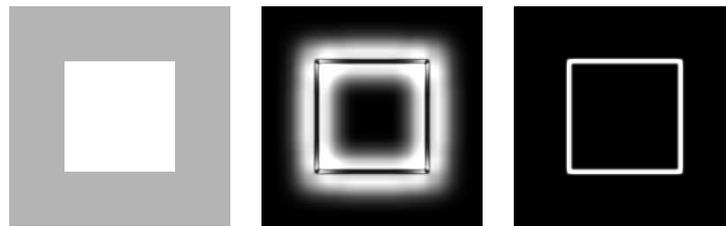
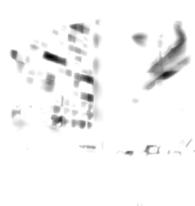
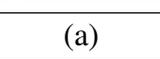
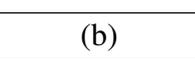
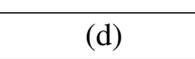


FIGURE 6.8 – Détection de contours. L'image d'origine (à gauche) est filtrée par une différence de filtre de Gabor (au centre), ou simplement par un filtre de Gabor (notre proposition, à droite).

Images LDR	Cartes de saillance d'images LDR par [Itti 00]	Cartes de saillance d'images HDR par [Itti 00]	Cartes de saillance d'images HDR par [Itti 00] modifié
(a)	(b)	(c)	(d)

Images LDR	Cartes de saillance d'images LDR par [Itti 00]	Cartes de saillance d'images HDR par [Itti 00]	Cartes de saillance d'images HDR par [Itti 00] modifié
			
			
			
			
			
			
			
(a)	(b)	(c)	(d)

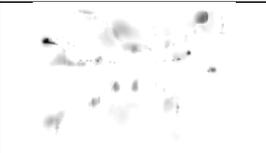
Images LDR	Cartes de saillance d'images LDR par [Itti 00]	Cartes de saillance d'images HDR par [Itti 00]	Cartes de saillance d'images HDR par [Itti 00] modifié
			
(a)	(b)	(c)	(d)

TABLE 6.2: Comparaison de différentes cartes de saillance. Images provenant de [Reinhard 05b].

6.4.4 Validation de l'algorithme d'Itti et Koch modifié

Avant de pouvoir exploiter ce nouvel algorithme pour le calcul de cartes de saillance d'images HDR, nous voulions valider cet algorithme par une expérimentation avec des sujets. Elle visait à comparer les zones de saillance calculées par l'algorithme avec les zones de l'image effectivement regardées par un panel de sujets, et à vérifier la correspondance. Nous en avons également profité pour comparer ces résultats aux cartes de saillance d'images produites par différents opérateurs de reproduction de tons.

Protocole de l'expérimentation

Nous avons construit une scène HDR dans un environnement contrôlé, composé de 2 sources de lumière et d'objets divers (figure 6.9). Cette scène HDR présentait un contraste de 3 480 000 :1. Les sujets ont été assis devant la scène HDR, qui était visible selon un angle de 20°, si bien que les sujets n'ont pas eu besoin de bouger la tête pour visualiser l'ensemble de la scène. Un *eye-tracker* utilisant des rayons infrarouges était de plus positionné devant les sujets, pour relever les points qu'ils fixaient sur la scène HDR. Ayant suivi un protocole similaire à [Itti 04], les consignes étaient d'observer la scène pendant 30 secondes, et de mémoriser les différents éléments présentés. La scène était d'abord cachée, avec ses lumières éteintes, alors que l'on effectuait la calibration de l'*eye-tracker* avec 8 LED présentes autour de la scène, plus une présente en son centre. Ensuite, la scène était dévoilée, les lumières allumées, et l'enregistrement des points de fixation des sujets était effectué pendant 30 secondes.



FIGURE 6.9 – Scène HDR utilisée pour l'expérimentation.

17 sujets ont participé à l'expérimentation (11 hommes, 6 femmes, d'une moyenne d'âge de 29 ans). Bien que certains sujets travaillent dans le domaine du traitement d'images, aucun d'entre eux ne connaissait le but de l'expérimentation. L'expérimentation entière ne durait pas plus de 15 minutes.

Résultats

Comme [LeMeur 06], nous avons construit des régions d'intérêt à partir des points de fixation des sujets. Chaque région d'intérêt est définie par un disque de rayon 1° sur lequel le regard du sujet est resté pendant au moins 100 *ms* consécutives. Toutes ces régions sont ensuite additionnées, pour former une distribution des régions d'intérêt. La carte ainsi construite est normalisée pour que la somme des coefficients soit égale à 1, et correspond ainsi à une distribution indiquant la probabilité que le regard d'un observateur se fixe sur un pixel donné (figure 6.10). Cette carte de saillance est notre référence. Elle est affichée en utilisant de fausses couleurs, pour mieux montrer les différentes valeurs de saillance.



FIGURE 6.10 – Cartes de fixation des sujets sur la scène HDR.

La seconde étape était de calculer les cartes de saillance de la scène HDR, pour comparer les performances des algorithmes d'[Itti 00] et d'[Itti 00] modifié. Une série

de photographies avec différents temps d'exposition a donc été réalisée, afin de reconstruire une image HDR de la scène HDR [Debevec 97, Jacobs 10]. Les cartes obtenues sont présentées sur la figure 6.11.

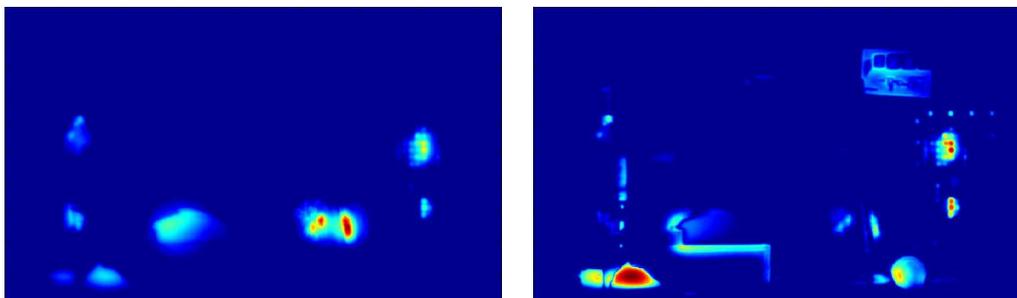


image HDR + [Itti 2000]

image HDR + [Itti 2000] modifié

FIGURE 6.11 – Cartes de saillance de l'image HDR.

Enfin, la troisième étape consistait à traiter l'image HDR par différents algorithmes de reproduction de tons, et à produire les cartes de saillance associées. Les opérateurs sélectionnés sont ceux de Tumblin et Rushmeier [Tumblin 93], noté O_1 , Irawan [Irawan 05], noté O_2 , Pattanaik et *al.* [Pattanaik 00], noté O_3 , Reinhard et *al.* [Reinhard 02], noté O_4 , Reinhard et *al.* [Reinhard 05a], noté O_5 , et Mantiuk et *al.* [Mantiuk 08], noté O_6 . L'implémentation de ces opérateurs provient de [PFStmo 10]. Les cartes obtenues sont présentées sur la figure 6.12.

Comme nous l'avons remarqué pour les images de la figure 6.2, l'algorithme d'Itti et Koch directement appliqué sur l'image HDR détecte principalement les zones très lumineuses, *ie.* les lampes (figure 6.11, à gauche). Il ne détecte les autres zones faiblement saillantes, sur la charte de couleurs et au niveau de la bouteille, que grâce à la carte de conspécuité des couleurs. Ce sont en effet des zones avec de forts contrastes de couleurs.

Cependant, si les modifications que nous avons apportées permettent effectivement d'obtenir un nombre plus important de zones saillantes (figure 6.11, à droite), celles-ci ne correspondent pas à la carte de référence que nous avons obtenue par l'expérimentation (figure 6.10). La carte de saillance de référence montre plusieurs zones étalées, alors que la carte calculée ne montre que des zones très locales. Les zones les plus saillantes d'après la carte de référence sont la charte colorée, la bouteille et la lampe de droite, alors que la zone la plus saillante d'après l'algorithme modifié est une zone sombre en bas à gauche qui n'apparaît pas dans la carte de référence. Des calculs ont été effectués pour mesurer la différence entre les cartes produites (erreur

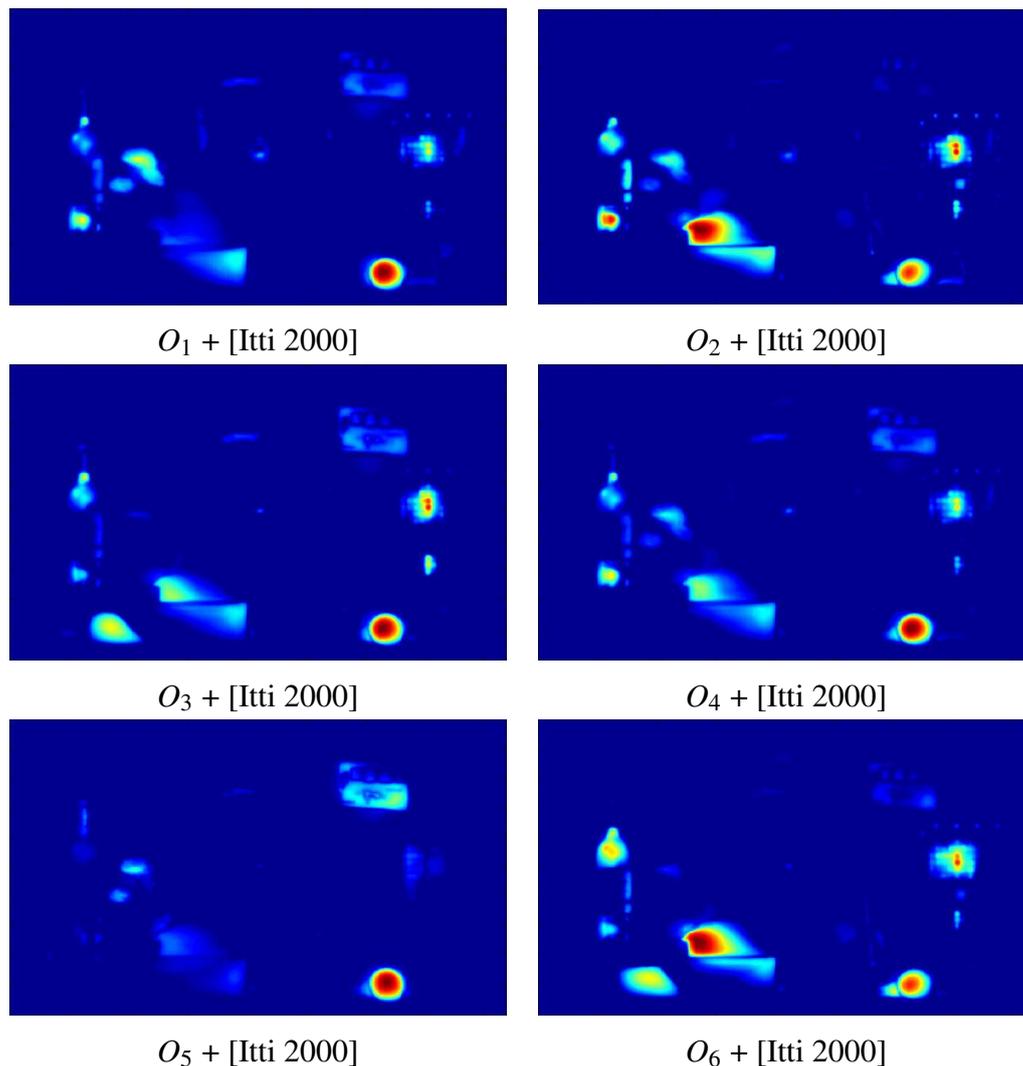


FIGURE 6.12 – Cartes de saillance des images traitées par les opérateurs de reproduction de tons.

quadratique moyenne, coefficient de Dice), mais n'ont pas montré une bonne correspondance. Le seul résultat satisfaisant, montré par le coefficient de Dice [?], est que la carte de saillance obtenue par l'algorithme d'[Itti 00] modifié est plus proche de la carte de référence que celle obtenue par [Itti 00]. Mais ce seul résultat ne nous permet pas de valider l'algorithme proposé, en tout cas pas sur l'image HDR de cette expérience. C'est pourquoi nous avons renoncé à utiliser cet algorithme pour évaluer les opérateurs de reproduction de tons.

En revanche, nous pouvons tout de même étudier l'influence des opérateurs de reproduction de tons sur les cartes de saillance calculées. Les différentes cartes présentées montrent qu'il y a bien une influence, car elles sont toutes différentes. Néanmoins,

elles ont de nombreux points communs. Par exemple, elles donnent toutes pour point le plus saillant le pot (en bas à droite de l'image). De plus, O_2 et O_6 pointent la lampe de gauche, O_2 , O_4 et O_6 capturent la bouteille de vin, tous les opérateurs sauf O_5 pointent la charte de couleurs. Les différences restantes entre ces différentes cartes concernent principalement des zones déclarées de faible saillance (avec une faible réponse dans la carte de saillance).

Enfin, nous pouvons étudier la correspondance entre les cartes de saillance issues des opérateurs de reproduction de tons avec la carte de référence donnée par les régions d'intérêt des sujets. La corrélation est moyenne, puisque la carte de référence indique la bouteille, la charte de couleurs et la lampe de droite (ainsi que la carte), alors que les cartes de saillance des opérateurs montrent le pot, la lampe de gauche, et la charte de couleurs.

6.4.5 Conclusion

Notre conclusion est que l'utilisation de la saillance comme critère d'évaluation pour évaluer les opérateurs de reproduction est difficile à mettre en œuvre, pour trois raisons :

- la première est qu'il n'y a pas d'algorithme de calcul de cartes de saillance validé pour les images HDR. Même si nous avons proposé un algorithme modifié qui semble fonctionner sur un large panel d'images, ses performances se sont montrées décevantes sur la scène HDR que nous avons expérimentée. Cet algorithme doit donc être utilisé avec précaution. Il peut de plus être amélioré, notamment en introduisant un aspect *top-down*, prenant en compte la tâche de l'observateur face à la scène observée, et éventuellement la signification des objets de la scène pour l'observateur.
- la seconde est que les cartes de saillance obtenues à partir de l'image HDR traitée par différents opérateurs ne sont pas très différentes, les zones détectées comme étant les plus saillantes étant souvent les mêmes. Nous en déduisons que les opérateurs de reproduction de tons utilisés n'ont pas véritablement changé les zones saillantes pour l'algorithme de calcul, et qu'il est donc difficile d'exploiter cette méthode pour réaliser une évaluation de différents opérateurs selon ce critère.
- enfin, la troisième est que les cartes de saillance des opérateurs de reproduction de tons ne correspondent que faiblement avec la carte de référence provenant des régions d'intérêt des sujets. Une explication serait que les sujets ont interprété la scène (en essayant par exemple de lire l'étiquette de la bouteille, ou les villes

présentes sur la carte routière centrale), alors que ce n'est pas pris en compte dans l'algorithme de calcul de la saillance.

Pour la simulation de conduite, l'utilisation d'un algorithme de calcul de cartes de saillance *bottom-up* pourra de plus se révéler insuffisant, car la perception visuelle des conducteurs est liée à leur tâche de conduite. Les cartes de saillance produites par cet algorithme risquent donc d'être éloignées de la vérité.

Au vu de ces raisons, nous avons renoncé à utiliser une métrique pour évaluer la reproduction de tons. C'est pourquoi nous nous sommes ensuite focalisés sur une évaluation effectuée directement par des expérimentations, où les sujets eux-mêmes notent la qualité des images présentées. Ces expérimentations sont développées dans la section suivante.

6.5 Evaluation de la reproduction de tons par des expérimentations

Puisque la métrique réalisée, quoique innovante, ne nous a pas permis d'évaluer les opérateurs de reproduction de ton de manière satisfaisante, nous nous sommes ensuite tournés vers une évaluation plus répandue dans la littérature : une évaluation subjective, réalisée par des expérimentations avec des sujets.

L'état de l'art sur les précédentes évaluations (section 4.3) a mis en avant plusieurs points :

- les évaluations déjà réalisées ne portaient que sur des images, jamais sur des vidéos ;
- si le critère est la fidélité, l'évaluation des sujets sera plus fiable s'ils comparent les images affichées à la réalité, plutôt que de comparer les images traitées par les opérateurs entre elles [Ashikhmin 06] ;
- aucun opérateur n'est meilleur que les autres, mais certains opérateurs se retrouvent régulièrement aux premières places des évaluations.
- les résultats d'une évaluation peuvent dépendre des images HDR utilisées [Kuang 07b].

Pour une application affichant un flux vidéo représentant un environnement virtuel de nuit, et où le but est d'avoir la meilleure fidélité possible par rapport au contenu de

chaque image HDR, nous avons voulu prendre en compte l'aspect dynamique dans notre évaluation. Les images HDR de luminance moyenne faible, et avec de forts contrastes entre les sources de lumière et les zones d'obscurité, sont déjà des cas difficiles à traiter pour les opérateurs de reproduction de tons. Mais dans un environnement virtuel dynamique, ces valeurs peuvent changer à chaque image, si bien que l'opérateur doit constamment s'ajuster. Nous pensons donc que l'adaptation temporelle jouera un rôle important dans cette évaluation.

Nous avons voulu utiliser des vidéos HDR de scènes réelles comme références. Mais les moyens actuels pour l'acquisition d'un signal HDR ne permettant pas de filmer directement une scène réelle (voir section 4.1.1), nous avons donc décomposé notre évaluation à travers deux expérimentations, l'une statique, l'autre dynamique (voir section 6.5.1).

Choix des opérateurs

Chaque évaluation portait sur une sélection de 8 opérateurs, que nous avons considérés comme les meilleurs d'après notre état de l'art et/ou les plus intéressants pour notre application : [Ward 94, Rahman 96, Ward 97, Reinhard 02, Durand 02, Choudhury 03, Reinhard 05b] et [Irawan 05]. L'opérateur iCAM [Fairchild 02] n'a pas été intégré, car il produit des images aux teintes rouges à partir d'images HDR contenant de faibles luminances. Ces opérateurs sont légèrement différents de ceux utilisés précédemment, car il ne proviennent pas de la même source ([Reinhard 05b]), exceptés [Ward 97] et [Irawan 05] que nous avons implémenté nous-mêmes.

6.5.1 Description des expérimentations

Expérimentation sur des images, par comparaison avec la scène réelle

La première expérimentation évaluait la fidélité d'images nocturnes. Les sujets devaient comparer les images produites par les 8 opérateurs avec la scène réelle telle qu'elle a été photographiée. Si certains opérateurs sont dotés d'une composante dynamique, elle n'entrait donc pas en compte ici.

Pour cette expérimentation, 10 scènes urbaines ont d'abord été photographiées de nuit, et les images HDR correspondantes ont été construites en utilisant plusieurs expositions de chaque scène ([Debevec 97, Jacobs 10]). Chaque image HDR a ensuite été traitée par les 8 opérateurs de reproduction de tons (voir figure 6.13). La même semaine que l'acquisition, de nuit, les sujets étaient ensuite amenés à comparer les

images traitées affichées sur un ordinateur portable (écran LCD), à l'endroit précis où chaque scène avait été photographiée. Les 13 sujets (9 hommes, 4 femmes) devaient noter les images de 0 (peu réaliste) à 10 (très réaliste). Ils s'entraînaient d'abord à noter sur les images de la première scène, et seules les 9 scènes suivantes ont été prises en compte pour l'évaluation. Pour chaque scène et chaque sujet, les images étaient présentées dans un ordre aléatoire. L'expérimentation durait 40 minutes.

Expérimentation sur des vidéos, sans référence

La seconde expérimentation évaluait la fidélité de vidéos nocturnes. C'est donc uniquement dans cette expérimentation que la composante dynamique des opérateurs a pu entrer en compte. Les vidéos utilisées présentaient des images de synthèse, il n'y avait donc pas de référence à la réalité. L'expérimentation évaluait le réalisme de chaque opérateur par rapport à l'image que les sujets se faisaient de la réalité. Nous sommes donc conscients qu'il y a un biais entre les résultats de cette l'évaluation et la fidélité des vidéos par rapport à la réalité, dû à la perception et à l'interprétation des vidéos par les sujets. C'est donc plutôt le sentiment de fidélité ressenti par les sujets qui est évalué ici.

Pour cette expérimentation, 5 flux d'images HDR ont été réalisés, 3 à partir d'une scène urbaine, 2 à partir d'une scène autoroutière réalisée par le LIVIC [LIVIC 10]. Chaque image HDR était ensuite traitée par les différents opérateurs. L'opérateur d'Irawan et *al.* étant le seul à intégrer un modèle d'adaptation temporelle, il traitait chaque image en fonction de l'état d'adaptation du modèle à l'image précédente, alors que les autres opérateurs traitaient chaque image indépendamment de la précédente. Les vidéos ont ensuite été créées à partir de chaque flux d'images provenant de chaque opérateur. Lors de l'évaluation, les 14 sujets (9 hommes, 5 femmes) s'entraînaient d'abord à noter 8 vidéos provenant de différents opérateurs (4 présentant la scène urbaine, 4 présentant la scène autoroutière), puis notaient les 40 vidéos présentées. Comme précédemment, les notes allaient de 0 (peu réaliste) à 10 (très réaliste). Pour chaque flux d'images HDR et pour chaque sujet, les vidéos provenant des opérateurs de reproduction de tons étaient présentées dans un ordre aléatoire. L'expérimentation a eu lieu dans une salle sombre, faiblement éclairée, et durait 20 minutes.



[Ward 94]



[Rahman 96]



[Ward 97]



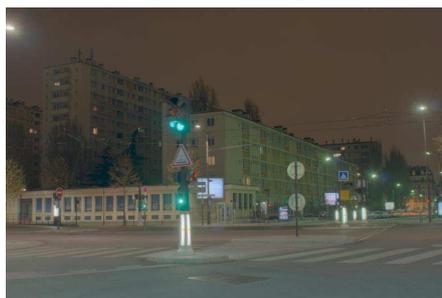
[Reinhard 02]



[Durand 02]



[Choudhury 03]



[Reinhard 05a]



[Irawan 05]

FIGURE 6.13 – Exemple d’une série de 8 images affichées successivement aux sujets, chacune correspondant à un opérateur de reproduction de tons.



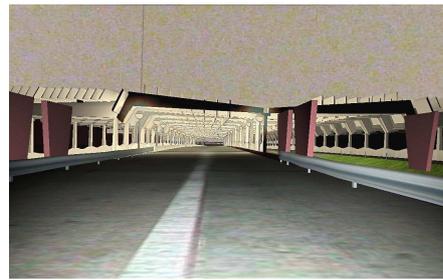
[Ward 94]



[Rahman 96]



[Ward 97]



[Reinhard 02]



[Durand 02]



[Choudhury 03]



[Reinhard 05a]



[Irawan 05]

FIGURE 6.14 – Exemple d’une série de 8 images affichées successivement aux sujets, chacune correspondant à un opérateur de reproduction de tons. Base de données 3D LIVIC-2010.

6.5.2 Résultats

La table 6.3 présente les moyennes des notes de chaque opérateur, pour chacune des expérimentations. Le rang de chaque opérateur, déterminé à partir de sa note moyenne, est également présenté. La première constatation est qu'aucune moyenne ne dépasse 7.5 / 10, et qu'aucun opérateur n'arrive à produire des images fidèles à la réalité quelle que soit l'image traitée. Pire, l'expérimentation 1, qui présente l'évaluation la plus fiable puisqu'elle compare des images des opérateurs avec la scène réelle, ne présente aucune moyenne supérieure à 6.1 / 10. Selon ses résultats, l'objectif de la reproduction de tons tel qu'il a été énoncé par Tumblin et Rushmeier (figure 4.3), qui est que l'observateur perçoive une image visuellement similaire à la réalité, n'est donc pas atteint, au moins pour des images de nuit.

Opérateurs	images	rang	vidéos	rang
[Ward 94]	2.65	[8]	3.27	[7]
[Rahman 96]	5.92	[2]	5.06	[4]
[Ward 97]	4.29	[5]	5.58	[3]
[Reinhard 02]	3.96	[6]	4.29	[6]
[Durand 02]	4.87	[4]	2.67	[8]
[Choudhury 03]	6.10	[1]	4.91	[5]
[Reinhard 05a]	3.90	[7]	5.62	[2]
[Irawan 05]	5.43	[3]	7.47	[1]

TABLE 6.3 – Score moyen des algorithmes de reproduction de tons sur les expérimentations statique (images) et dynamique (vidéos).

Ce résultat était attendu, et montre que la dynamique offerte par les écrans standards est insuffisante pour rendre les images visuellement similaires aux scènes qu'elles représentent. Néanmoins, nous cherchons à déterminer s'il y a un opérateur meilleur que les autres pour traiter des images de nuit selon un critère de fidélité par rapport à la réalité.

Une analyse de la variance (ANOVA) a été effectuée pour les deux expérimentations. Le seuil de significativité p a été fixé à 0.05. Le test post-hoc HSD (*Honestly Significant Difference*) de Tukey, ou test de la différence franchement significative, a de plus été réalisé pour déterminer si les différences de performances entre chaque paire d'algorithmes sont significatives ou non.

Expérimentation sur des images

Cette expérimentation amenait les sujets à comparer des scènes réelles avec leurs images dont les tons ont été reproduits sur un écran d'ordinateur. Seul l'aspect statique des opérateurs est donc pris en compte.

Dans cette expérimentation, les variables indépendantes étaient les algorithmes (8) et les scènes (9), les variables dépendantes étaient les sujets (13). Les résultats de l'analyse de variance montrent qu'il n'y a pas d'effet principal de l'image ($p = 0.186$), alors qu'il y a une forte influence de l'opérateur ($p = 10^{-6}$). Même si les performances des opérateurs peuvent varier d'une image à l'autre, il n'y a donc pas de variation significative sur les images de nuit utilisées ici.

Dans un premier temps, l'analyse de variance ne permettait pas de montrer des différences significatives entre les opérateurs. Mais puisqu'il n'y a pas de différence significative des notes d'une image à l'autre, nous avons réalisé une nouvelle analyse de variance sur les moyennes données par chaque sujet à chaque opérateur : cette méthode permet de mieux différencier les opérateurs.

	C03	R96	I05	D02	W97	R02	R05	W94
C03	–	NS	NS	*	*	*	*	*
R96	NS	–	NS	NS	*	*	*	*
I05	NS	NS	–	NS	NS	*	*	*
D02	*	NS	NS	–	NS	NS	NS	*
W97	*	*	NS	NS	–	NS	NS	*
R02	*	*	*	NS	NS	–	NS	*
R05	*	*	*	NS	NS	NS	–	NS
W94	*	*	*	*	*	*	NS	–

TABLE 6.4 – Test HSD de Tukey, jugeant si les différences entre chaque paire d'opérateurs est significative (*) ou non (NS : non-significatif), pour l'expérimentation sur des images. C03 : [Choudhury 03], R96 : [Rahman 96], I05 : [Irawan 05], D02 : [Durand 02], W97 : [Ward 97]. R02 : [Reinhard 02], R05 : [Reinhard 05a], W94 : [Ward 94].

Les résultats du test HSD de Tukey sont présentés sur la table 6.4. On peut voir que les performances d'opérateurs se suivant dans le classement ne peuvent pas toujours être considérés comme statistiquement différents (à cause de moyennes trop proches et d'écart-types trop élevés). Nous avons cependant pu classer les opérateurs selon 5

groupes, chaque groupe réunissant les opérateurs qui n'ont pas de performances significativement différentes entre eux. Par cette méthode, nous parvenons à distinguer statistiquement tous les opérateurs, car aucun opérateur n'appartient au(x) même(s) groupe(s) que les autres.

Les 3 premiers opérateurs [Choudhury 03, Rahman 96, Irawan 05] ne peuvent pas être distingués d'après les notes qu'ils ont obtenues lors de cette expérience, et appartiennent donc au premier groupe. Le second regroupe [Rahman 96, Irawan 05, Durand 02], le troisième [Irawan 05, Durand 02, Ward 97]. Le quatrième inclut 4 opérateurs [Durand 02, Ward 97, Reinhard 02, Reinhard 05a]. Le dernier groupe comprend [Reinhard 05a, Ward 94]. Les différents groupes auxquels appartiennent chaque opérateur sont présentés sur la table 6.5.

Opérateurs	C03	R96	I05	D02	W97	R02	R05	W94
groupe	1	1 & 2	1, 2 & 3	2, 3 & 4	3 & 4	4	4 & 5	5

TABLE 6.5 – Groupes d'appartenance des 8 opérateurs, sur l'expérience portant sur des images.

Les deux premiers opérateurs ont été appréciés des sujets, car ils respectaient la luminosité de la scène (l'image était correctement exposée). [Reinhard 02], et dans une moindre mesure [Ward 97, Reinhard 05a] et [Irawan 05] étaient généralement trop lumineux (les images semblent « surexposées »), [Durand 02] et [Ward 94] trop sombres (les images semblent « sous-exposées »). Lorsqu'on leur pose la question à la fin de l'expérimentation, les sujets utilisaient comme critères la luminosité, l'apparence naturelle de l'image, les contrastes, les couleurs, et parfois la taille et l'aspect des halos autour des sources.

Expérimentation sur des vidéos

Cette expérimentation amenait les sujets à évaluer la fidélité de scènes virtuelles nocturnes. Par l'intermédiaire du rendu HDR présenté au chapitre précédent, les tons de 5 vidéos HDR ont été successivement reproduits par les 8 mêmes opérateurs qu'à l'expérimentation précédente, donnant ainsi 40 vidéos à évaluer. L'aspect dynamique des opérateurs, s'il y en a un, est donc pris en compte.

Dans cette expérimentation, les variables indépendantes étaient les algorithmes (8) et les vidéos (5), les variables dépendantes étaient les sujets (14). Contrairement à l'expérimentation précédente, les résultats de l'analyse de variance montrent qu'il y a une influence significative de la vidéo ($p = 0.0199$), et toujours une forte influence

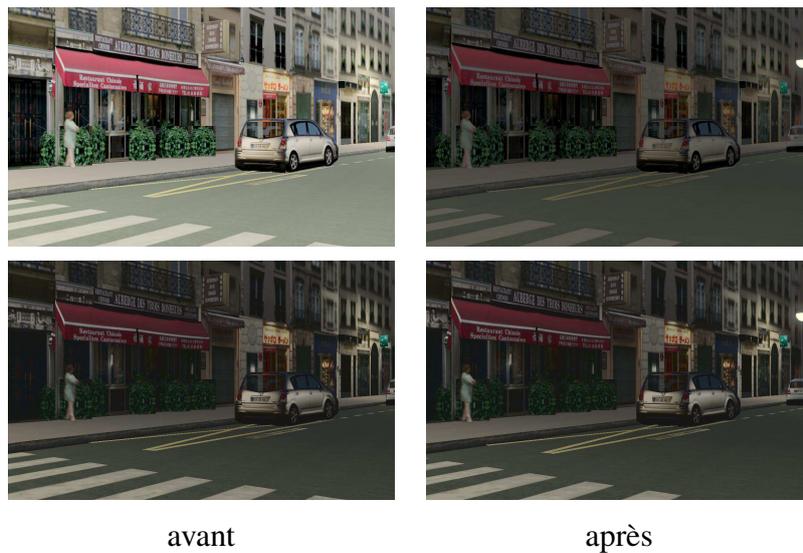


FIGURE 6.15 – Comparaison du comportement des opérateurs [Reinhard 05a] (en haut) et [Irawan 05] (en bas) face à une augmentation brusque de la luminance maximale, ici celle de la source du lampadaire. Base de données 3D INRETS-MSIS-2007.

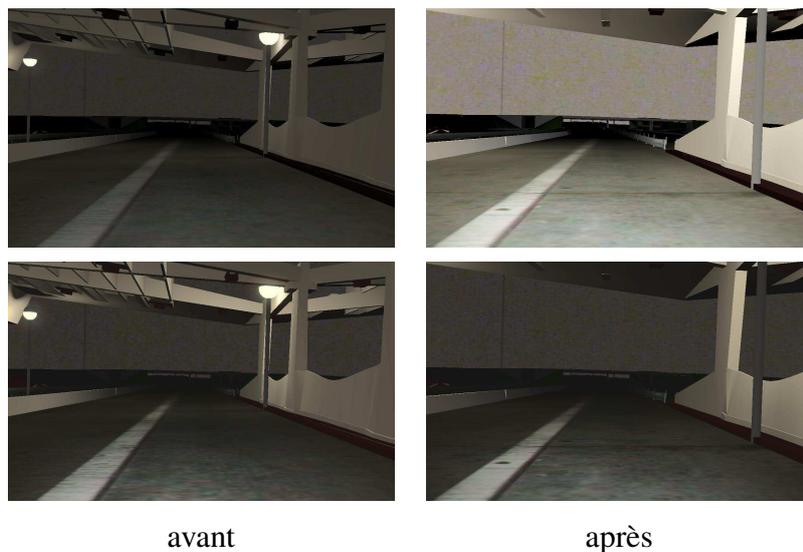


FIGURE 6.16 – Comparaison du comportement des opérateurs [Reinhard 05a] (en haut) et [Irawan 05] (en bas) face à une baisse brusque des luminances moyenne et maximale. Base de données 3D INRETS-MSIS-2007.

de l'opérateur ($p = 10^{-6}$). L'influence des vidéos sur les notes peut s'expliquer par la raison suivante : 3 vidéos utilisées présentent une dynamique de luminance « stable » au cours de la durée de chaque vidéo, tandis que 2 présentent un changement abrupt et

important de la luminance maximale, provoqué par la disparition ou l'apparition d'au moins une source lumineuse dans le champ de vision. La figure 6.15 illustre les images produites par les deux premiers opérateurs de cette évaluation, sur les images précédant et suivant l'apparition d'une source lumineuse dans le champ de vision. Les opérateurs « victimes », du point de vue de la fidélité, de ce brusque changement sont [Durand 02, Reinhard 05a, Choudhury 03, Rahman 96, Ward 94], tandis que [Irawan 05, Ward 97, Reinhard 02] traitent correctement ce passage.

Le second changement abrupt intervient à la sortie du tunnel d'autoroute (figure 6.17). L'intérieur du tunnel est éclairé, alors que la route extérieure ne l'est pas. Ici, ce sont à la fois la luminance maximale et la luminance moyenne qui chutent. Sans adaptation temporelle, l'image devient brusquement plus lumineuse à la sortie du tunnel. A l'inverse, l'adaptation est plus progressive et moins marquée si l'opérateur intègre un modèle d'adaptation temporelle.

Les résultats du test HSD de Tukey sont présentés sur la table 6.6. De même que pour l'expérimentation précédente, les performances d'opérateurs se suivant dans le classement ne peuvent pas être considérés comme statistiquement différentes (à cause de moyennes trop proches et des écarts-type trop élevés), excepté [Irawan 05] qui est significativement préféré aux autres opérateurs.

	I05	R05	W97	R96	C03	R02	W94	D02
I05	—	*	*	*	*	*	*	*
R05	*	—	NS	NS	NS	*	*	*
W97	*	NS	—	NS	NS	*	*	*
R96	*	NS	NS	—	NS	NS	*	*
C03	*	NS	NS	NS	—	NS	*	*
R02	*	*	*	NS	NS	—	NS	*
W94	*	*	*	*	*	NS	—	NS
D02	*	*	*	*	*	*	NS	—

TABLE 6.6 – Test HSD de Tukey, jugeant si les différences entre chaque paire d'opérateurs est significative (*) ou non (NS : non-significatif), pour l'expérimentation sur des vidéos. I05 : [Irawan 05], R05 : [Reinhard 05a], W97 : [Ward 97], R96 : [Rahman 96], C03 : [Choudhury 03], R02 : [Reinhard 02], W94 : [Ward 94], D02 : [Durand 02].

Pour cette expérimentation, nous avons également classé les opérateurs selon 5 groupes, chaque groupe réunissant les opérateurs qui n'ont pas de perfor-

mances significativement différentes entre eux. Le meilleur groupe ne comprend que l'opérateur d'[Irawan 05], tandis que le second rassemble les opérateurs de [Reinhard 05a, Ward 97, Rahman 96, Choudhury 03]. Le troisième inclut les opérateurs de [Rahman 96, Choudhury 03, Reinhard 02], et le quatrième [Reinhard 02, Ward 94]. Le dernier groupe comprend les opérateurs de [Ward 94, Durand 02]. L'appartenance de chaque opérateur aux différents groupes est représenté sur la table 6.7. Cette méthode permet de différencier les opérateurs, sauf [Reinhard 05a] et [Ward 97] qui sont équivalents, de même que [Rahman 96] et [Choudhury 03].

Opérateurs	I05	R05	W97	R96	C03	R02	W94	D02
groupe	1	2	2	2 & 3	2 & 3	3 & 4	4 & 5	5

TABLE 6.7 – Groupes d'appartenance des 8 opérateurs, sur l'expérience portant sur des vidéos.

Sur les vidéos, l'opérateur de [Durand 02] est celui qui a obtenu les plus mauvaises notes, car il n'y a pas de cohérence temporelle dans le traitement des images successives. Même s'il n'y a pas de brusque changement dans la dynamique de luminance, les images produites sont tantôt sombres, tantôt très sombres (figure 6.17), ce qui introduit de brusques et forts changements de luminosité au cours de la vidéo. Nous n'avons pas d'explication évidente à ce stade.



FIGURE 6.17 – Comportement de l'opérateur ([Durand 02]) sur deux images successives, alors qu'il n'y a pas de changement significatif dans le contenu de l'image. Base de données 3D INRETS-MSIS-2007.

Résultats sur les deux expérimentations

Les appartenances de chaque opérateur aux différents groupes sont présentées sur la table 6.8. Comme nous avons utilisé les paramètres par défaut de chaque algorithme, les opérateurs pourraient donner de meilleurs résultats avec un réglage plus fin. Leurs

implémentations n'ayant pas toutes été réalisées par nous-mêmes, des erreurs ont également pu persister lors de la réalisation des images.

Opérateurs	images	vidéos
[Irawan 05]	[1,2,3]	[1]
[Choudhury 03]	[1]	[2,3]
[Rahman 96]	[1,2]	[2,3]
[Ward 97]	[3,4]	[2]
[Reinhard 05a]	[4,5]	[2]
[Durand 02]	[2,3,4]	[5]
[Reinhard 02]	[4]	[3,4]
[Ward 94]	[5]	[4,5]

TABLE 6.8 – Groupes d'appartenance des 8 opérateurs sur les 2 expérimentations.

Néanmoins, d'après les résultats obtenus, et pour des images d'environnements urbains nocturnes, l'opérateur [Irawan 05] est le seul à faire parti du premier groupe des deux expérimentations. Cela signifie qu'aucun autre opérateur n'est meilleur que lui, quelle que soit l'expérimentation considérée. Cet opérateur est notamment meilleur que les autres dans l'expérimentation sur les vidéos, et une raison serait qu'il est le seul opérateur intégrant une modèle d'adaptation visuelle, permettant de simuler l'adaptation aux brusques changements lumineux. En revanche, il a parfois été critiqué dans l'expérimentation avec des images, certains sujets jugeant ses images trop lumineuses et trop contrastées.

Ces résultats font aussi l'hypothèse que les sujets ont noté les images présentées selon leur fidélité par rapport à la réalité. Des biais peuvent exister dans les résultats, non seulement dans l'expérimentation sur des vidéos car les sujets n'avaient pas de référence réelle, mais aussi dans l'expérimentation sur des images, car les sujets pouvaient avoir tendance à comparer les images entre elles, plutôt que de ne les comparer qu'à la réalité. Enfin, dans les deux expérimentations, les sujets pouvaient également noter les images selon leur préférence, plutôt que par rapport à la fidélité. Pour minimiser ces biais sur l'évaluation, la consigne de noter les images en fonction de leur fidélité par rapport à la réalité était fréquemment rappelée.

6.6 Conclusion

Dans ce chapitre, nous avons cherché à évaluer le rendu HDR présenté au chapitre précédent, en évaluant séparément ses trois composantes : le rendu de l'image HDR, la simulation de l'éblouissement et la reproduction de tons.

En comparant les images HDR virtuelles produites avec des images HDR réelles, nous avons pu constater une bonne correspondance des ordres de grandeur des luminances et des contrastes entre les différentes zones des images nocturnes. Si nous savions à l'avance qu'*OpenGL* et le modèle de Phong ne permettent pas d'obtenir des grandeurs exactes, nous avons pu voir qu'une modélisation physique de la scène permet d'obtenir des résultats proches de la réalité, du point de vue des luminances.

L'évaluation de la simulation de l'éblouissement a montré que les objectifs de réalisme n'étaient pas encore tout à fait atteints. Si les deux simulations de l'éblouissement étudiées ont montré qu'elles permettent de rendre des zones plus lumineuses aux yeux des observateurs qu'elles ne le sont réellement, le réalisme de ces simulations n'est pas parfait. A ce jour, la seule simulation temps réel dans notre application, exploitant des filtres gaussiens, suggère à l'observateur qu'il est ébloui, plutôt que de reproduire exactement la luminance de voile, les halos et les stries caractéristiques de l'éblouissement.

Enfin, l'évaluation sur la reproduction de tons a été la plus détaillée. D'abord, nous avons pu voir que l'évaluation d'opérateurs à travers une métrique basée sur la saillance est difficilement réalisable, parce qu'il n'existe pas d'algorithme de saillance à la fois *top-down* et *bottom-up* pour les images HDR. Nous avons proposé notre algorithme de saillance *bottom-up* adapté aux images HDR, mais ses performances restent insuffisantes pour en faire une métrique pour la reproduction de tons.

Nous nous sommes ensuite tournés vers l'évaluation de la reproduction de tons via des expérimentations avec sujets, et avons abouti à plusieurs conclusions. La première, obtenue d'après les notes moyennes des opérateurs, est qu'aucun d'entre eux ne donne à l'observateur des impressions visuellement similaires à l'écran et dans la réalité. Malgré le fait que nous ayons choisi des opérateurs parmi les meilleurs, il reste encore un écart entre les images affichées et la réalité. La seconde conclusion est l'importance de l'adaptation temporelle dans le traitement de flux vidéo, à la fois pour garder une

cohérence dans l’affichage d’images successives (lorsqu’il n’intervient pas de grand changement de luminances), mais aussi pour pouvoir s’adapter aux changements de luminances, en un temps familier pour l’observateur. Enfin, les deux expérimentations effectuées ont permis de montrer qu’aucun opérateur n’est meilleur que l’opérateur d’Irawan et *al.* pour le traitement d’images urbaines nocturnes, et à l’inverse, que cet opérateur est significativement meilleur dans le traitement de flux vidéos virtuels et également nocturnes.

Chapitre 7

Conclusion

En 2007, le LCPC a souhaité améliorer le visuel des simulateurs de conduite, en leur intégrant un rendu HDR. L'objectif est le même que celui des opérateurs de reproduction de tons : avoir une reproduction visuellement similaire de la réalité à l'écran, afin que le conducteur se comporte de manière identique dans un simulateur et sur la route.

Un rendu HDR

Dans cette thèse, nous nous sommes intéressés au problème de la visualisation réaliste d'environnements virtuels, et notamment, des scènes routières. Nous avons traité des images HDR représentant des scènes nocturnes, un des cas les plus difficiles à afficher car les contrastes y sont particulièrement importants. Pour mettre en œuvre le rendu HDR, nous avons redéveloppé l'ensemble du pipeline de rendu *OpenGL*, de l'exploitation de la modélisation de la scène jusqu'à son affichage à l'écran. La méthode proposée permet d'utiliser des environnements virtuels déjà existants, et ne demande qu'à ajouter une description des sources de lumière en unités physiques pour le rendu. L'image HDR résultante, en luminance, est ensuite traitée par une simulation de l'éblouissement puis par un opérateur de reproduction de tons. L'ensemble de ce rendu est temps réel, et est intégré dans les simulateurs du LEPSiS.

Le rendu de l'image HDR est réalisé en utilisant les propriétés physiques des sources de lumière, des objets de la scène, et de la lumière. Même si *OpenGL* ne permet pas de réaliser des simulations exactes, nous sommes restés aussi proches de la réalité physique que possible pour le rendu des luminances des images HDR. La comparaison d'images réelles avec des images virtuelles montre d'ailleurs que les ordres de grandeur des luminances sont bien respectés.

En ce qui concerne la simulation de l'éblouissement et la reproduction de tons, des choix ont été fait sur la sélection des algorithmes, pour maximiser le réalisme perceptif. A propos de l'éblouissement, le modèle de Ritschel et *al.* est celui qui montre les meilleurs résultats visuels, mais n'est pas tout à fait temps réel à ce jour. Nous avons donc utilisé ceux de Kawase et Spencer et *al.* Le premier suggère efficacement à l'observateur la présence de source(s) lumineuse(s), même si ses performances visuelles sont limitées, alors que le second a de meilleures performances visuelles, mais est moins efficace pour suggérer la luminosité des sources. Le choix entre ces deux algorithmes est à faire en fonction de l'application, sachant que dans notre implémentation, seul celui de Kawase fonctionne en temps réel à ce jour. A propos de la reproduction de tons, des expérimentations ont montré que l'opérateur d'Irawan et *al.* était le plus approprié pour la simulation de conduite, et qu'il pourrait également se révéler très performant dans d'autres applications interactives. Cet opérateur est doté d'un model visuel humain incluant l'adaptation temporelle, et compresse la dynamique de luminance avec un ajustement de l'histogramme de l'image HDR.

Enseignements

Divers enseignements peuvent être tirés de ce travail. Le premier est le lien étroit qu'il doit y avoir entre la modélisation d'une scène et l'exploitation qui est faite pour son rendu. Pour réaliser un rendu HDR réaliste, il faut décrire les paramètres des sources lumineuses en unités physiques. Les valeurs de luminance des sources contenues dans les textures HDR doivent de plus être en rapport avec la quantité de lumière qui est émise par ses sources. Si l'intensité émise par les sources est importante pour définir l'éclairage des objets à proximité de chaque source, la luminance des sources l'est également car elle entre en jeu dans la simulation de l'éblouissement de l'observateur face à ces sources. L'étape de modélisation est donc primordiale si l'on veut obtenir un rendu HDR de qualité.

Une telle modélisation peut en revanche poser des problèmes pratiques. D'une part, les logiciels d'infographie ne contiennent pas forcément les paramètres pour modéliser finement les lumières (intensité, direction, forme d'émission, atténuation), et d'autre part, l'exportation de ces données du logiciel d'infographie et leur importation dans le moteur de rendu n'est généralement pas immédiate. La modélisation nécessaire au rendu HDR a donc un coût, tant au niveau de la modélisation qu'au niveau de son exploitation. Il faut donc prévoir du temps pour la mise en place de ce rendu HDR.

Pour des scènes routières nocturnes, la modélisation des sources lumineuses, de la propagation de la lumière puis de sa réflexion sur les objets de la scène permet d'obtenir des images HDR aux luminances du même ordre de grandeur que celles rencontrées dans la réalité. Ce n'était pas un résultat évident car la modélisation et le moteur de rendu utilisés possèdent encore de nombreuses imperfections, forcément sources d'erreur dans les valeurs de luminance produites.

Un autre enseignement que nous a apporté ce travail concerne l'importance de la composante dynamique d'un opérateur de reproduction de tons pour le traitement de vidéos HDR. Cela a pu être mis en évidence grâce au rendu HDR développé, qui nous a permis de proposer la première évaluation d'opérateurs de reproduction de tons sur des vidéos. S'il était prévisible qu'un opérateur avec une composante dynamique amènerait de meilleurs résultats, nous avons vu que des opérateurs performants sur des images pouvaient voir leurs performances chuter sur des vidéos, car ils n'arrivent pas à traiter des changements de luminance dans la scène. Pour qu'un opérateur de reproduction de tons puisse traiter des vidéos de manière réaliste pour l'observateur, il semble donc qu'il doit intégrer *a minima* un modèle d'adaptation temporelle.

Perspectives sur le rendu HDR

Ce rendu HDR a été développé dans le cadre de la simulation de conduite de nuit, mais il peut également être utilisé dans de nombreuses applications. D'abord, il peut bien entendu être utilisé de jour. Cela a été testé sur la scène urbaine, il suffit de modéliser un éclairage diurne (via une texture HDR représentant le ciel et le soleil) et un éclairage de la scène correspondant pour obtenir des simulations de jour. Le gain obtenu par le rendu HDR par rapport à un rendu LDR est cependant moindre, puisque les contrastes sont moins importants de jour que de nuit, excepté en cas de gros changement dans la dynamique de luminance, par exemple lors d'un passage dans un tunnel. Ensuite, le rendu HDR peut également être utilisé dans tout type de simulation réaliste, comme des simulations d'éclairage ou des simulations architecturales. Dans ces applications, s'il n'y a pas d'objet en mouvement, le rendu HDR de l'environnement peut être réalisé une unique fois et stocké dans une texture : seuls la simulation de l'éblouissement et la reproduction de tons seraient ensuite mis à jour à chaque image, en fonction des déplacements de l'observateur dans la scène virtuelle. Enfin, même si la fidélité par rapport à la réalité n'est pas leur objectif premier, les jeux vidéo dits « réalistes » peuvent aussi tirer parti de ce rendu HDR.

Même si elle est déjà correcte, la précision de la simulation d'éclairage ne peut qu'augmenter avec dans les prochaines années, avec des progrès à venir dans la modélisation (comme la généralisation de l'emploi de BRDF, notamment pour améliorer le terme spéculaire) comme dans le rendu (avec l'apparition de moteurs de rendu à lancer de rayons temps réel). On devrait ainsi tendre vers les valeurs de luminance très proches de la réalité, et arriver également à intégrer les ombres, qui permettent d'améliorer le réalisme des simulations.

Avec du temps, l'intégration de la simulation de l'éblouissement de Ritschel et *al.*, la plus réaliste, est également envisageable. Le principal obstacle est l'utilisation de la transformée de Fourier en temps réel sur de grandes images, mais l'utilisation de la carte graphique pour cette tâche (par exemple avec la librairie CUFFT de Cuda [Nvidia 10]) devrait permettre d'obtenir un temps d'exécution satisfaisant.

Pour notre application, la simulation de conduite, nous avons choisi l'opérateur d'Irawan et *al.* pour son traitement des contrastes et son modèle visuel, et les évaluations réalisées ont montré que ce choix était pertinent. Mais il faut bien garder à l'esprit que cet opérateur ne donne pas à l'observateur une sensation visuellement similaire à l'observateur, il n'est pas parfait. A l'avenir, il est possible qu'un autre opérateur puisse se montrer plus performant pour notre application. Il pourra en outre être facilement implémenté et intégré au rendu HDR.

Enfin, une autre perspective intéressante est l'utilisation d'un écran HDR pour un simulateur de conduite. Si l'écran HDR est suffisamment lumineux, nous n'aurions plus besoin d'utiliser une simulation de l'éblouissement ni un opérateur de reproduction de tons, et pourrions directement afficher les images HDR à l'écran lorsqu'elles sont rendues. Pour le LEPSiS, le rendu HDR développé permettrait la création du premier simulateur de conduite HDR, qui devrait se révéler plus performant pour des études de perception de l'environnement routier.

Perspectives sur les évaluations

De nombreux travaux restent à effectuer sur l'évaluation des images et vidéos manipulées. Pour évaluer les performances des opérateurs en situation dynamique, nous avons utilisé des vidéos de synthèse, mais sans référence à la réalité. Cette évaluation a permis de mettre en évidence l'importance du modèle visuel de l'opérateur d'Irawan et *al.*, mais elle a été évaluée par rapport au sentiment de fidélité éprouvé par les sujets plutôt que par rapport à la fidélité sur la réalité. Là encore, l'utilisation d'un écran HDR pourrait se montrer efficace, car les sujets pourraient alors évaluer des vidéos traitées

par des simulations de l'éblouissement et des opérateurs de reproduction de tons avec la vidéo de référence HDR, affichée sur l'écran HDR.

D'autre part, si les évaluations subjectives sont devenues courantes, en faisant appel à des sujets pour noter ou classer des images provenant d'opérateurs de reproduction de tons, il n'existe toujours pas de métrique permettant d'évaluer la qualité de ces mêmes images de manière objective et automatique. Une métrique reste donc toujours à développer pour évaluer la qualité des opérateurs de reproduction de tons (éventuellement en fonction d'une application donnée). Mais d'autres métriques pourraient également apparaître, dont par exemple une métrique pour évaluer la qualité d'images HDR. Elle permettrait de montrer qu'une image produite par une nouvelle méthode de rendu HDR est meilleure (ou non) par rapport aux autres méthodes de la littérature, évitant ainsi de devoir recourir aux évaluations subjectives avec sujets, souvent longues et fastidieuses.

Conclusion

La modélisation physique utilisée permet de simuler dans des scènes virtuelles des luminances proches de celles de la réalité. Les algorithmes utilisés pour leur visualisation à l'écran permettent de plus aux observateurs d'avoir la meilleure perception possible de ces luminances, malgré les capacités d'affichage limitées des écrans. Les notes obtenues sur les deux expérimentations restant toutes inférieures à 7.5 / 10, nous ne pouvons cependant pas dire que la perception de l'observateur sera la même à l'écran et dans la réalité. Pour les expérimentations sur simulateur de conduite, un biais persiste donc, qu'il convient de prendre en compte en fonction de son impact sur les expérimentations. Par rapport au rendu LDR qui existait avant notre travail, nous pensons cependant que les expérimentateurs pourront profiter, via ce rendu HDR, d'un écart plus faible entre la simulation et la réalité.

Bibliographie

- [Adelson 82] E. H. Adelson. *Saturation and adaptation in the rod system*. Vision Research, vol. 22, pages 1299–1312, 1982.
- [Alpern 97] M. Alpern. *Rhodopsin kinetics in the human eye*. Journal of Physiology, vol. 217, pages 447–471, 1997.
- [APR 10] APR. *Les données de l'accidentologie*. www.preventionroutiere.asso.fr/accidentologie_securite_routiere.aspx, 2010. Association Prévention Routière.
- [Ashikhmin 06] M. Ashikhmin & J. Goyal. *A reality check for tone mapping operators*. ACM Transactions on Applied Perception, vol. 3, no. 4, pages 399–411, 2006.
- [Autodesk 10] Autodesk. *Autodesk Products*. <http://usa.autodesk.com/>, 2010.
- [Baker 49] H. Baker. *The course of foveal light adaptation measured by threshold intensity increment*. Journal of Optometry, vol. 39, pages 172–179, 1949.
- [Blackwell 46] H. R. Blackwell. *Contrast thresholds of the human eye*. Journal of the Optical Society of America, vol. 36(11), pages 624–643, 1946.
- [Blender 10] Blender. *Blender homepage*. www.blender.org/, 2010.
- [CAD 07] CAD. *Rendu photo-réaliste et temps-réel*. CAD Magazine, no. 139, pages 38–43, 2007.
- [Cadík 08] M. Cadík, M. Wimmer, L. Neumann & A. Artusi. *Evaluation of HDR tone mapping methods using essential perceptual attributes*. In Journal of Visual Communication and Image Representation, volume 32, pages 330–349, 2008.
- [Campbell 66] F. Campbell & R. Gubisch. *Optical quality of the human eye*. Journal of Physiology 186, pages 558–578, 1966.
- [Chiu 93] K. Chiu, M. Herf, P. Shirley, S. Swamy, C. Wang & K. Zimmerman. *Spatially Nonuniform Scaling Functions for High Contrast Images*. In Proceedings of Graphics Interface 1993, pages 245–253, 1993.
- [Choudhury 03] P. Choudhury & J. Tumblin. *The trilateral filter for high contrast images and meshes*. In Proceedings of the Eurographics Symposium on Rendering, pages 186–196, 2003.

- [CISR 10] CISR. *Dossier de presse*. www.securiteroutiere.equipement.gouv.fr/cnsr/2_documents_page_travaux/CISR_fev_2008.pdf, 2010. Comité interministériel de la sécurité routière.
- [Couleur 10] Blog Couleur. *Le blog de la couleur*. <http://www.blog-couleur.com/>, 2010.
- [Crytek 10] Crytek. *The Crytek webpage*. www.crytek.com/technology/cryengine-2/specifications/, 2010.
- [Darty 10] Darty. *Catalogue Darty*. www.darty.com, 2010.
- [Debevec 97] P. Debevec & J. Malik. *Recovering high dynamic range radiance maps from photographs*. In Proceedings of SIGGRAPH, pages 369–378. ACM, 1997.
- [Drago 02] F. Drago, W. L. Martens, K. Myszkowski & H.-P. Seidel. *Perceptual Evaluation of Tone Mapping Operators with Regard to Similarity and Preference*. Rapport technique, Max Planck Institute, 2002.
- [Drago 03] F. Drago, K. Myszkowski, T. Annen & N. Chiba. *Adaptive Logarithmic Mapping For Displaying High Contrast Scenes*. In Proc. of EUROGRAPHICS 2003, volume 22 of *Computer Graphics Forum*, pages 419–426, 2003.
- [Dumont 00] E. Dumont. *Semi-Monte Carlo light tracing applied to the study of road visibility in fog*. In Monte Carlo and Quasi-Monte Carlo Methods 1998, pages 177–187. Springer-Verlag, 2000.
- [Durand 00] F. Durand & J. Dorsey. *Interactive tone mapping*. In Proceedings of Eurographics Symposium on Rendering, pages 219–230, 2000.
- [Durand 02] F. Durand & J. Dorsey. *Fast bilateral filtering for the display of high-dynamic-range images*. In Proceedings of SIGGRAPH, pages 257–266. ACM, 2002.
- [Fairchild 02] M.D. Fairchild & G.M. Johnson. *Meet iCAM : A next-generation color appearance model*. IS&T/SID 10th Color Imaging Conference, pages 33–38, 2002.
- [Fattal 02] R. Fattal, D. Lischinski & M. Werman. *Gradient domain high dynamic range compression*. In Proceedings of SIGGRAPH, pages 249–256. ACM, 2002.
- [Ferwerda 96] J. A. Ferwerda, S. N. Pattanaik, P. Shirley & D. P. Greenberg. *A Model of Visual Adaptation for Realistic Image Synthesis*. Proceedings of SIGGRAPH, pages 249–258, 1996.
- [Giannopulu 08] I. Giannopulu, R. Bertin, R. Brémond, S. Espié & Z. Kapoula. *Visuomotor strategies using driving simulators in virtual and pre-recorded environment*. Advances in Transportation Studies, vol. 14, pages 49–56, 2008.
- [Got 10] C. Got. *Editorial*. www.securite-routiere.org/editoriaux/2010fevrier.html, 2010.

- [Grall 97] Y. Grall. *La fonction visuelle*. Techniques de l'ingénieur, 1997.
- [Grassvalley 10] Grassvalley. *Viper FilmStream Camera*. <http://www.grassvalley.com/products/viper>, 2010.
- [Grave 06] J. Grave. Conception et évaluation d'un opérateur de reproduction de tons pour des études de visibilité routières. Thèse. LCPC, 2006.
- [Grave 08] J. Grave & R. Brémond. *A tone mapping algorithm for road visibility experiments*. ACM Transactions on Applied Perception, vol. 5, no. 2, page art. 12, 2008.
- [Gruyer 10] D. Gruyer, N. Hiblot, P. Desouza, G. Le Boudec & B. Monnier. *A new generic virtual platform for cameras modeling*. In VISION conference, 2010.
- [Hollins 73] M. Hollins & M. Alpern. *Dark Adaptation and Visual Pigment Regeneration in Human Cones*. Journal of General Physiology, no. 62(4), pages 430–447, 1973.
- [Hood 86] D. C. Hood & M. A. Finkenstien. Sensitivity to light, chapitre 5, Handbook of perception and human performance., Wiley, NY, 1986.
- [Hunt 95] R. W. G. Hunt. The reproduction of color, 5th ed. Fountain Press Ltd., 1995.
- [Irawan 05] P. Irawan, J. A. Ferwerda & S. R. Marschner. *Perceptually based tone mapping of high dynamic range image streams*. In Proceedings of Eurographics Symposium on Rendering, pages 231–242, 2005.
- [Itti 98] L. Itti, C. Koch & E. Niebur. *A Model of Saliency-Based Visual Attention for Rapid Scene Analysis*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 20, no. 11, pages 1254–1259, 1998.
- [Itti 00] L. Itti & C. Koch. *A saliency-based search mechanism for overt and covert shifts of visual attention*. Vision Research, vol. 40, pages 1489–1506, 2000.
- [Itti 01] L. Itti & C. Koch. *Computational Modeling of Visual Attention*. Nature Reviews Neuroscience, vol. 2, no. 3, pages 194–203, 2001.
- [Itti 03] L. Itti, N. Dhavale & F. Pighin. *Realistic Avatar Eye and Head Animation Using a Neurobiological Model of Visual Attention*. In Proc. SPIE Int. Symp. on Optical Science and Technology, volume 5200, pages 64–78. SPIE Press, Aug 2003.
- [Itti 04] L. Itti. *Automatic Foveation for Video Compression Using a Neurobiological Model of Visual Attention*. IEEE Transactions on Image Processing, vol. 13, no. 10, pages 1304–1318, 2004.
- [Itti 05] L. Itti, G. Rees & J. K. Tsotsos. (ed.) neurobiology of attention. Elsevier, 2005.

- [Jacobs 10] A. Jacobs. *WebHDR*. <http://luminance.londonmet.ac.uk/webhdr/about.shtml>, 2010.
- [Johnson 03] G. Johnson & M. Fairchild. *Rendering hdr images*. pages 36–41, 2003.
- [Kakimoto 05] M. Kakimoto, K. Matsuoka, T. Nishita, T. Naemura & H. Harashima. *Glare Generation Based on Wave Optics*. Computer Graphics Forum, vol. 24, no. 2, pages 185–193, 2005.
- [Kang 03] S. B. Kang, M. Uyttendaele, S. Winder & R. Szeliski. *High Dynamic Range Video*. ACM Transactions on Graphics, vol. 22(3), 2003.
- [Kawase 04] M. Kawase. *Practical implementation of high dynamic range rendering*. In Game Developers Conference, 2004.
- [Koblentz 96] Koblentz. *Analyse du comportement du conducteur sur autoroutes*. 1996.
- [Krawczyk 05] G. Krawczyk, K. Myszkowski & H.-P. Seidel. *Perceptual effects in real-time tone mapping*. In Proceedings of SCCG, pages 195–202. ACM, 2005.
- [Kuang 07a] J. Kuang, G.M. Johnson & M.D. Fairchild. *iCAM06 : A refined image appearance model for HDR image rendering*. Journal of Visual Communication and Image Representation, vol. 18, pages 406–414, 2007.
- [Kuang 07b] J. Kuang, H. Yamaguchi, C. Liu, G. M. Johnson & M. D. Fairchild. *Evaluating HDR rendering algorithms*. ACM Transactions on Applied Perception, vol. 3, no. 3, pages 286–308, 2007.
- [Ledda 04] P. Ledda, L. P. Santos & A. Chalmers. *A local model of eye adaptation for high dynamic range images*. In Proceedings of AFRIGRAPH, pages 151–160. ACM, 2004.
- [Ledda 05] P. Ledda, A. Chalmers, T. Troscianko & H. Seetzen. *Evaluation of tone mapping operators using a high dynamic range display*. In Proceedings of SIGGRAPH, pages 640–648. ACM, 2005.
- [LeMeur 06] O. LeMeur, P. LeCallet, D. Barba & D. Thoreau. *A coherent computational approach to model bottom-up visual attention*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 28, no. 5, pages 802–817, 2006.
- [Li 99] X. Li & A. Gilchrist. *Relative area and relative luminance combine to anchor surface lightness values*. Perception & Psychophysics 61, vol. 5, pages 771–785, 1999.
- [Lindbloom 09] B. J. Lindbloom. *Useful Color Equations*. www.brucelindbloom.com, 2009.
- [LIVIC 10] LIVIC. *Laboratoire sur les Interactions Véhicules-Infrastructure-Conducteurs*. www.inrets.fr/ur/livic/, 2010.

- [Lugand 09] E. Lugand. Pratique de l'image-based lighting. Utilisation des photographies à haute dynamique pour l'illumination des objets de synthèse. Rapport de fin d'études. Ecole Nationale Supérieure Louis Lumière, 2009.
- [Mann 95] S. Mann & R. W. Picard. *Being Undigital with Digital Cameras : Extending Dynamic Range by Combining Differently Exposed Pictures*. IS&T's 48th Annual Conference, 1995.
- [Mantiuk 08] Rafal Mantiuk, Scott Daly & Louis Kerofsky. *Display Adaptive Tone Mapping*. ACM Transactions on Graphics, vol. 27, no. 3, page article no. 68, 2008.
- [Matković 97] K. Matković. Tone mapping techniques and color image difference in global illumination. The Institute of Computer Graphics, Technische Universität Wien, 1997.
- [Mayeur 10] A. Mayeur, R. Brémond & C. Bastien. *The effect of the driving activity on target detection as a function of the visibility level*. Transportation Research, part F, vol. 13, pages 115–128, 2010.
- [MDEVO 10] MDEVO & Maison départementale de l'éducation du Val d'Oise. *Le simulateur de conduite 2 roues motorisés*. www.cddp95.ac-versailles.fr/route/spip.php?article126, 2010.
- [Meyzonette 02] J. L. Meyzonette. Radiométrie et détection optique. Edp sciences edition, 2002.
- [Mitchell 06] J. Mitchell, G. McTaggart & C. Green. *Shading in valve's source engine*. In Proceedings of SIGGRAPH, pages 129–142. ACM, 2006.
- [Mitsunaga 99] T. Mitsunaga & S. K. Nayar. *Radiometric Self Calibration*. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 1999.
- [Moon 45] P. Moon & D. E. Spencer. *The visual effect of non uniform surrounds*. Journal of the Optical Society of America, vol. 35, no. 3, pages 233–248, 1945.
- [Naka 66] K. I. Naka & W. A. H. Rushton. *S-potential from colour units in the retina of fish (cyprinidae)*. Journal of Physiology, vol. 185, pages 536–555, 1966.
- [Nvidia 10] Nvidia. *CUFFT Library*. http://developer.download.nvidia.com/compute/cuda/3_0/toolkit/docs/CUFFT_Library_3.0.pdf, 2010.
- [Oblivion 10] Oblivion. *The Oblivion : The Elder Scrolls IV official webpage*. <http://oblivion.fr.ubi.com/>, 2010.
- [Ogre3D 10] Ogre3D. *The Ogre3D webpage*. www.ogre3d.org/, 2010.
- [OSG 10] OSG. *The OpenSceneGraph webpage*. www.openscenegraph.org, 2010.

- [OsgPPU 10] OsgPPU. *The osg Post Processing Units webpage*. <http://projects.tevs.eu/osgPPU>, 2010.
- [Pattanaik 98] S. N. Pattanaik, J. A. Ferwerda, M. D. Fairchild & D. P. Greenberg. *A Multiscale Model of Adaptation and Spatial Vision for Realistic Image Display*. In Proceedings of SIGGRAPH, pages 287–298. ACM, 1998.
- [Pattanaik 00] S. N. Pattanaik, J. Tumblin, H. Yee & D. P. Greenberg. *Time-dependent visual adaptation for fast realistic image display*. In Proceedings of SIGGRAPH, pages 47–54. ACM, 2000.
- [Peli 90] E. Peli. *Contrast in Complex Images*. Journal of the Optical Society of America, vol. 7(10), pages 2032 – 2040, 1990.
- [PFStmo 10] PFStmo. *PFStmo : tone mapping operators*. <http://www.mpi-inf.mpg.de/resources/tmo/>, 2010.
- [Philips 99] Philips. *Philirama 1998/99*. Philips, 1999.
- [Phong 75] T. B. Phong. *Illumination for computer generated pictures*. Communication to the ACM, vol. 18, no. 6, pages 311–317, 1975.
- [Pixim 10] Pixim. *Pixim Digital Pixel System*. <http://www.pixim.com/>, 2010.
- [Polyphony 10] Digital Polyphony. *Gran Turismo 5*. <http://www.gran-turismo.com/>, 2010.
- [Rahman 96] Z.-U. Rahman, D. J. Jobson & G. A. Woodell. *A multiscale retinex for color rendition and dynamic range compression*. In SPIE Proceedings : Applications of Digital Image Processing XIX, vol. 2847, 1996.
- [Reinhard 02] E. Reinhard, M. Stark, P. Shirley & J. Ferwerda. *Photographic tone reproduction for digital images*. In Proceedings of SIGGRAPH. ACM, 2002.
- [Reinhard 05a] E. Reinhard & K. Devlin. *Dynamic Range Reduction Inspired by Photoreceptor Physiology*. IEEE Transactions on Visualization and Computer Graphics, vol. 11, no. 1, 2005.
- [Reinhard 05b] E. Reinhard, G. Ward, S. N. Pattanaik & P. Debevec. *High dynamic range imaging : acquisition, display, and image-based lighting*. Morgan Kaufmann, 2005.
- [Remande 09] C. Remande. *"Moins d'éclairage pour moins d'accidents ?" Est-ce raisonnable ?* Lux, vol. 252, pages 1–7, 2009.
- [Ritschel 09] T. Ritschel, M. Ihrke, J. Revall Frisvad, J. Coppens, K. Myszkowski & H.P. Seidel. *Temporal Glare : Real-Time Dynamic Simulation of the Scattering in the Human Eye*. Computer Graphics Forum, vol. 28, no. 2, pages 183–192, 2009.
- [Rodieck 03] R. W. Rodieck. *La vision*. De Boeck, 2003.

- [Rushton 86] W. A. H. Rushton & D. I. A. MacLeod. *The equivalent background of bleaching*. Perception, vol. 15, pages 689–703, 1986.
- [Schlick 94] C. Schlick. *Quantization Techniques for Visualization of High Dynamic Range Pictures*. In Fifth Eurographics Workshop on Rendering, pages 7–20. Springer-Verlag, 1994.
- [Schoettle 03] B. Schoettle, M. Sivak, M. J. Flannagan & W. J. Kosmatka. *A Market-Weighted Description of Low-Beam Headlighting Patterns in Europe*. Rapport technique, University of Michigan Transportation Research Institute, 2003.
- [Spencer 95] G. Spencer, P. Shirley, K. Zimmerman & D. P. Greenberg. *Physically based glare effect for digital images*. In Proceedings of SIGGRAPH, pages 325–334. ACM, 1995.
- [Spheron 10] Spheron. *SpheroCam HDR*. <http://www.spheron.com/>, 2010.
- [Spillman 90] L. Spillman & J. S. Werner. *Visual perception : The neurophysiological foundations*. Academic Press, 1990.
- [Stevens 60] S. S. Stevens & J. C. Stevens. *Brightness Function : Parametric Effects of Adaptation and Contrast*. Journal of the Optical Society of America, vol. 53, no. 11, 1960.
- [Stevens 63] S. S. Stevens & J. C. Stevens. *Brightness Function : Effects of Adaptation*. Journal of the Optical Society of America, vol. 53, no. 3, pages 375–385, 1963.
- [Tumblin 93] J. Tumblin & H.E. Rushmeier. *Tone reproduction for realistic images*. IEEE Computer Graphics and Applications, vol. 13, no. 6, pages 42–48, 1993.
- [Tumblin 99] J. Tumblin & G. Turk. *LCIS : A Boundary Hierarchy For Detail-preserving Contrast Reduction*. In A. Rockwood, editeur, Siggraph 1999, Computer Graphics Proceedings, Annual Conference Series, pages 83–90, 1999.
- [Walther 06] D. Walther. www.saliencytoolbox.net/, 2006.
- [Wanvik 10] P. O. Wanvik. *Effects of road lighting : An analysis based on Dutch accident statistics 1987 - 2006*. In Accident Analysis & Prevention, volume 41(1), pages 123 – 128, 2010.
- [Ward 94] G. Ward. *A contrast based scale-factor for image display*, pages 415–421. Graphic Gems IV, Academic Press Professional, San Diego, CA, 1994.
- [Ward 97] G. Ward, H. Rushmeier & C. Piatko. *A visibility matching tone reproduction operator for high dynamic range scenes*. IEEE Transactions on Visualization and Computer Graphics, vol. 3, no. 4, pages 291–306, 1997.

- [Ward 98] G. Ward & R. Shakespeare. *Rendering with radiance : the art and science of lighting visualization*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1998.
- [Ward 03] G. Ward. *Fast, Robust Image Registration for Compositing High Dynamic Range Photographs from Hand-held Exposures*. *Journal of Graphics Tools*, vol. 8(2), pages 17–30, 2003.
- [Wikipedia 10a] Wikipedia. *Lux (unité)*, *article*. http://fr.wikipedia.org/wiki/Lux_
- [Wikipedia 10b] Wikipedia. *Simulateur de conduite*. http://fr.wikipedia.org/wiki/Simulateur_de_conduite, 2010.
- [Yoshida 05] A. Yoshida, V. Blanz, K. Myszkowski & H.-P. Seidel. *Perceptual evaluation of tone mapping operators with real world scenes*. In *Proceedings of the SPIE*, volume 5666, pages 192–203, 2005.
- [Yoshida 07] Akiko Yoshida, Volker Blanz, Karol Myszkowski & Hans-Peter Seidel. *Testing Tone Mapping Operators with Human-Perceived Reality*. *Journal of Electronic Imaging*, vol. 16, no. 1, pages 1–14, 2007.
- [Yoshida 08] A. Yoshida, M. Ihrke, R. Mantiuk & H.-P. Seidel. *Brightness of the glare illusion*. pages 83–90, 2008.
- [Zacharias 10] V. Zacharias. *Notions de photométrie*. http://perso.magic.fr/zacharias/art_tech/techniques/photometrie/photometrie.htm, 2010.