HAL
open science

L'échantillonnage spatio-chromatique dans la rétine humaine et les caméras numériques

Brice Chaix de Lavarène

To cite this version:

Brice Chaix de Lavarène. L'échantillonnage spatio-chromatique dans la rétine humaine et les caméras numériques. Traitement du signal et de l'image [eess.SP]. Université Joseph-Fourier - Grenoble I, 2007. Français. NNT : . tel-00276483

HAL Id: tel-00276483

https://theses.hal.science/tel-00276483

Submitted on 30 Apr 2008

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.

Université Joseph Fourier

No. attribué par la bibliothèque

THESE

pour obtenir le grade de

DOCTEUR DE L'UNIVERSITE JOSEPH FOURIER

Spécialité : Signal - Image - Parole - Télécom

préparée au laboratoire "Grenoble Image Parole Signal Automatique" (GIPSA-Lab)

dans le cadre de l'École Doctorale :
Électronique, Électrotechnique, Automatique, Traitement du Signal

présentée et soutenue publiquement

par

Brice CHAIX DE LAVARÈNE

le 22 novembre 2007

Titre :

L'ÉCHANTILLONNAGE SPATIO-CHROMATIQUE DANS LA RÉTINE HUMAINE ET LES CAMÉRAS NUMÉRIQUES

Directeur de thèse : Jeanny Hérault
Co-Directeur de thèse : David Alleysson

COMPOSITION DU JURY

Pierre-Yves Coulon	:	Président
Christine Fernandez-Maloigne	:	Rapporteur
Sabine Süsstrunk	:	Rapporteur
Michel Paindavoine	:	Examinateur
Jeanny Hérault	:	Directeur de thèse
David Alleysson	:	Co-Directeur de thèse

Remerciements

J'adresse tout d'abord mes remerciements à mes encadrants de thèse David Alleysson et Jeanny Hérault pour m'avoir permis de travailler sur ce sujet passionnant qu'est la couleur. Je leur suis reconnaissant pour m'avoir fait confiance et pour avoir su me donner une grande autonomie durant l'accomplissement de ces travaux.

Je remercie également les membres du Jury d'avoir accepté d'évaluer mes travaux, et en particulier Christine Fernandez-Maloigne et Sabine Süsstrunk qui en ont été les rapporteurs.

Je remercie tous les membres du laboratoire, permanents, thésards ou stagiaires avec qui j'ai eu l'occasion de discuter ou d'échanger des idées. En particulier Nicolas, avec qui j'ai partagé mon bureau pendant ces trois ans, Nathalie pour ses remarques constructives sur mes travaux, Marino, qui veille sur les doctorants à Viallet, ainsi que l'ensemble de l'équipe *perception*.

Je remercie chaleureusement Pauline pour son soutien de tous les instants et sa relecture attentive du manuscrit. Je remercie le groupe La Vérue pour les bons moments passés ensemble (Mokozore, Antxoka, Numéro 2 et bien sûr Boulette), l'équipe Boxson ainsi que tous les amis musiciens que j'ai pu rencontrer durant ces trois ans, y compris (et surtout) les *c'est ça ouais*. Je témoigne également mon amitié à Tiseb et à mes amis parisiens (ou peut-être devrais-je dire citoyens du monde !).

Enfin, je dédie cette thèse aux membres de ma famille.

Table des matières

Chapitre 1 Introduction **9**

Chapitre 2 La vision des couleurs dans le système visuel humain **13**
1 La trivalence de la vision des couleurs 13
1.1 Historique 13
1.2 Notions d'intensité et de chrominance 18
2 Anatomie du système visuel humain de bas niveau 21
2.1 Optique de l'oeil 22
2.2 Le corps genouillé latéral : des voies fonctionnelles parallèles . . 22
2.3 La rétine 23
2.4 Le cortex visuel primaire 36
3 Performances et limites du système visuel humain 37
4 Les modèles neurocomputationnels de vision des couleurs 41
5 Conclusion 46

Chapitre 3 L'acquisition d'une image couleur par un appareil photographique numérique **49**
1 Anatomie d'un appareil photographique numérique 50
1.1 Le capteur photosensible 50
1.2 La matrice de filtres couleur 51
1.3 Le système optique 53
1.4 Autres traitements algorithmiques 54
2 Etat de l'art du démosaïçage dans les caméras numériques 54
2.1 Méthodes simples 56
2.2 Les modèles spectraux 57
2.3 Les modèles spatiaux 65
2.4 Evaluation des performances des algorithmes 69
3 Conclusion 70

Chapitre 4 Echantillonnage irrégulier et interpolation **73**
1 Théorèmes d'échantillonnage 73
1.1 Echantillonnage régulier 73
1.2 Transposition au cas irrégulier : le théorème de Yen 76

1.3 Comparaison d'une interpolation à partir d'un sous-échantillonnage régulier vs. irrégulier . . . 79
2 Analyse spectrale . . . 79
2.1 Approche intuitive . . . 80
2.2 Echantillonnage de Poisson . . . 82
2.3 Echantillonnage à gigue . . . 82
2.4 Echantillonnage de Poisson à disque . . . 85
2.5 Conséquence de la contrainte de grille régulière . . . 86
2.6 L'échantillonnage irrégulier ou le théorème de Shannon local . . . 87
3 Algorithmes d'interpolation de signaux échantillonnés irrégulièrement . . . 90
3.1 Interpolation de Shepard . . . 90
3.2 Interpolation affine par triangles . . . 91
3.3 Exploitation de la contrainte de bande limitée . . . 91
3.4 Convolution normalisée . . . 93
3.5 Splines . . . 98
3.6 Interpolation orientée contour . . . 99
3.7 Evaluation des algorithmes . . . 99
4 Génération de mosaïques aléatoires . . . 101
4.1 Mosaïque de type Poisson . . . 101
4.2 Mosaïque de type Poisson à disque . . . 101
4.3 Matrice irrégulière basée sur un motif périodique . . . 102
5 Conclusion . . . 103

Chapitre 5 Démosaïçage généralisé et implications pour le système visuel humain 105
1 Modèle intensité-chrominance dans le cas d'une mosaïque quelconque . . . 106
2 Comparaison des performances des différents motifs d'échantillonnage pour un capteur couleur . . . 107
2.1 Méthode linéaire de démosaïçage par minimisation de l'erreur quadratique . . . 107
2.2 Evaluation des performances des différentes mosaïques . . . 117
3 Nouveau modèle de séparation des canaux spatiaux et chromatiques . . . 120
3.1 Approche par estimation de l'intensité . . . 120
3.2 Approche par estimation de la chrominance . . . 131
4 Conséquences du modèle pour le système visuel humain . . . 140
4.1 Structure des champs récepteurs des cellules ganglionnaires . . . 141
4.2 L'implication du cortex visuel primaire . . . 146
4.3 Apprentissage non supervisé de la mosaïque des cones . . . 147
5 Conclusion . . . 148

Chapitre 6 Conclusion et perspectives 151

Annexe A Base d'images utilisée 157

Annexe B Quelques exemples de résultats **159**
1 Recontruction minimisant l'erreur quadratique moyenne par apprentissage sur la base d'images Kodak 159
2 Reconstruction utilisant le modèle proposé 160

Chapitre 1

Introduction

DEUX théories de vision des couleurs se sont affrontées au XIXème siècle. La première, initiée par Young, défendait l'idée que trois récepteurs chromatiques, correspondant chacun à une couleur primaire - rouge, vert ou bleu - étaient nécessaires et suffisants pour expliquer la perception des couleurs. Les expériences d'égalisation des couleurs à partir de trois primaires, réalisées par Maxwell et Helmholtz, corroboraient cette théorie. Cependant, celle-ci n'explique pas pourquoi il est impossible de *percevoir* une couleur comme étant un mélange de bleu et de jaune, ou comme un mélange de rouge et de vert. La seconde théorie, suggérée par Goethe et reprise plus tard par Hering, considérait qu'il n'y avait pas trois, mais quatre couleurs fondamentales, formant deux couples antagonistes : le bleu et le jaune, ainsi que le rouge et le vert.

Grâce aux progrès de nombreuses disciplines telles que l'anatomie ou la physiologie au cours du XXème siècle, il s'avère que ces deux théories sont complémentaires et interviennent à deux niveaux différents de la vision des couleurs. La rétine est tapissée de trois types de cônes, chacun sensible à une gamme de longueurs d'onde de la lumière visible. Ces photorécepteurs sont à la base de la trivalence de notre vision des couleurs. A partir de ces trois informations chromatiques, le système visuel crée trois axes perceptifs. L'un est achromatique, il rend compte de l'intensité lumineuse. Les deux autres sont les oppositions de couleurs rouge moins vert et bleu moins jaune. C'est pourquoi un rouge ne peut paraître verdâtre, ou encore qu'un bleu ne peut paraître jaunâtre.

S'il peut sembler aller de soi que la présence de trois types de photorécepteurs induit un espace des couleurs à trois dimensions, le système visuel est confronté à un problème de taille. A chaque position spatiale de la rétine, un seul type de cône est présent. Un cône isolé ne peut pas renseigner ni sur l'intensité ni sur la couleur de la lumière qui l'éclaire, car chaque photorécepteur est susceptible de capter une information lumineuse quelle que soit la longueur d'onde de celle-ci. Ainsi, il est impossible de savoir si une variation du signal issu d'un cône isolé correspond à une variation d'intensité lumineuse ou à une variation de couleur. Ce paradigme s'appelle le principe d'univariance de Rushton. Par conséquent, les signaux issus des cônes voisins sensibles aux autres gammes de longueurs d'onde doivent être utilisés pour construire le canal achromatique et les oppositions de couleur. Or, si deux cônes sont à des positions spatiales différentes, non seulement l'intensité de la lumière, mais également sa chromaticité, ont pu varier entre ces deux positions, posant ainsi des problèmes de cohérence des signaux comparés.

Les modèles actuels de vision des couleurs ne permettent pas d'expliquer le niveau de performance atteint par le système visuel dans la construction sans ambiguïté des voies d'intensité lumineuse et d'opposition de couleurs.

En plus de ce problème de traitement du signal, les fonctions des neurones intervenant dans la construction des voies perceptives ne sont pas entièrement comprises. En particulier, les types de connexions entre neurones rétiniens ne sont pas clairement établies et l'électrophysiologie ne permet pas à l'heure actuelle de trancher. Un modèle théorique pertinent de construction des voies d'oppositions de couleurs permettrait ainsi d'inférer les relations entre neurones.

Par ailleurs, les systèmes artificiels d'acquisition d'images couleur dits "mono-capteur" sont confrontés à la même problématique d'échantillonnage spatio-chromatique que le système visuel. Leur capteur est en effet couvert d'une matrice de filtres rouges, verts ou bleus, permettant l'échantillonnage à chaque pixel d'une seule composante couleur sur les trois nécessaires à la représentation d'une image en couleur. Ainsi, un algorithme doit interpoler chacun des canaux couleur pour que l'image en couleur puisse effectivement être affichée. Cette étape algorithmique est appelée démosaïçage[1] et peut générer des artefacts de reconstruction plus ou moins visibles selon la méthode employée.

De notre point de vue, la reconstruction d'une image à trois composantes couleur par pixel dans une caméra est équivalente à la construction des trois voies perceptives dans le système visuel humain. Le domaine des caméras numériques est donc un formidable champ d'expérimentation des modèles neurocomputationnels de vision des couleurs. Alleysson *et al.* ont proposé un modèle d'échantillonnage spatio-chromatique qui justifie le passage par l'espace intensité-chrominance pour la problématique du démosaïçage. De ce modèle est dérivé un algorithme simple et efficace basé sur l'estimation de l'intensité et des oppositions de couleurs. Cependant cette méthode s'applique aux mosaïques dont les filtres sont disposés de manière régulière, comme par exemple la mosaïque de Bayer, mais échoue lorsque l'arrangement est irrégulier. Un arrangement irrégulier des photorécepteurs est pourtant ce qui caractérise la rétine. Nous proposons dans cette thèse une généralisation de la méthode par intensité et oppositions de couleurs à toutes sortes de mosaïques. L'algorithme proposé, en plus d'être appliquable quelle que soit la configuration spatiale des filtres de couleur, se révèle très efficace d'un point de vue algorithmique. Nous discutons également des implications de ce modèle au niveau de la construction des voies perceptives du système visuel.

Enfin, la qualité de l'image reconstruite dépend également de l'arrangement spatial de la mosaïque. Nous établissons un comparatif des performances de différentes mosaïques et proposons une mosaïque offrant des performances supérieures à celles de la mosaïque de Bayer. Par ailleurs, lorsqu'une image échantillonnée régulièrement souffre de repliement spectral, le phénomène de Moiré apparaît, caractérisé par des motifs basses fréquences très réguliers et très visibles. Il est connu depuis longtemps que l'échantillonnage irrégulier permet de casser la cohérence de ces motifs et d'en réduire par conséquent la visibilité. Nous étudions dans cette thèse dans quelle mesure l'irrégulier peut s'appliquer à l'échantillonnage de la couleur. Nous verrons que la nature de l'aléatoire qui caractérise l'arrangement des cônes, si l'on en croit les études dans le domaine de l'imagerie rétinienne, n'est pas optimal.

[1] traduit du néologisme anglais *demosaicing* ou *demosaicking*

La thèse est structurée comme suit. Nous établissons au Chapitre 2 un état de l'art de la vision des couleurs. Nous décrivons en particulier l'anatomie et le comportement du système visuel, de la rétine au cortex visuel primaire. On distingue, à la sortie de la rétine, trois voies cellulaires parallèles. La voie magnocellulaire véhiculerait une information d'intensité de faible résolution spatiale. La voie koniocellulaire serait consacrée à l'opposition bleu-jaune. Enfin, la voie parvocellulaire véhiculerait à la fois les détails d'intensité lumineuse et l'opposition rouge-vert. Nous relatons également les principaux modèles neurocomputationnels de construction des oppositions de couleurs.

Le troisième chapitre est consacré à l'étude des caméras numériques. Après avoir décrit brièvement la structure d'un dispositif mono-capteur, nous dégageons les principales caractéristiques des algorithmes de démosaïçage. Nous nous attardons en particulier à la description du modèle spectral d'Alleysson *et al.*, qui servira de base au travaux présentés dans cette thèse.

Nous étudions au Chapitre 4 les propriétés de l'échantillonnage irrégulier. Nous verrons que ce dernier a pour effet de réduire la visibilité du repliement spectral par rapport à un échantillonnage régulier. Nous comparons également quelques algorithmes d'interpolation d'images échantillonnées irrégulièrement. Nous en conclurons que le critère décisif à la résolution du problème de démosaïçage irrégulier ne résidera pas dans le choix d'un algorithme d'interpolation particulier, mais plutôt dans la manière dont on saura mettre à profit la corrélation spectrale entre canaux couleur. Nous finissons le chapitre en présentant des mosaïques irrégulières vérifiant différentes propriétés statistiques, qui seront utilisées au Chapitre 5.

Au Chapitre 5, deux méthodes pour résoudre le problème du démosaïçage irrégulier sont proposées. L'une est basée sur un apprentissage supervisé permettant la minimisation linéaire de l'erreur quadratique. Nous utilisons cette méthode pour comparer les performances de différentes mosaïques. Toutefois, il y des restrictions à l'application de cette méthode. La mosaïque doit être composée du pavage d'un motif de taille réduite, du fait de la complexité algorithmique de la phase d'apprentissage. Ainsi, cette méthode ne s'applique pas au cas des mosaïques ne présentant aucune périodicité. Nous proposons donc une seconde méthode, sans restriction d'application, et qui est une généralisation de l'algorithme d'Alleysson *et al*. Cependant, alors qu'Alleysson estimait directement une intensité pleinement résolue, nous montrons qu'il peut être judicieux d'estimer uniquement les basses fréquences d'intensité dans un premier temps, puis d'estimer la chrominance et enfin l'intensité pleinement résolue. Nous discuterons de l'interprétation de ce modèle au niveau de la structure du système visuel humain, en particulier au niveau de la structure des champs récepteur des cellules ganglionnaires.

Chapitre 2

La vision des couleurs dans le système visuel humain

NOUS dressons dans ce chapitre un état de l'art de la vision des couleurs chez l'humain. La littérature dans ce domaine est extrêmement vaste, tant le nombre de discipline impliquées est grand. Cet état de l'art se veut donc modeste.

Nous commençons par faire un bref historique de la façon dont les scientifiques sont parvenus à la théorie des oppositions de couleurs et de la trivalence de la vision des couleurs. Il est remarquable qu'Hering, Young, ou encore Helmoltz, aient deviné dès le XIXème siècle les mécanismes sous-jacents de la vision des couleurs, bien avant que l'on connaisse précisément la structure anatomique et physiologique du système visuel humain. Nous décrivons ensuite les concepts d'intensité et de chrominance pour une image en couleur.

D'importantes avancées ont été faites ces cinquante dernières années. Ces avancées sont liées à des travaux dans des disciplines variées telles que l'anatomie, l'électrophysiologie, l'analyse moléculaire, l'optique adaptative (permettant des images rétiniennes à l'échelle des cônes), la psychophysique, ou encore le traitement du signal. L'étude de la science des couleurs est par nature un grand défi pluridisciplinaire. Nous décrivons donc, dans un second temps, l'anatomie et le comportement du système visuel humain que ces avancées ont permis de mettre en évidence, de la rétine jusqu'au cortex visuel primaire.

Cependant, l'état actuel des connaissances ne permet pas de déterminer tous les mécanismes fondamentaux de la vision des couleurs. Ainsi, plusieurs théories sur les mécanismes de construction des oppositions de couleurs ont ainsi vu le jour au cours de la seconde moitié du XXème siècle. Nous décrivons quelques unes de ces théories pour clore le chapitre.

1 La trivalence de la vision des couleurs

1.1 Historique

L'histoire de la science physique des couleurs commença en 1666 lorsque Newton entreprit son expérience de décomposition de la lumière. Cette expérience (Figure 2.1) consistait en la projection d'un rayon de lumière sur un écran après passage par un prisme. Le prisme, dont

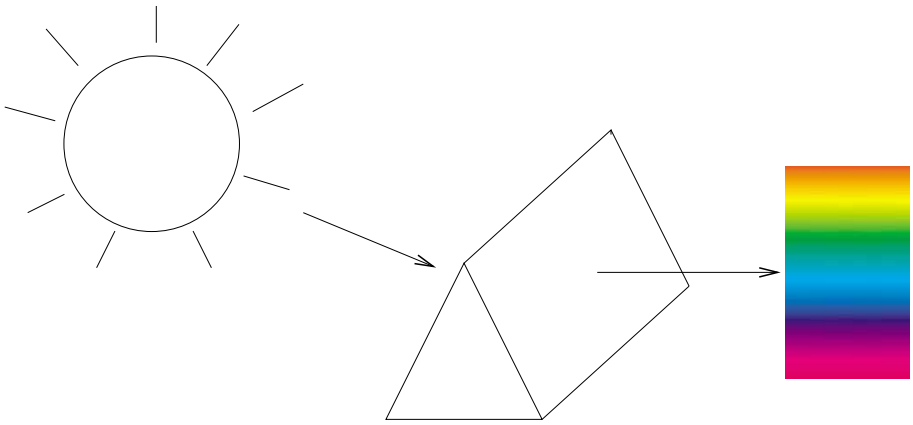


FIG. *2.1: Expérience de décomposition de la lumière de Newton.*

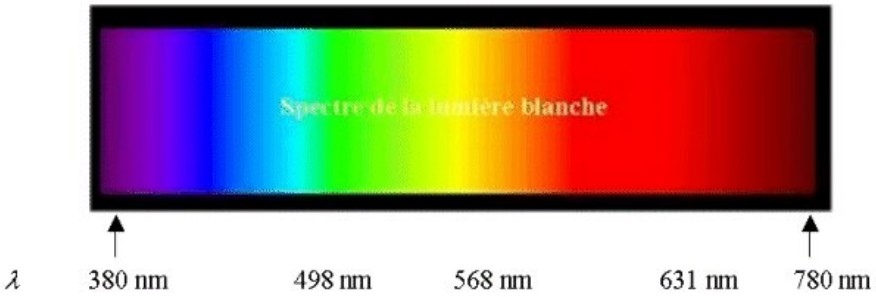


FIG. *2.2: Spectre de la lumière visible.*

l'indice de réfraction dépend de la longueur d'onde, décomposait la lumière en un spectre coloré (Figure 2.2) allant du violet (couleur la plus réfringible) au rouge (couleur la moins réfringible). L'expérience de Newton montra que la lumière blanche pouvait être décomposée en une série de raies spectrales correspondant chacune à une couleur. Il y avait pour Newton sept couleurs primaires, les sept couleurs de l'arc-en-ciel.

Au début du XIXème siècle, Thomas Young, émit l'hypothèse qu'il n'était pas nécessaire que la rétine contienne une infinité de types de photorécepteur, chacun accordé à une couleur, et que seulement trois types de photorécepteurs suffisait pour créer les sensations de n'importe quelle couleur. Ainsi il y aurait trois couleurs primaires. C'était la naissance de la théorie trichromatique de la vision. Newton avait compris que le mélange de deux couleurs pures en créait une 3ème et que cette couleur composite ne pouvait pas être distinguée de la même couleur non composite, c'est-à-dire la même couleur monochromatique et indivisible. Cependant, Newton en cherchait la cause dans la physique intrinsèque de la lumière, c'est-à-dire une cause extérieure à la perception humaine. Young fit un grand pas en avant idéologique dans la compréhension des couleurs en imputant la cause non pas à la physique de la couleur mais à la structure de notre système visuel. Hermann von Helmoltz reprit l'idée de Young en 1850 [von Helmholtz 67], et poussa la théorie plus loin, si bien qu'on parla désormais de la théorie de Young-Helmoltz. Une lumière rouge stimulerait plus fortement le récepteur associé aux gammes de longeurs d'onde proche du rouge, et ainsi le cerveau créerait la sensation rouge.

Mais c'est à James Maxwell que l'on doit la démonstration expérimentale, en 1859, de la trivalence de la vision des couleurs. Par son expérience d'égalisation (Figure 2.3), Maxwell

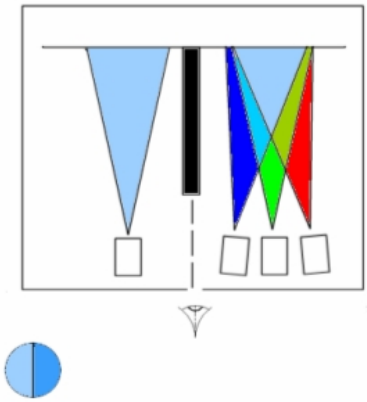


FIG. 2.3: *Expérience d'égalisation des couleurs de Maxwell. Un observateur doit égaliser une lumière par la combinaison de trois primaires. L'observateur contrôle l'intensité de chacune des primaires.*

montra qu'une lumière quelconque pouvait être égalisée visuellement par l'ajustement de l'intensité de trois lumières particulières, appelées primaires. Les primaires doivent être indépendantes, c'est-à-dire choisie de telle manière qu'avec deux d'entre elles on ne puisse obtenir la troisième. Les primaires sont généralement réparties sur le spectre, et on les appelle rouge, vert et bleu pour leur gamme de longueur d'onde préférentielle. Maxwell représentait les couleurs sur un triangle équilatéral dont les primaires constituaient les sommets. Cependant on se rend compte des limites de cette représentation à deux dimensions. Par exemple un jaune saturé sera situé à l'extérieur du triangle, au-delà de la ligne rouge-vert.

Dès la fin du XVIIIème siècle, Goethe, qui s'intéressait à la perception des couleurs, avait remarqué que la théorie Newtonnienne des couleurs ne permettait pas de rendre compte de nombreuses illusions colorées, et que la notion de couleur ne pouvait pas être séparée de la notion de perception. Pour illustrer cela, citons un passage du traité des couleurs de Goethe [Goethe 10] : "*Un soir, me trouvant dans une auberge, je regardai quelque temps une servante de taille harmonieuse, au teint blanc éblouissant, aux cheveux noirs, et vêtue d'un corselet écarlate. Elle était entrée dans ma chambre, et je la fixais à une certaine distance, et dans la pénombre. Dès qu'elle fut sortie, je distinguais sur le mur blanc en face de moi, un visage noir entouré d'une auréole claire, et les vêtements de la silhouette nettement dessinée étaient d'un beau vert marin.*" Goethe opposait dans sa théorie des couleurs dites positives, le rouge et le jaune, à des couleurs dites négatives, le bleu et le vert. Sa théorie ne fut pas prise au sérieux par la communauté scientifique de l'époque. Il fallut attendre 1874, pour que le physiologiste allemand Ewald Hering [Hering 74] exposa une théorie similaire à celle de Goethe, et qui semblait en opposition avec celle de Young-Helmholtz. Hering avait constaté que le mélange - additif - du rouge et du vert, ainsi que celui du bleu et du jaune, donnaient une teinte désaturée, proche de l'achromatique. Comme si l'action de ces couleurs s'annulaient mutuellement. Par ailleurs Hering remarqua également qu'aucune couleur ne pouvait être à la fois rougeâtre et verdâtre, ou encore bleuâtre et jaunâtre. Une autre constatation de Hering est que le mélange du rouge et du vert donnait du jaune, mais aucune nuance de rouge ni de vert ne peut être perçue lorsque l'on observe du jaune. Hering en déduisit qu'il n'y avait pas trois types de détecteur dans la rétine, mais plutôt quatre, respectivement pour le rouge, le vert, le bleu et le jaune. Le rouge serait antagoniste (en opposition) au vert et le bleu serait antagoniste au jaune. C'est la théorie

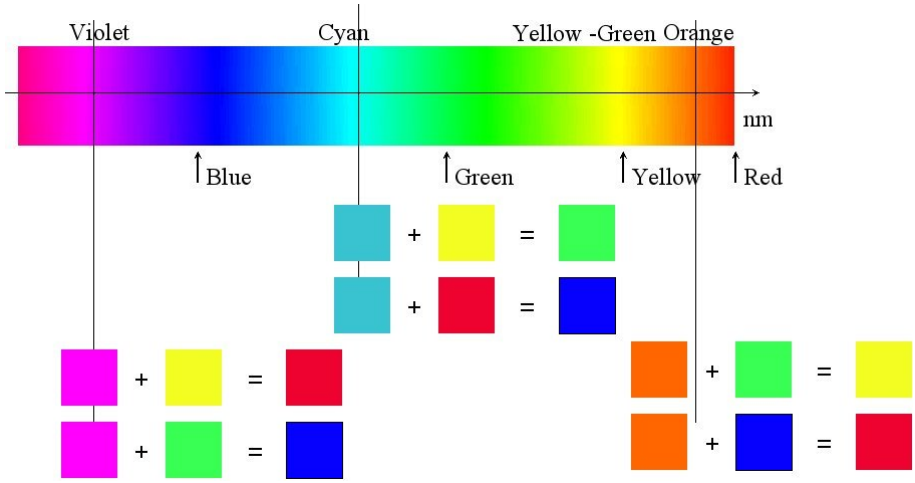


FIG. *2.4: Illustration de l'expérience de Jameson et Hurvich. L'ajout d'une certaine quantité de couleur à une lumière monochromatique permet d'obtenir des teintes uniques.*

des *oppositions de couleurs*. Mais dès lors que les couleurs s'annulent pour donner naissance au blanc, Hering considérait que c'est justement parce que les couleurs antagonistes rendaient leurs effets mutuellement impossibles que le blanc pouvait s'exprimer. Ainsi le blanc et sa couleur antagoniste, le noir, étaient dans la théorie de Hering des couleurs à parts entières, qui s'ajoutait aux quatre précédentes. La couleur serait alors composée de trois couples de couleurs antagonistes.

Ainsi deux théories du mécanisme de la vision des couleurs coexistèrent à la fin du XIXème siècle. Pour l'une, la théorie de Young-Helmholtz, la rétine était composée de trois types de photorécepteurs, et l'excitation des trois récepteur donnait le blanc. Pour l'autre, la théorie de Hering, un récepteur ne pouvait pas mesurer à la fois une couleur et sa couleur antagoniste, et des oppositions étaient ainsi créées entre le bleu et le jaune, le rouge et le vert, et le blanc et le noir.

L'avancée des connaissances sur le système visuel humain au XXème siècle nous montre que les deux théories sont en fait complémentaires et qu'elles interviennent à deux niveaux différents dans la perception de la couleur. Il est remarquable que les théoriciens de la couleur avaient anticipé au XIXème siècle les mécanismes rétiniens sous-jacents de la vision des couleurs. On sait maintenant que la couleur est mesurée par trois types de récepteurs au niveau de la rétine. Ces trois composantes chromatiques sont ensuite codées en oppositions de couleurs par le système visuel. La façon dont les trois composantes sont codées, c'est-à-dire la façon dont elles sont additionnées ou soustraites pour donner naissance à un canal achromatique et deux canaux d'oppositions de couleurs, a donné lieu a plusieurs modèles. Ces modèles sont rappelés p. 41. Nous reviendrons sur les modèles d'oppositions de couleurs au regard du modèle d'échantillonnage spatio-chromatique développé dans cette thèse.

La question concernant l'origine de la sensation jaune - c'est-à-dire si elle provient d'un mélange des excitations issues des récepteurs associés au rouge et au vert ou si elle provient d'un récepteur spécifiquement associé à la couleur jaune - peut être facilement tranchée expérimentalement. Elle a été résolue par des expériences de rivalité binoculaire. En effet, si en stimulant du rouge dans un oeil et du vert dans l'autre, une sensation jaune est perçue, alors il est clair que le

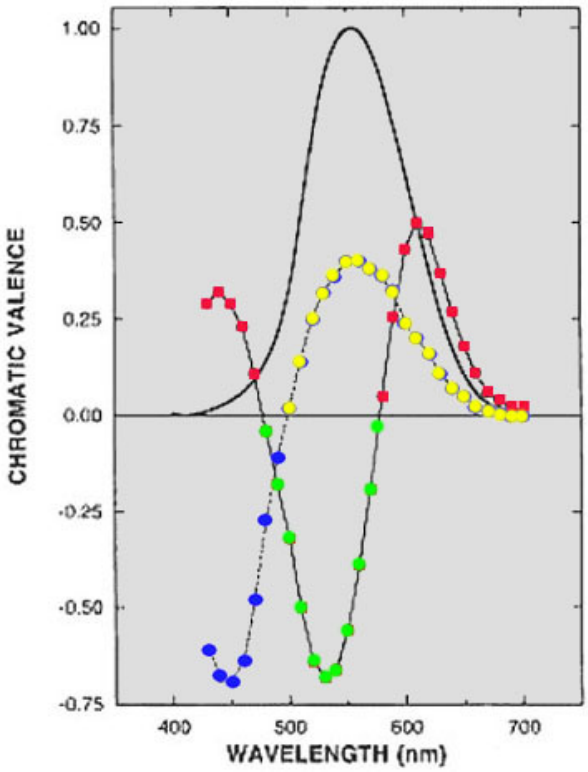


FIG. *2.5: Résultat de l'expérience de Jameson et Hurvich. Cette expérience permet de quantifier le codage en oppositions de couleurs.*

jaune résulte du mélange du rouge et du vert au niveau du cerveau, et non d'un récepteur jaune. C'est effectivement ce qui est observé [Hecht 28]. La fusion d'images binoculaires permet donc de corroborer la théorie de Young-Helmoltz.

Par ailleurs, plusieurs expériences psychophysiques permettent de mettre en évidence le phénomène d'oppositions de couleurs, en accord avec la théorie de Hering. L'expérience *d'annulation des teintes*, menée par Leo Hurvich et Dorothea Jameson en 1957, permit de mesurer précisément les quatre couleurs de base, appelées *unique hues* ou teintes uniques [Hurvich 57]. On parle de teinte unique lorsque celle-ci n'est pas perçue comme composée de plusieurs teintes. Par exemple la couleur orange est perçue comme un mélange de rouge et de jaune, ce n'est donc pas une teinte unique. La couleur jaune, au contraire, ne semble pas perçue comme un mélange de couleurs, c'est une teinte unique. L'expérience de Jameson et Hurvich est basée sur le fait que l'on peut annuler une des composantes chromatiques d'une couleur monochromatique en lui ajoutant la même quantité de la couleur opposée (Figure 2.4). On peut ainsi quantifier, à chaque longueur d'onde, les proportion des quatre couleurs bleu, jaune, rouge et vert. Le résultat de cette expérience est donné Figure 2.5. On constate que le bleu-jaune d'une part, et le rouge-vert d'autre part, sont en opposition. A une longueur d'onde donnée, on ne peut pas avoir une couleur composée à la fois de rouge et de vert, ou de bleu et de jaune.

1.2 Notions d'intensité et de chrominance

Le système visuel code l'information selon un axe de intensité et deux axes d'oppositions de couleur. Examinons dans ce paragraphe ces deux concepts en les illustrant sur une image numérique couleur.

Du fait de la trivalence de notre perception des couleurs et de la possibilité d'égaliser des couleurs, nous pouvons reproduire la couleur à l'aide de trois primaires. Pour la synthèse additive[1], les couleurs de base sont le rouge, le vert et le bleu. Ce sont les couleurs qui composent par exemple les phosphores d'un écran de télévision. Ainsi, une image en couleur est en fait composée de trois plans couleur C_R, C_V et C_B (Figure 2.6).

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

FIG. *2.6: Une image numérique couleur (a) est composée de trois plans couleur rouge (b), vert (c) et bleu (d).*

Il est possible d'effectuer un changement de base dans cet espace des couleurs pour décrire notre image en une composante d'intensité lumineuse φ, et deux composantes de chrominance ψ_R et ψ_B (ou oppositions de couleurs), portant l'information chromatique. L'intensité φ est définie comme un barycentre des trois canaux couleur :

$$\varphi = \sum_{i \in \{R,V,B\}} p_i C_i$$

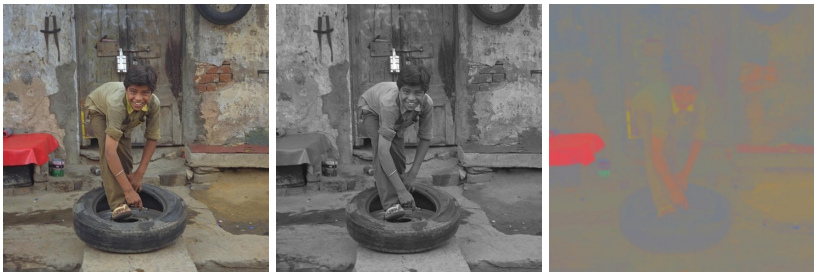
avec p_i des coefficients positifs dont la somme vaut 1. Le signal de chominance est alors définie comme la différence entre les plans couleur et l'intensité :

$$\begin{cases} \psi_R = C_R - \varphi = (1 - p_R)C_R - (p_V C_V + p_B C_B) \\ \psi_B = C_B - \varphi = (1 - p_B)C_B - (p_R C_R + p_V C_V) \end{cases}$$

On remarque que la chrominance est effectivement un signal d'opposition de couleur. On peut également calculer le signal d'opposition ψ_V, comme différence entre C_V et φ, cependant comme les trois composantes de chrominance sont linéairement dépendante, la connaissance de deux d'entre elles est suffisante. La Figure 2.7 montre les composantes d'intensité et de chrominance de l'exemple de la Figure 2.6.

Remarque : si on veut formaliser les notions d'intensité-chrominance pour une image couleur, on peut concevoir l'ensemble des images couleurs comme un espace affine de dimension tridimensionnel $\varepsilon = (E, V, f)$. E serait l'*ensemble* des images couleur (dont les intensités sont

[1] Il sera toujours question de synthèse additive, dans ce rapport.



(a) (b) (c)

FIG. *2.7: Décomposition d'une image couleur (a) en intensité (b) et chrominance (c). Une valeur constante est ajoutée pour l'affichage de la chrominance.*

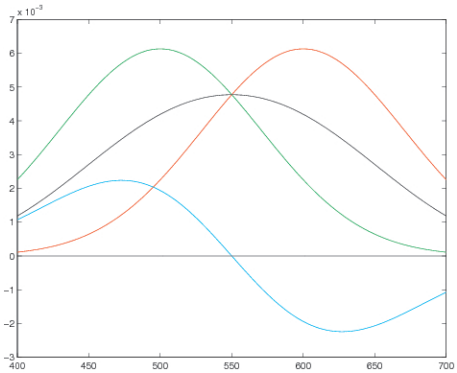


FIG. *2.8: Exemple de fonctions de sensibilités du rouge (en rouge), du vert (en vert), de l'intensité (en noir) et de la chrominance (en cyan).*

positives où nulles). V serait l'*espace vectoriel* des images couleurs dont les intensités peuvent être négatives. f serait l'application qui a deux points de E associe un vecteur dans V composé de la différence coordonnée par coordonnée. Une image couleur C et son intensité φ sont alors deux *points* de l'ensemble, tandis que la chrominance $\overrightarrow{\psi}$ est un *vecteur* de V :

$$f(C, \varphi) = \overrightarrow{\psi}$$

Autrement dit :

$$C = \varphi + \overrightarrow{\psi}$$

L'intensité est un point particulier de l'ensemble des images couleurs, puisqu'il est obligatoirement situé sur le segment de droite défini par les coefficients p_i. Le sous-espace affine associé à l'intensité est de dimension 0, puisque son espace vectoriel est réduit au vecteur nul ($\overrightarrow{\psi} = \overrightarrow{0}$). Par conséquent, on peut considérer que l'intensité est un scalaire, par opposition à la chrominance qui est un vecteur.

Que le système d'acquisition soit naturel ou qu'il soit artificiel, les valeurs $C_i(x, y)$ (avec $i \in \{R, V, B\}$ et (x, y) la position spatiale) sont mesurées en intégrant le produit entre l'irradiance de la scène $S(x, y, \lambda)$ et les fonctions de sensibilité $\sigma_i(\lambda)$, par rapport à la longueur d'onde λ :

$$C_i(x, y) = \int S(x, y, \lambda)\sigma_i(\lambda)d\lambda$$

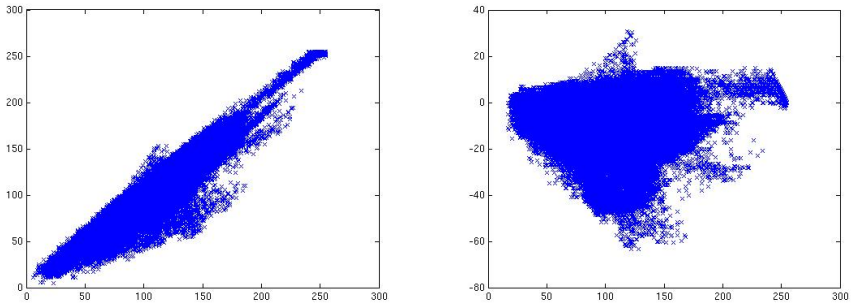


FIG. *2.9: Nuages de points de l'image "Pneu" dans le domaine (B,V) (à gauche) et dans le domaine (ψ_B =B-Lum,Lum) (à droite). Les plans bleu et vert présentent une forte composante principale dans la direction B=V, ils sont donc très corrélés. Au contraire, les deux composantes principales du nuage de points représentant la chrominance bleue en fonction de l'intensité sont $\psi_B = 0$ et Lum=128. Ces deux plans sont donc bien plus décorrélés.*

Le passage dans la base[2] intensité-chrominance comme l'acquisition des mesures avec des fonctions de sensibilité virtuelles σ_φ et σ_{psi_i}, calculées comme des combinaisons linéaires des σ_i :

$$\begin{cases} \varphi(x,y) &= \displaystyle\int_\lambda S(x,y,\lambda) \sum_{i\in\{R,V,B\}} p_i\sigma_i(\lambda)d\lambda \\ \psi_i(x,y) &= \displaystyle\int_\lambda S(x,y,\lambda)(\sigma_i(\lambda) - \sum_{j\in\{R,V,B\}} p_j\sigma_j(\lambda))d\lambda \end{cases}$$

S'il est possible d'avoir une fonction de sensibilité σ_φ, les fonctions σ_{ψ_i} sont en revanche une vue de l'esprit. Il est en effet impossible pour un capteur de mesurer des valeurs d'intensité lumineuse négatives, telle que ce serait le cas avec σ_{ψ_i}. En revanche, leur allure spectrale permet de bien comprendre ce que représente la chrominance. Nous avons représenté sur la Figure 2.8 deux exemples de fonctions de sensibilité σ_R et σ_V et les fonctions σ_φ et σ_{ψ_V} qui leur sont associées, avec $p_R = p_V = 1/2$. Comme on peut le voir, σ_φ est une fonction large à valeur positive, qui va permettre d'effectuer une moyenne spectrale de S. On peut la qualifier d'*intégrateur spectral*. Elle va nous renseigner sur la quantité d'énergie lumineuse qui a été mesurée. La fonction σ_{ψ_V} est quant à elle positive dans la région spectrale correspondant au pic de sensibilité du vert, et négative dans la région spectrale correspondant au pic de sensibilité du rouge. On peut la qualifier de *différentiateur spectral*. Elle nous renseigne sur la répartition spectrale de S. Une chrominance positive indique que l'énergie est majoritairement dans les longeurs d'onde du vert. De même, une chrominance négative indique que l'énergie est majoritairement dans le rouge.

L'intérêt de la base intensité-chrominance par rapport à la base RVB, est la décorrélation des signaux. Du fait que les fonctions de sensibilités spectrales σ_i se recouvrent, les signaux C_i sont très corrélés. Les C_i contiennent beaucoup d'information redondante, par exemple une ombre sera visible sur les trois canaux. L'intensité et la chrominance sont davantage décorrélés, puisque l'intensité portera une information spatiale (commune aux trois canaux C_i), tandis que la chrominance contient l'information spécifique à chaque canal couleur. Ainsi, une ombre,

[2]On parle souvent d'*espace* intensité-chrominance, par abus de langage.

qui est une information d'intensité lumineuse, sera portée uniquement par l'intensité. Cette représentation est donc plus efficace. Si l'on considère les fonctions de sensibilité, on constatera que le produit scalaire (l'intégral du produit) de σ_φ et de σ_{ψ_i} sera généralement proche de zéro, ce qui en fait une base orthogonale. Afin d'illustrer cette décorrélation, la Figure 2.9 montre deux nuages de points, correspondant à l'image Figure 2.6(a). L'un montre la position des pixels dans le plan défini par (C_B, C_V). Ce nuage a une composante principale dans la direction $C_R = C_B$, traduisant la forte corrélation des deux canaux. L'autre montre les pixels dans le plan défini par (ψ_B, φ). Ce nuage est davantage décorrélé. Par exemple pour une $\psi_B = 0$, la variance sur φ est maximale. Cette proriété de décorrélation est fondamentale dans la problématique de l'échantillonnage spatiochromatique, comme le montre le modèle proposé par Alleysson *et al.*, décrit au Chapitre 3, Paragraphe 2.2.3.

2 Anatomie du système visuel humain de bas niveau

Par bas niveau, nous entendons la rétine, le corps genouillé latéral et le cortex visuel primaire (Figure 2.10). Cet état de l'art met l'accent sur les propriétés spatio-chromatiques des neurones du système visuel.

Il est important d'insister sur le fait que les mesures électrophysiologiques effectuées sur les cellules du systèmes visuels sont très "aléatoires". Il y a une grande différence entre les modèles psychophysiques et la réalité de la physiologie qui est bien plus désordonnée. Ainsi, certains points font toujours l'objet de controverses au sein de la communauté. En particulier, la sélectivité des connexions des neurones -*une cellule donnée reçoit-elle une information uniquement des cônes L ou bien reçoit-elle une information de plusieurs types de cônes ?*- reste un débat ouvert [Lee 07], à l'exception peut-être des connexions des cellules horizontales [Dacey 96b]. Il s'agit donc d'avoir un esprit critique vis-à-vis des mesures rapportées dans le domaine, et de ne rien prendre comme acquis.

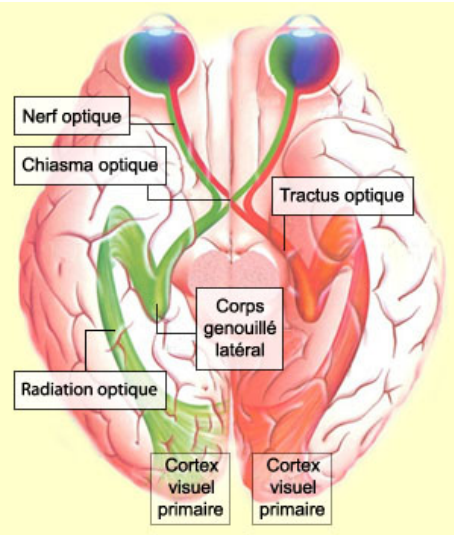


FIG. *2.10: Schéma global du système visuel humain [MacGill 02].*

2.1 Optique de l'oeil

Avant de se projeter sur la rétine, une image doit traverser le système optique de l'oeil (Figure 2.11), constitué de la la cornée, du cristallin, de la pupille et du corps vitré . La fonction de transfert optique de l'oeil dépend de la longueur d'onde. Elle est en particulier modifiée pour les courtes longueurs d'onde, dans le bleu. Il y aurait ainsi de l'aberration chromatique dans les longeurs d'onde du bleu [Cottaris 03].

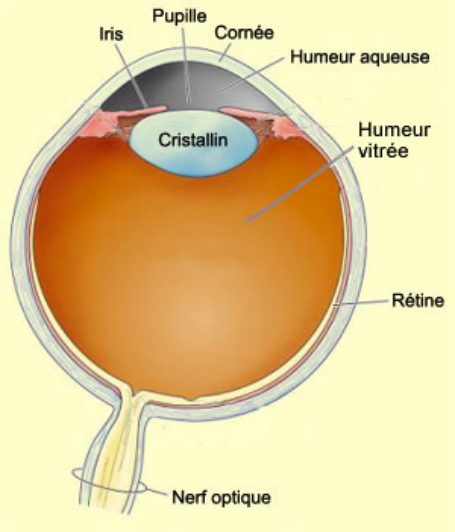


FIG. *2.11: Schéma d'une coupe de l'oeil [MacGill 02]*

Mais, même si l'optique de l'oeil était parfaite, l'image projetée sur la rétine aurait une résolution limitée, due à la diffraction occasionnée par la pupille. En effet, l'image projetée est convoluée par le module carré de la transformée de Fourier de l'ouverture de la pupille (diffraction à l'infini de Fraunhofer). Si celle-ci est circulaire, alors l'image d'un point brillant (autrement dit la réponse impulsionnelle du système) est la tache d'Airy, un sinus cardinal de révolution au carré. Ainsi la réponse impulsionnelle du système optique de l'oeil sera plus fine si le diamètre de la pupille est grand, et plus large pour un petit diamètre. La meilleure résolution est donc obtenue pour de grands diamètres de pupille. Cependant les aberrations optiques de l'oeil sont également plus importantes pour de grands diamètres de pupille. Il existe donc un optimum pour les diamètres de pupille intermédiaires [Thibos 00].

La fonction de transfert optique de l'oeil, c'est-à-dire la manière dont chaque amplitude est atténuée par le système optique de l'oeil avant d'arriver sur la rétine, dépend de l'excentricité. La fréquence de coupure est la plus haute dans l'axe de la fovea, région centrale de la rétine, où elle se situe autour de 60 cycles/degré [Campbell 66]. Elle diminue avec l'excentricité, de manière plus ou moins rapide selon les études, elle serait toujours de 50 cyc./deg. à 35 degrés d'excentricité [Anderson 91].

2.2 Le corps genouillé latéral : des voies fonctionnelles parallèles

Dans notre description anatomique du système visuel, nous commençons par le corps genouillé latéral (CGL), qui est un relais entre la rétine et le cortex visuel, situé dans le thalamus.

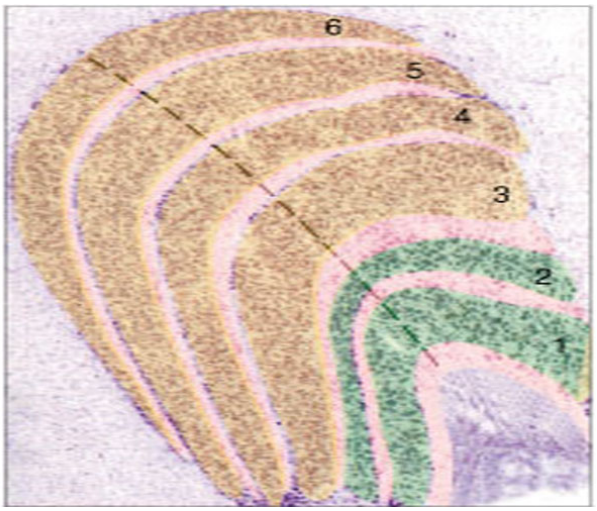


FIG. *2.12: Le corps genouillé latéral est structuré en plusieurs couches. Ces couches constituent trois voies cellulaires parallèles.*

La logique voudrait que l'on décrive le système visuel en suivant le parcours du signal lumineux au travers des neurones. Cependant, il existe au niveau du CGL des voies fonctionnelles très importantes, construites par les neurones rétiniens. Une description ici de ces voies cellulaires nous permettra d'y faire référence au cours de la description de des cellules rétiniennes.

Le CGL se décompose en six couches cellulaires (Figure 2.12). Le CGL doit son nom à la façon dont sont pliées ces couches qui rappelle la forme d'un genou. Les couches 1 et 2 (dont les neurones sont plus gros que ceux des couches 3 à 6) constituent la voie magnocellulaire. Les réponses des neurones de cette voie sont transitoires (elles varient rapidement dans le domaine temporel) et véhiculent une information de basses fréquences spatiales et achromatiques. Les couches 3 à 6, dont les neurones sont de taille plus petite, constituent la voie parvocellulaire. Elle véhiculerait les détails achromatiques et l'opposition de couleurs rouge-vert. Elle serait tonique (maintenue) et moins rapide que la voie magnocellulaire. Entre chacune de ces couches, des petits neurones constituent la voie koniocellulaire. Cette voie serait spécialisée dans le traitement de l'opposition bleu-jaune.

Il existe une spécificité remarquable des cellules du corps genouillé qui laisse penser que le CGL ne serait qu'un relais entre entre la rétine et le cortex visuel. Pour chaque cellule du CGL, une cellule afférente est extrêmement dominante [Alonso 06], et ainsi les champs récepteurs du CGL sont largement indistinguables des champs récepteurs des cellules ganglionnaires. Cependant, il y a une importante rétroaction de la part du cortex visuel sur le corps genouillé latéral, ce qui indique que des traitements sont accomplis dans du CGL [Ichida 02]. Ce mécanisme de rétroaction reste mal compris.

2.3 La rétine

L'identification et la classification des neurones rétiniens a commencé vers la fin du XIXème siècle avec les travaux de Santiago Ramón y Cajal, qui a appliqué la méthode de Camillo Golgi qui permettait la coloration de cellules. Bien que la majeure partie des différentes cellules aient sans doute été identifée aujourd'hui, on s'attend à en découvrir encore d'autres [Masland 01].

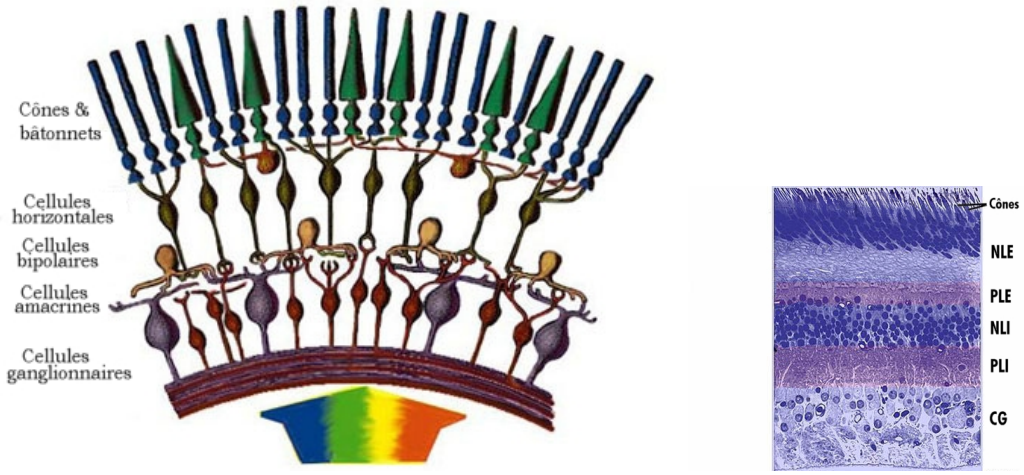


FIG. *2.13: Schéma et coupe de la rétine. La rétine est organisée en 5 couches de celllules. Attention, le sens de la lumière incidente est de bas en haut.*

Notre connaissance de la rétine humaine provient tout d'abord de l'étude du chat, puis, par la suite, de l'étude du macaque. La circuiterie du macaque est en effet très proche de celle de l'humain, et le macaque présente les mêmes performances visuelles [Dacey 00a].

La rétine est composée de 5 classes de neurones (Figure 2.13). Tout d'abord les photorécepteurs recueillent l'information lumineuse et la transcrivent en information électrique et/ou chimique. Cette information est relayée par les cellules bipolaires aux cellules ganglionnaires qui constituent la sortie de la rétine. La liaison photorécepteurs-bipolaires constitue la couche plexiforme externe, tandis que la liaison bipolaire-ganglionnaire constitue la couche plexiforme interne.

A ces neurones, qui véhiculent l'information verticalement d'une couche de neurone à l'autre (si on considère la rétine posée à plat), il faut ajouter deux classes de neurones effectuant des inhibition latérales : les cellules horizontales, se situant au niveau de l'intéraction photorécepteur/bipolaires, et les cellules amacrines, se situant au niveau de l'intéraction bipolaires/ganglionnaires. Pour chaque classe de cellule rétinienne, il existe une grande variété de sous-classes, qui varient par leurs caractéristiques (forme, champ récepteur, dentrites...). La figure 2.14 montre différents exemples de ces neurones. La classe la plus diverse est la classe des amacrines.

La rétine effectuerait un codage en oppositions de couleurs. Les cellules ganglionnaires qui véhiculerait ces oppositions de couleurs se répartissent en deux classes. La première porterait une opposition bleu-jaune par *différence* entre le signal issu des cônes S et un mélange de ceux issus des cônes L et M. La seconde véhiculerait, selon une idée largement répendue au sein de la communauté de vision des couleurs, une opposition rouge-vert des signaux issu des cônes L et M. D'après Dacey, les problèmes majeurs restants à résoudre seraient le nombre de voies neuronales qui transmettent des oppositions de couleurs, et le mécanisme sous-jacent permettant de créer les oppositions observées au niveau des cellules ganglionnaires [Dacey 00a].

Nous passons en revue ici les différents neurones rétiniens, leur répartition, leur proportion, leur comportement en terme de champ récepteur. Mais commençons tout d'abord par définir ce qu'est un neurone, puisque c'est l'élément de base de l'architecture de la rétine.

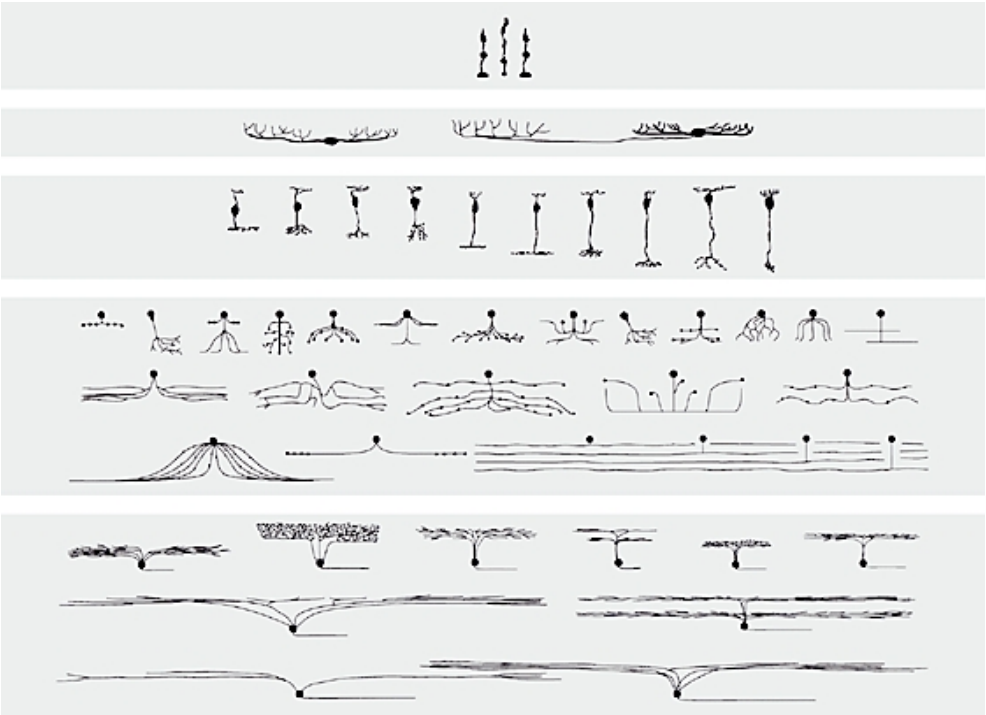


FIG. *2.14: Schéma des différentes classes de cellules présentes dans la rétine [Masland 01]. De haut en bas : les photorécepteurs, les horizontales, les bipolaires, les amacrines et les ganglionnaires. Les sous-classes ne sont pas toutes représentées.*

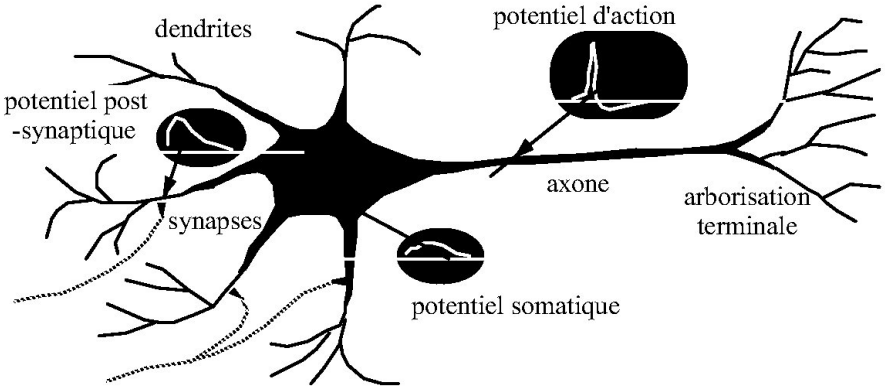


FIG. *2.15: Schéma d'un neurone.*

2.3.1 Le neurone : structure élémentaire

Un neurone est une structure nerveuse élémentaire composée d'un corps cellulaire (ou soma), d'un ensemble de ramifications courtes et nombreuses, les dendrites, et d'une fibre unique, l'axone. Le neurone reçoit des influx nerveux à travers les dendrites et les achemine à l'aide de son axone. Les neurones sont connectés entre eux à travers des synapses qui transmettent une information par l'intermédiaire d'un médiateur chimique ou directement sous forme électrique entre deux neurones. Chaque synapse est dite excitatrice ou inhibitrice, suivant que son influence est d'augmenter ou de diminuer l'état d'excitation du neurone auquel elle est reliée.

Le neurone ne peut pas transmettre un signal relatif qui aurait des valeurs négatives car ses

synapses sont soit excitatrices soit inhibitrices (une concentration chimique ne peut pas être négative). Ainsi certains neurones seront dits ON, ils transmettent la partie positive du signal, et d'autres seront dits OFF, ils transmettent la partie négative du signal.

Le système nerveux n'est pas constitué uniquement de neurones, mais également de cellules dites gliales. Ces cellules peuvent avoir différents rôles, comme par exemple métabolique, ou encore d'isolation. On trouve trois types de cellules gliales dans la rétine, dont les cellules de Müller. Ces dernières ont une forme allongée et s'étendent depuis le corps vitreux de l'oeil jusqu'au pied des photorécepteurs, traversant ainsi toute l'épaisseur de la rétine. Il a été récemment découvert que ces cellules joueraient un rôle analogue à celui de fibres optiques artificielles [Franze 07], c'est-à-dire qu'elles transmettraient le flux lumineux d'un point de l'espace à un autre sans dégradation. Ceci permettrait de diminuer la dispersion de la lumière due à la présence des neurones rétiniens, puisque la lumière incidente doit traverser toute l'épaisseur de la rétine avant d'arriver aux photorécepteurs, les neurones sensibles à la lumière.

Nous utiliserons abondamment, dans la suite de ce travail, la notion de champ récepteur associé à un neurone. Le champ récepteur d'un neurone est la portion de la rétine qui influence (par excitation ou inhibition) l'activité du neurone lorsqu'elle est soumise à un stimulus visuel. Bien sûr l'intensité de la réponse du neurone va dépendre de la position du stimulus à l'intérieur du champ récepteur. Cette réponse ne sera pas uniforme. Nous nous intéresserons donc tout particulièrement à la "forme" du champ récepteur, c'est-à-dire l'intensité avec laquelle le neurone réagit en fonction de la position du stimilus dans le champ récepteur. On peut faire une analogie entre le concept de champ récepteur d'un neurone et celui de réponse impulsionnelle d'un système dans le domaine du traitement du signal. La plupart des neurones visuels présentent un champ récepteur consitué d'un antagonisme centre/pourtour. Le centre sera excitateur pour les cellules ON (resp. inhibiteur pour les cellules OFF) tandis que le pourtour du champ récepteur sera inhibiteur (resp. excitateur). La forme du champ récepteur peut également dépendre de la longueur d'onde du stimulus. Le neurone peut être excité par certaines longueurs d'onde, et inhibées par d'autres.

Les formes des champs récepteur est importante pour la compréhension du système visuel, car elles correspondent à différents filtrages effectués par la rétine. L'électro-physiologie permet d'enregistrer directement à l'aide d'électrodes la réponse des neurones.

2.3.2 Les photorécepteurs

Il existe deux types de photorécepteurs : les bâtonnets et les cônes. Les bâtonnets sont responsables de la vision nocturne. Ils sont très sensibles aux faibles intensités lumineuses, mais saturent lorsque l'intensité lumineuse est importante (en plei soleil, par exemple). Possédant tous la même sensibilité spectrale, les bâtonnets ne sont pas sensibles à la couleur. C'est pourquoi nous ne percevons pas la couleur en vision nocture. Il y a environ 100 millions de bâtonnets, et 4 à 6 millions de cônes sur une rétine [Curcio 90], ces proportions variant d'un individu à l'autre. Bien que les bâtonnets soient en nombre bien supérieur, ce ne sont pas les photorécepteurs principaux. Ils seraient apparus après les cônes, au cours de l'évolution [Okano 92].

Les cônes sont quant à eux responsables de la vision diurne. Ils contiennent une opsine (une molécule de protéïne) qui détermine leur sensibilité spectrale. Au cours de l'évolution,

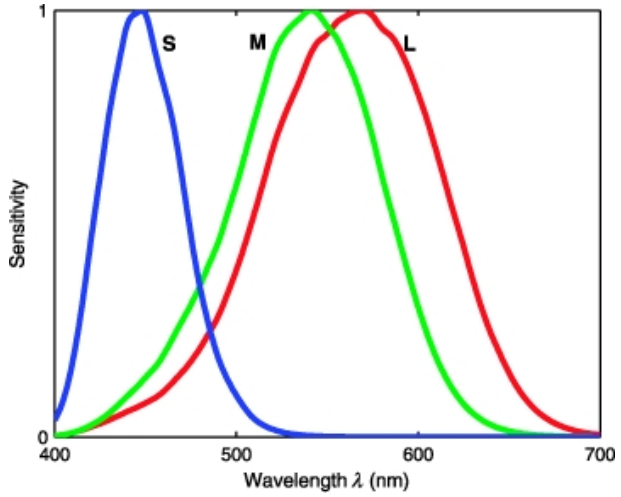


FIG. *2.16: Sensibilités spectrales des cônes L, M et S [Stockman 00].*

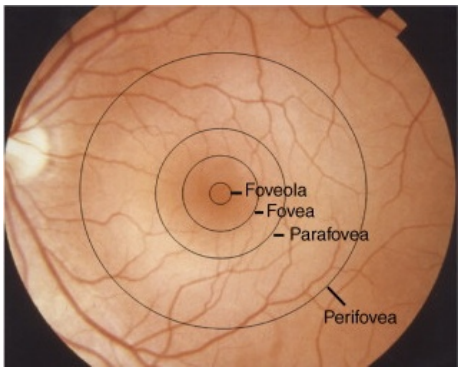


FIG. *2.17: La macula est constituée de trois régions concentriques : la fovea, la parafovea et la périfovea. Image publique du "National Eye Institute".*

l'unique opsine présente initialement s'est scindée en deux opsines, l'une sensible aux courtes longueurs d'onde (dans le bleu) et l'autre sensible aux grandes longueurs d'onde (dans le rouge/vert) [Okano 92]. Ces cônes bleus, ou "S" pour *Short wavelengths* (courtes longueurs d'onde), représentent moins de 10% de l'ensemble de la population des cônes et sont morphologiquement différents de l'autre type de cônes qui s'est différencié plus tard en cônes L, pour *Long wavelengths* (grandes longueurs d'onde), sensibles dans les gammes du rouge, et en cônes M, pour *Middle wavelengths* (moyennes longueurs d'onde), sensibles dans les gammes du vert [Calkins 98]. Les cônes S, M et L présentent des pics de sensibilité respectivement à 430nm, 530nm et 560nm [Solomon 07], mais leurs fonctions de sensibilité, en particulier celles des cônes L et M, couvrent presque l'ensemble du spectre visible (Figure 2.16). Il est important de noter que la fonction de sensibilité des cônes S est très décorrélée de celles des cônes L et M, tandis que ces deux dernières sont au contraire extrêmement corrélées. Par ailleurs, les cônes S ne contribueraient pas, ou très peu à la fonction d'intensité $V(\lambda)$. Cette dernière peut-être modélisée par la combinaison des fonctions de sensibilités des cônes L et M [Lennie 93].

Le centre de la rétine, appelé la macula (Figure 2.17), comprend trois régions : la fovea, la parafovea et la périfovea [Marmorstein 07]. La fovea est au centre de la macula et possède un diamètre d'environ 1,5mm. A l'intérieur de la fovea, une région de 0,5-0,8 mm de diamètre, ne

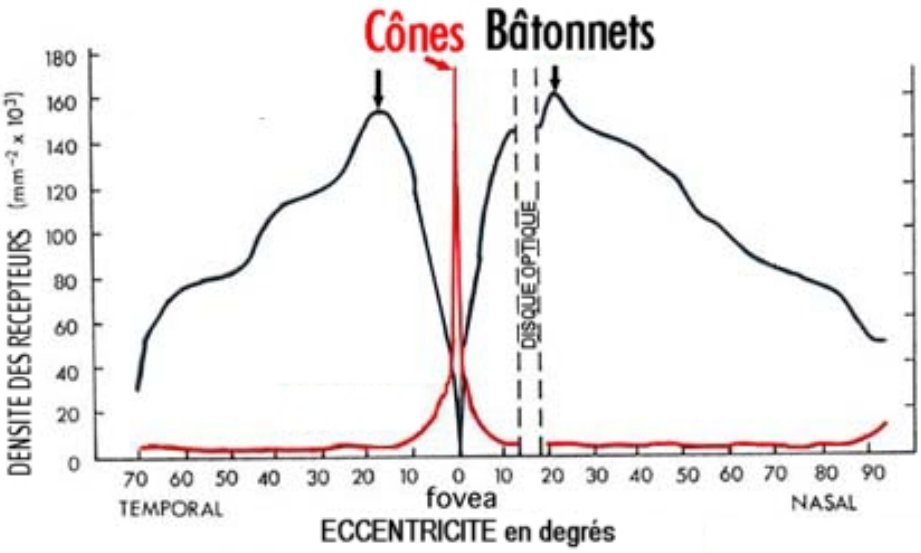


FIG. *2.18: Répartition des bâtonnets et de cônes en fonction de l'excentricité de la rétine*

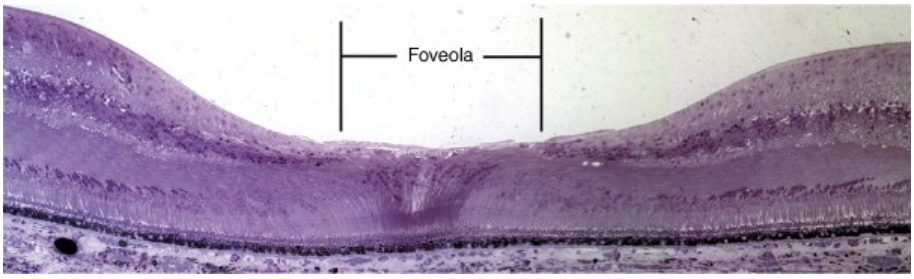


FIG. *2.19: Coupe de la retine au niveau de la foveola. La rétine présente une excavation permettant à la lumière incidente d'atteindre directement les cônes, dans cette région où l'acuité visuelle est maximale. Image tirée de [Marmorstein 07].*

contient aucun vaisseau capillaire. A l'intérieur de cette région se trouve la foveola, une région de 0,35mm de diamètre qui contient uniquement des cônes (Figure, la rétine présentant une excavation à cet endroit 2.19). La parafovea est une couronne de 0,5 mm autour de la fovea. Enfin, la périfovea est la couronne de 1,5mm autour de la parafovea.

La densité maximale des cônes en fovea est très variable selon les individus, de 100 000 à 300 000 cônes par millimètre carré [Curcio 90]. La densité des cônes diminue très rapidement avec l'excentricité (Figure 2.18). A 1mm du point de densité maximale, la densité perd un point de magnitude [Curcio 90]. De plus, elle n'est pas isotrope, elle est supérieure du côté nasal que du côté temporal de 40-45% et est légèrement supérieur en périphérie inférieure qu'en périphérie supérieure. Les deux yeux d'un même individu ont le même nombre de photorécepteurs, et ont une topographie similaire mais non identique [Curcio 90]. En fovea, les cônes sont disposés sur une grille hexagonale compacte quasi-régulière [Yellott 90]. La régularité locale est à peu près à l'échelle du 10ème de degré. La fréquence de Nyquist calculée par rapport à la densité des cônes est de l'ordre de 60 cycles/degré. Cette fréquence de Nyquist diminue brutalement avec l'excentricité (elle suit la densité des cônes), et se stabilise à 20 cycles/degré à partir de 10 degrés d'excentricité Les cônes sont disposés de manière de plus en plus irrégulière en fonction de l'excentricité. La nature régulière ou irrégulière est peut-être simplement due au nombre de

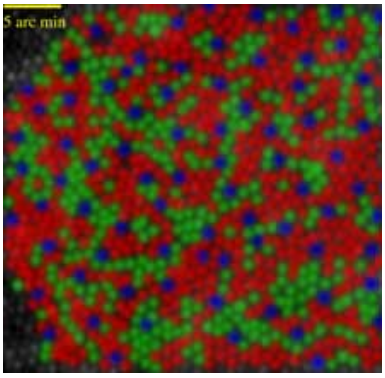


FIG. *2.20: Exemple d'image rétinienne de macaque [Roorda 01]. Les cônes L, M et S sont représentés par les fausses couleurs rouge, verte et bleue. Les cônes L et M sont arrangés très irrégulièrement. La mosaïque des cônes S est irrégulière mais clairsemée.*

degrés de liberté autorisé par la densité. Par exemple, si vous tenez un ensemble de crayons dans votre main et que vous les serrez, alors leur arrangement sera de forme hexagonale compacte, puisque la distance entre deux crayons étant réduite à zéro, ils n'ont aucun degré de liberté. En revanche, si vous ouvrez votre main et que vous vous trouvez en état d'apesanteur, alors au bout d'un certain lapse de temps, les crayons auront occupés un plus gros volume d'espace et la distance entre deux crayons sera aléatoire, donc irrégulière.

Des techniques sophistiquées permettent de visualiser les différentes classes de cônes [Packer 96, Roorda 01]. La Figure 2.20 montre une images rétiniennes obtenue sur un macaque. Roorda *et al.* utilisent par exemple un système d'optique adaptative pour pouvoir réaliser des images rétiniennes, à l'image de la technique utilisée en astronomie. L'idée consiste à corriger en temps réel les déformations du front d'onde de la lumière incidente. Les auteurs rapportent que les cônes S sont distribués de manière irrégulière chez les sujets humains, mais pas chez les macaques. D'autres auteurs affirment que la mosaïque des cônes S serait quasi-régulière [Curcio 91]. Il semblerait en fait que les cônes S soient disposés aléatoirement, mais avec une zone de garde entre deux cônes voisins, permettant ainsi d'avoir une densité constante à faible échelle.

Les cônes S sont totalement absents au centre de la foveola et ont un maximum de densité sur une couronne de 0.1 à 0.3 millimètre du centre de la fovea. De plus, d'après [Curcio 91], les distances inter-cônes les plus courtes sont entre les cônes S et les cônes L/M. Les mosaïques des cônes S et des cônes L/M seraient également statistiquement indépendantes. Cela suggère que la mosaïques S et les mosaïques L/M aient deux processus de formation différents, ce qui est en concordance avec les différentes étapes de l'évolution, puisque la différentiation entre L et M est survenue plus tardivement que la différentiation en deux types de cônes de plus longue longueur d'onde et de plus courte longueur d'onde.

D'après Roorda *et al* [Roorda 01], les cônes L et M sont arrangés de manière aléatoire, créant ainsi des petites régions, que nous qualifierons d'amas ("*clusters*" en anglais), où un seul type de photorécepteur est présent. Hofer *et al.* [Hofer 05] confirment ces petits regroupements d'un seul type de cône L ou M. L'optique adaptative permet de produire des stimuli très compacts dans le domaine spatial, de l'ordre de la taille de ces amas. La perception de la couleur de tels stimuli exposés de manière très brève varierait d'un flash à l'autre [Williams 07], probablement selon la proportion de chaque type de cônes atteints. Cette expérience montre la nécessité

de l'intégration spatiale et/ou temporelle pour une perception correcte de la couleur.

Il est également rapporté par Roorda *et al.* [Roorda 01] que les cônes L et M sont disposés de manière aléatoire et que leur proportion varie grandement d'un individu à l'autre. Hofer *et al.* [Hofer 05] rapportent des rapports L :M allant de 1,1 :1 à 16,5 :1. Cette variation expliquerait également les variabilités de la fonction d'intensité d'un individu à l'autre [Lennie 93]. Au contraire, les proportions de cônes S varient très peu d'un individu à l'autre [Hofer 05].

Les variations inter-individuelles des densités en photorécepteurs L et M ainsi que la variation à l'intérieur même de la rétine aurait pour conséquence de modifier le codage de l'intensité. Dacey [Dacey 00b] a mené des expériences sur la réponse des cellules horizontales sur des cellules provenant de différents individus et sur différentes régions à l'intérieur des rétines. Ces expériences révèlent une grande variabilité dans les réponses, reflétant la variabilité des proportions de cônes L et M. La même constatation aurait été faite au niveau des cellules ganglionnaires par Diller *et al* [Dacey 00a]. Ainsi la notion d'intensité *physiologique* - c.à.d. mesurée dans les cellules rétiniennes - dépendrait du nombre de cônes L et M dans le champ récepteur des cellules, et serait donc variable selon les individus ainsi que selon la région considérée à l'intérieur même d'une rétine.

Par ailleurs, du fait des variations inter-individuelles du ratio L/M, on pourrait s'attendre à ce que la longueur d'onde de la teinte jaune unique ("*unique yellow*") soit différente d'un individu à l'autre. Le jaune unique est la couleur pour laquelle le rouge et le vert sont équilibrés (ils s'annulent alors car ils sont en opposition). Le ratio L/M devrait donc influer directement sur la longeur d'onde de cette annulation rouge/vert. Pourtant il a été vérifié expérimentalement que ce n'est pas le cas [Brainard 00]. De plus, la position du jaune unique est influencée par l'environnement. En effet, des décalages significatifs ont été observés chez des sujets portant plusieurs jours d'affilés des lunettes pourvues de filtres rouges ou verts. Ainsi, la position du jaune unique serait déterminée par un mécanisme adaptatif valable tout au long de la vie, guidé par la statistique chromatique de l'environnement.

Il est intéressant de remarquer qu'il existe une zone de la rétine d'environ 1 millimètre de diamètre totalement dépourvue de photorécepteurs : c'est ce que l'on appelle la tache aveugle. Cette tache aveugle est due au fait que le nerf optique doivent traverser la rétine afin de se raccorder aux signaux de sortie de celle-ci. En effet les cellules photoréceptrices ne sont pas à la surface de la rétine mais sous plusieurs couches de neurones (Figure 2.13). Ainsi la lumière incidente qui excitent les cellules photosensibles (les cônes et les bâtonnets) traverse d'abord d'autres couches de neurones de la rétine, que nous décrivons dans la suite, et qui servent au traitement du signal mesuré par ces cellules photosensibles. Cette tache aveugle peut être facilement mise en évidence (Figure 2.21). Elle n'est pas présente dans toutes les espèces d'animaux. Par exemple, les pieuvres n'en ont pas [Gopalan 01] car leurs photorécepteurs sont, plus logiquement, directement à la surface de leur rétine.

FIG. *2.21: Mise en évidence de la tache aveugle. Fermez votre oeil droit et regarder le point rouge. Approchez doucement vers le point rouge. A une certaine distane la ligne doit apparaître continue.*

2.3.3 Les cellules bipolaires : liaison entre les photorécepteurs et les ganglionnaires

Les cellules bipolaires transmettent l'information issues des cônes aux cellules ganglionnaires. Il existerait entre 9 et 11 variétés de cellules bipolaires [Masland 01]. Chaque cône serait en contact avec plusieurs types de cellules bipolaires, quelle que soit la région sur la rétine. Le signal des cônes serait ainsi parallélisé sous différents formats, chaque format dépendant du type de bipolaire associé.

La couche plexiforme interne est composée en fait de différentes sous-couches synaptiques. Certaines cellules ganglionnaires ont des dentrites dans une première couche, que viennent contacter des cellules bipolaires, d'autre dans une deuxième couche, en relation avec une autre variété de bipolaire, et ainsi de suite [Masland 01]. Les cellules bipolaires se classent également en fonction du nombre de dendrites ainsi qu'en fonction de leurs récepteurs de neurotransmetteurs (distinction moléculaire).

Il existe des cellules bipolaires type ON et de type OFF, en même nombre, véhiculant respectivement la partie positive et la partie négative du signal issu de la couche plexiforme externe. Bien que le signal issu directement de la transduction de la lumière soit positif, il peut prendre des valeurs négatives du fait de la rétroaction des cellules horizontales. Ces deux classes, ON et OFF, se découpent encore en cellules à réponse rapide ("*ON/OFF-transient*"), véhiculant l'information haute fréquence temporelle, donc le mouvement, ainsi qu'en cellules à réponse plus lente ("*ON/OFF-sustained*") qui véhiculeraient les informations de basse fréquence temporelle [Awatramani 00, DeVries 00]. Awatramani *et al.* ont démontré que cette distinction entre *sustained* et *transient* est une caractéristique inhérente aux cellules bipolaires, c'est-à-dire qu'elle n'est pas causée par une rétroaction d'autres neurones comme on le supposait auparavant.

Les bipolaires dites diffuses, sont connectées à plusieurs cônes, entre 5 et 10. Elles participent à l'élaboration de la voie magnocellulaire. Les bipolaires dites *midget* (à cause de leur petite morphologie, *midget* signifiant miniature, en anglais) et intervenant dans la construction de la voie parvocellulaire, sont connectées quant à elles à un seul cône L ou M. Certaines bipolaires sont connectées spécifiquement à des cônes S (entre 1 et 5 cônes S) [Kouyama 92] et participent à l'élaboration de la voie koniocellulaire.

Les cônes bleus sont associés à une circuiterie spécifique. Ils se connectent à un seul type de bipolaires, qui à leur tour ne se connectent qu'à un seul type de ganglionnaires, les ganglionnaires bistratifiées [Kouyama 92, Dacey 94]. Ces cellules ganglionnaires spécifiques aux cônes S reçoivent également de l'information des bipolaires issues des cônes de plus grandes longueurs d'onde L et M, et effectuent une différence entre ces deux informations (S d'un côté et L/M indifférenciés de l'autre), c'est une opposition bleu-jaune. Ce principe d'opposition de couleurs entre les cônes S et les cônes de plus grande longueur d'onde semble être commune à tous les mammifères (qu'il y ait un ou deux types de cônes de grande longueur d'onde) [Jacobs 93]. Etonnamment, il ne semble pas y avoir de bipolaires OFF pour les cônes S, du fait du pédicule particulier des cônes S ne leur permettant pas de délivrer de signal OFF [Kolb 97].

Les bipolaires autres que celles spécifiques aux cônes S peuvent être connectées à la fois à des cônes L et des cônes M. Elles ne sont pas chromatiquement sélectives, contrairement aux bipolaires S [Masland 01].

Il existe en moyenne un plus grand nombre de bâtonnets que de cônes par cellule bipolaire. De nombreux bâtonnets convergent en effet vers une même cellule bipolaire. L'acuité spatiale nocturne est donc "sacrifiée" au bénéfice de la sensibilité. Bien qu'il existe quelques cellules bipolaires spécifiques aux bâtonnets, ces derniers sont connectées aux mêmes cellules bipolaires que les cônes, la synapse se faisant sur l'axone terminal des bipolaires des cônes. Le système associé aux bâtonnets semble avoir été "greffé" sur le système préexistant associé aux cônes [Masland 01]. Ainsi, il existe une intéraction entre cônes et bâtonnets en vision mésopique, c'est-à-dire lorsque l'intensité lumineuse est suffisamment faible pour que les bâtonnets puissent fonctionner, mais suffisamment forte pour que les cônes restent actifs. Les bâtonnets provoquent des décalage spectraux des teintes uniques (les "*unique hues*", p. 17), mais toutes les teintes uniques ne sont pas déviées dans le même sens (vers les grandes ou vers les faibles longeurs d'onde), ce qui rend leur modélisation problématique [Buck 00].

2.3.4 Les cellules ganglionnaires

Les cellules ganglionnaires constituent la sortie de la rétine, leur axone formant le nerf optique jusqu'au corps genouillé latéral. La rétine contient environ 1,5 millions de cellules ganglionnaires, réparties entre 10 à 15 classes [Masland 01].

De l'ordre de 80% de ces cellules sont des cellules dites *midget*, qui se projettent sur la voie parvocellulaire du corps genouillé latéral [Perry 84]. Les cellules *midget* présentent une structure de champ récepteur de type centre/pourtour avec un antagonisme de couleurs de type rouge-vert : certaines cellules sont excitées par le signal issu des cônes L et inhibées par les signal issu des cônes M, et *vice versa*. Elles sont donc supposées véhiculer l'information d'opposition rouge-vert. C'est du moins la thèse majoritaire au sein de la communauté scientifique, car il existe une controverse quant à la transmission neuronale des oppositions rouge/vert. La thèse minoritaire étant que des cellules ganglionnaires spécifiques à cette opposition n'ont pas encore été découvertes [Dacey 96a, Calkins 96]. Nous reviendrons sur cette controverse lors de la description des différents modèles neurocomputationnels de vision des couleurs.

Smith et al. [Smith 92] ont mesuré les réponses de ces cellules chez le macaque en fonction d'un stimulus hétérochromatique modulé. Un tel stimulus consiste en la modulation de deux lumières colorées, en l'occurence une rouge et une verte, dont on contrôle la fréquence et la phase relative. Le décalage de phase entre les deux modulations permet de doser le niveau de chromaticité du stimulus. Le stimulus est purement achromatique lorsque les deux signaux varient en phase, et il est purement chromatique lorsqu'ils varient en opposition. Leurs résultats montrent que les cellules *midget* ne sont pas sensibles aux faibles fréquences achromatiques, tandis qu'elles sont sensibles aux faibles fréquences chromatiques. Elles sont en revanche très sensibles aux hautes fréquences achromatiques, et modérément sensibles aux hautes fréquences de stimuli chromatiques.

Dans la fovea, une cellule ganglionnaire *midget* est associée à une seule cellule bipolaire *midget*, qui est elle-même associée à un seul cône [Masland 01]. On parle de "*private line*" en anglais, que nous traduirons par ligne directe. Cela donne le maximum de résolution spatial au niveau des ganglionnaires *midget*. En périphérie de la mosaïque, les ganglionnaires ont un champ dendritique bien plus large et sont, par conséquent, connectés à un plus grand

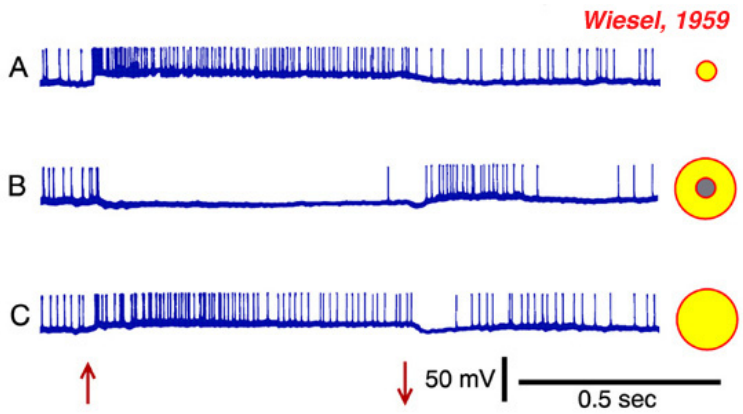


FIG. 2.22: *Activité électrique d'une cellule ganglionnaire de chat, pour trois type de stimulus [Wiesel 59]. (A) Le stimulus est un point correspondant au centre du champ récepteur du neurone. L'activité est maximale. (B) Le stimulus est une couronne. Le neurone est inhibé. (C) Le stimulus est un disque. L'activité est soutenue, mais moins forte qu'en (A).*

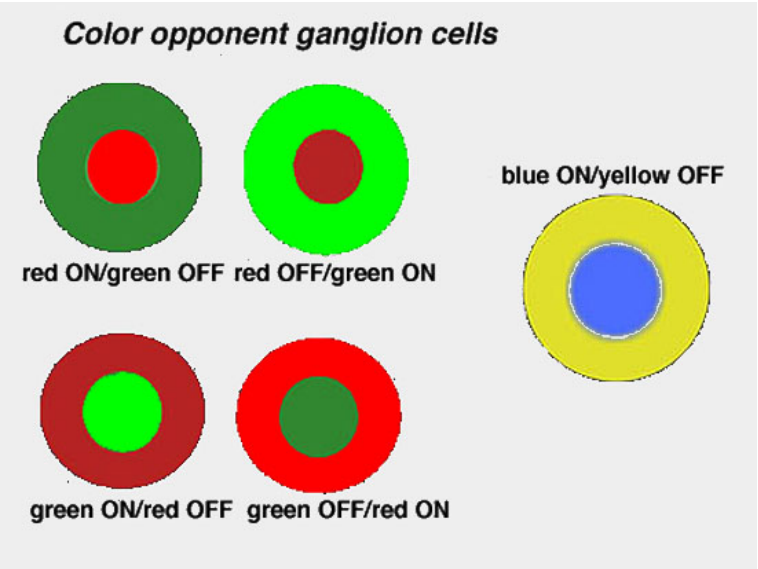


FIG. 2.23: *Champ récepteur des cellules ganglionnaires présentant un antagonisme spectral.*

nombre de bipolaires, jusqu'entre 30 et 40 bipolaires [Dacey 00a]. Cet acroissement du nombre de connexions par cellule ganglionnaire est valable aussi bien dans le centre de leur champ récepteur que dans le pourtour [Dacey 99].

Une question qui n'est pas clairement tranchée concerne la sélectivité des champs récepteurs ([Smith 92, Lankheet 98]). Le pourtour du champ récepteur des ganglionnaires *midget* en fovea pourrait être connecté aux bipolaires L et M sans discrimination. Dacey [Dacey 96a] soutient cette hypothèse ("*mixed surround*") car les cellules horizontales - qui constituent l'essentiel du pourtour du champ récepteur des ganglionnaires - ne présentent pas de sélectivité entre L et M. De plus, il mentionne le fait que le bien plus grand poids synaptique associé au centre permet d'avoir une opposition. La même constatation de non-sélectivité est faite au niveau des cellules amacrines [Calkins 96]. Dans la périphérie de la rétine, les cellules ganglionnaires *midget* reçoivent de l'information d'un bien plus grand nombre de cellules bipolaires. En particulier

le centre du champ récepteur n'est plus associé à une seule bipolaire, mais à plusieurs. Les expériences psychophysiques montrent que la sensibilité aux variations chromatiques décroît avec l'excentricité deux fois plus vite que la sensibilité aux variations en intensité [Mullen 91]. D'après Dacey [Dacey 99] et Mullen *et al.* [Mullen 96], la non-sélectivité des connexions expliquerait cette baisse de sensibilité, puisqu'il n'y aurait plus d'antagonisme chromatique, le centre et le pourtour portant tous deux un signal de type (L+M). Cependant, il aurait été mesuré à la sortie de ces cellules ganglionnaires une sensibilité chromatique aussi forte en périfovea qu'en fovea [Martin 01]. Le fait qu'en fovea, une cellule ganglionnaire est associé à un seul cône en son centre tandis qu'en périfovea une cellule ganglionnaire est associé à un plus important nombre de cônes, et ce à sensibilité chromatique égale, permettrait de conclure à la sélectivité des connexions, du moins dans le centre du champ récepteur, et peut-être aussi dans le surround. La diminution de sensibilité serait alors imputable à des causes post-rétiniennes. Ce problème de spécificité ou non du champ récepteur des cellules ganglionnaires est une question centrale dans la compréhension des mécanismes de la vision des couleurs chez les primates.

La deuxième classe principale de cellules ganglionnaires est la classe des ganglionnaires parasol. Elles représentent 10% des cellules ganglionnaires. Leur morphologie est différente des *midget*. Leur arbre dendritique est bien plus étendu, sommant ainsi une information sur une plus grande région. Ces cellules se projettent sur la couche magnocellulaire du corps genouillé latéral [Perry 84]. Elles ne présentent pas d'opposition de couleurs, mais ont au contraire une sensibilité spectrale étendue du fait qu'elles effectuent la somme des signaux des cônes L et M. Là encore, les signaux mis en jeu dans le pourtour des champs récepteurs ne sont pas clairements définis. Smith *et al.* [Smith 92], par exemple, émettent l'hypothèse que le pourtour pourrait être défini par une opposition de couleurs L-M, tandis que le centre serait de type L+M. Mais cela ne satisfait pas complètement les observations, comme c'est le cas également de l'hypothèse où le centre et le pourtour seraient défini par L+M. Le pourtour des cellules du champ récepteur magnocellulaire est donc complexe, et fait intervenir des oppositions de couleurs. Une description détaillée peut être trouvée dans l'article de Lee [Lee 96].

D'après Sun *et al.*, les cellules ganglionnaires *midget* ainsi que parasol n'auraient aucune connexion avec les bipolaires spécifiques aux cônes S [Sun 06].

Une troisième classe de ganglionnaires, découverte après les deux types de ganglionnaires principaux, est quant à elle spécifique aux cônes S. Elle permet de coder l'opposition de couleurs bleu-jaune [Dacey 94, Calkins 98, Dacey 00a]. Ces petites cellules représentent 3% de la population des cellules ganglionnaires. La découverte tardive de ces cellules motive peut-être l'hypothèse d'une cellule non encore découverte codant l'opposition rouge-vert.

Les autres classes de ganglionnaires sont en nombre bien inférieur et sont moins connues. Les taches qu'elles effectuent ne nécessitent donc pas leur présence en grand nombre. Il a été récemment découvert que certaines cellules ganglionnaires (entre 1% et 3%) sont intrinsèquement sensibles à la lumière. Elles pourraient avoir un rôle dans la régulation du rythme circadien (notre rythme biologique) ainsi que dans le contrôle réflexe de l'ouverture de la pupille [Berson 03, Gooley 03]).

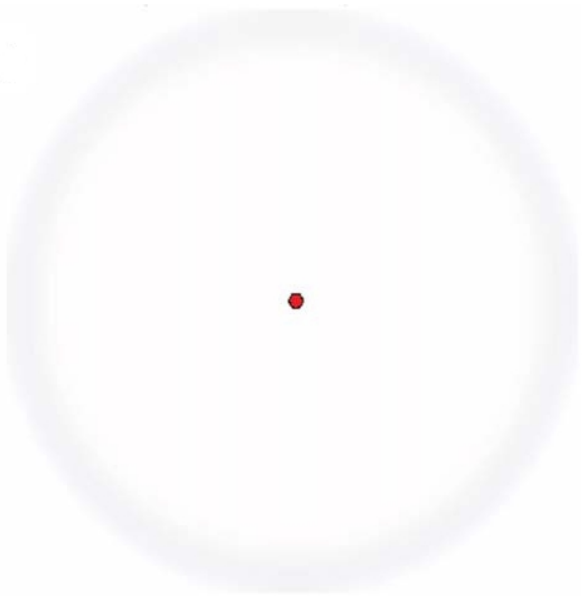


FIG. *2.24: Illusion visuelle illustrant le caractère passe-haut temporel du système visuel [Martinez-Conde 06]. Fixez le rond rouge et stabilisez votre regard. L'anneau présent en périphérie du champ visuel devrait disparaître.*

2.3.5 Les cellules horizontales et amacrines : inhibitions latérales

Les cellules horizontales agissent au niveau de la couche plexiforme interne et établissent des connexions latérales entres cônes, bâtonnets et bipolaires. Elles auraient un rôle de rétroaction sur les cônes permettant d'ajuster la dynamique des cônes et de supprimer des variations d'illumination [Meylan 07]. Cela permet en particulier de percevoir des détails dans une zone sombre d'une scène globalement très illuminée.

Elles seraient par ailleurs impliquées dans la construction de la partie périphérique du champ récepteur des cellules ganglionnaires [Mangel 87]. Les cellules horizontales ne présentent pas d'opposition spectrale [Dacey 96b]. Ceci constitue un argument en faveur de l'hypothèse de mixité du champ récepteur périphérique des cellules ganglionnaires.

Les cellules H1 ne seraient connectées qu'aux cônes L et M, et non aux S. Les cellules H2 seraient quant à elles connectées aux trois types de cônes, et répondraient très fortement aux stimulations provenant des cônes S [Dacey 96a]. Ainsi les cellules H1 seraient impliquées dans la formation des voies parvocellulaire et magnocellulaire, tandis que les cellules H2 correspondent à la circuiterie spécifique associée aux cônes S, la voie koniocellulaire.

On dénombre pas moins de 19 types de cellules amacrines [Masland 01]. Ces cellules établissent des inhibitions latérales au niveau de la couche plexiforme interne, c'est-à-dire au niveau de la synapse reliant les bipolaires aux ganglionnaires. Ces cellules sont les moins connues des cellules rétiennes. Les amacrines ne contacteraient pas les cônes de manière sélective [Calkins 96]. Elles auraient un rôle de filtrage spatio-temporel, modélisé comme un filtre passe-haut dans le domaine temporel par Beaudot *et al.* [Beaudot 94]. Ce comportement passe-haut pourrait expliquer le rôle des microsaccades oculaires. L'absence de microsaccades oculaires a comme conséquence d'atténuer le signal, car l'image est alors un signal temporellement constant. Ceci est particulièrement observable en périphérie du champ visuel (Figure 2.24).

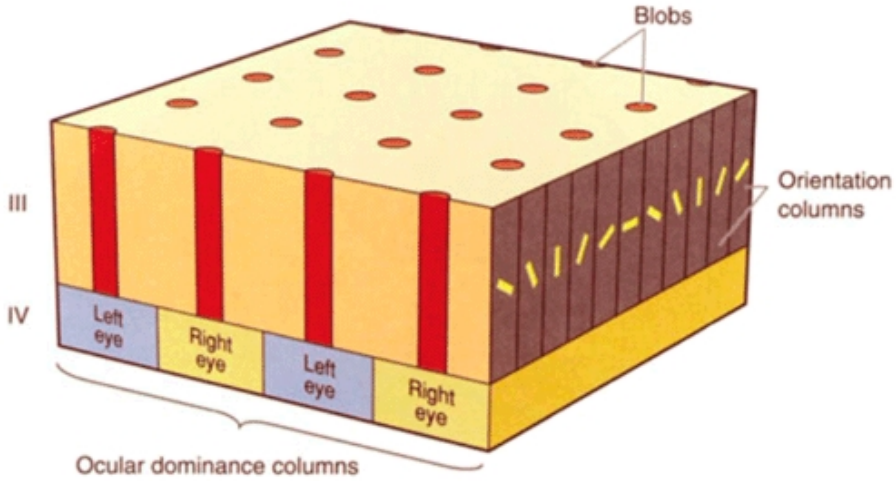


FIG. *2.25: Un module cortical (ou hypercolonne) de V1. Chaque module cortical permet d'analiser une région spatiale du champ visuel.*

2.4 Le cortex visuel primaire

Il est intéressant de constater que l'on retrouve une forme de mosaïque au niveau du cortex visuel primaire. Nous avons vu qu'au niveau de la rétine, la mosaïque établissait un multiplexage spatial des mesures d'intensité lumineuse dans différentes bandes spectrales, dans lesquels les composantes achromatiques et d'oppositions de couleurs étaient mélangées. Au niveau du cortex visuel primaire, la mosaïque code un multiplexage spatial de l'information d'intensité et des informations d'oppositions de couleurs. L'intensité étant elle-même décomposée en juxtaposition d'orientations locales.

Le cortex V1 peut être vu comme une nappe à deux dimensions constituée de plusieurs couches. V1 est constitué de micro-colonnes corticales (dans le sens de l'épaisseur) chacune sensible à une orientation (Figure 2.25). A mesure que l'on se déplace dans une colonne (donc perpendiculairement au cortex) la bande de fréquence, à laquelle les neurones sont sensibles, augmente. Lorsque l'on se déplace parallèlement au cortex, ce sont les orientations qui varient. Les colonnes sensibles aux orientations ne sont pas, ou peu, sensibles à la couleur. Au contraire, il existe des colonnes qui n'ont pas d'orientation privilégiée et qui sont sensibles à la couleur, avec un comportement passe-bas. On parle de "blobs" de cytochrome oxydase, car ce marqueur se concentre spécifiquement dans ces colonnes. La colonne de blob et les intercolonnes d'orientations qui l'entoure forment une macro-colonne de dominance oculaire. Elles sont successivement associée à l'oeil droit et à l'oeil gauche. Enfin le regroupement de 4 x 4 colonnes de dominance oculaire représente un module cortical (Figure 2.25). Il correspond à l'information contenue dans une portion de la rétine de 0.1 degré de champ en fovea, et de 6.0 degrés à 45 degrés d'excentricité.

Cependant, une telle dichotomie en cellules sensibles aux orientations achromatiques et en cellules sensibles aux couleurs basse-fréquences doit impérativement être nuancée. Johnson *et al.* [Johnson 01] ont récemment fait des mesures montrant que, effectivement, les cellules sensibles uniquement à l'intensité (60% des cellules mesurées) ne sont pas sensibles aux basses fréquences (95% d'entre elles), que les cellules sensibles uniquement à des stimuli équilumi-

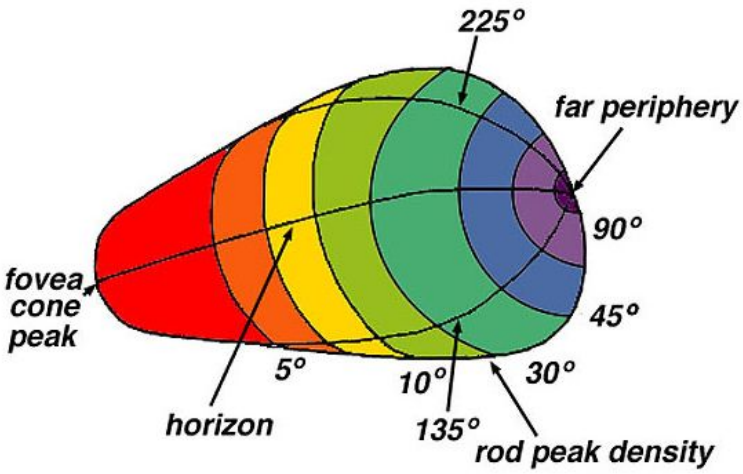


FIG. *2.26: Projection log-polaire sur le cortex visuel primaire V1. La zone correspondant à la fovea est sur-représentée*

nants sont sensibles majoritairement aux basses fréquences (74% de ces cellules), mais que la majorité des cellules sensibles à la couleur est également sensible à l'intensité et est de plus sélective à certaines fréquences, généralement passe-bandes. Ainsi, l'information d'oppositions de couleurs et l'information d'intensité ne serait pas complètement séparées à ce niveau du système visuel. Deux études antérieures [Thorell 84, Lennie 90] auraient également découvert ce genre de cellules, mais dans une plus faible proportion du fait, affirment les auteurs, de l'utilisation de stimuli contenant davantage de contraste en intensité qu'en couleur.

La carte de la rétine se projette en log-polaire sur le cortex visuel primaire [Schwartz 80]. Cette projection a pour effet d'allouer une surface en relation avec la densité d'échantillons. Ainsi la région correspondant à la fovea est davantage représentée que la périphérie de la rétine (Figure 2.26). On parle de magnification corticale.

Par ailleurs, on distingue différentes couches, numérotées de I à IV (un blob occupe les couches I à IVB), vers lesquelles se projettent les voies fonctionnelles issues du corps genouillé latéral [Chatterjee 03] (Figure 2.25). La voie magnocellulaire se projetterait sur la couche IVCα. La voie parvocellulaire se projetterait sur la couche IVCβ, puis sur dans les blobs. Enfin, la voie koniocellulaire, ne portant que de l'information de couleur, se projetterait directement sur les blobs.

3 Performances et limites du système visuel humain

Après avoir passé en revue les différentes caractéristiques du système visuel, intéressons-nous maintenant à ses performances en terme de résolution des images perçues.

En fovea, le filtrage optique est ajusté à la fréquence de Nyquist des cônes, soit environ 60 cycles/degré (Figure 2.28 et 2.29), si on les considère dans leur globalité [Yellott 90]. Le théorème de Shannon est respecté. Cependant, si l'on considère les sous-mosaïques L, M et S séparément, alors la fréquence de Nyquist est largement diminuée et est très inférieure à la fréquence de Nyquist. Pourtant, il n'a été rapporté aucune manifestation d'*aliasing* en fovea. Le modèle d'échantillonnage d'Alleysson *et al.* [Alleysson 05] permet d'expliquer en partie

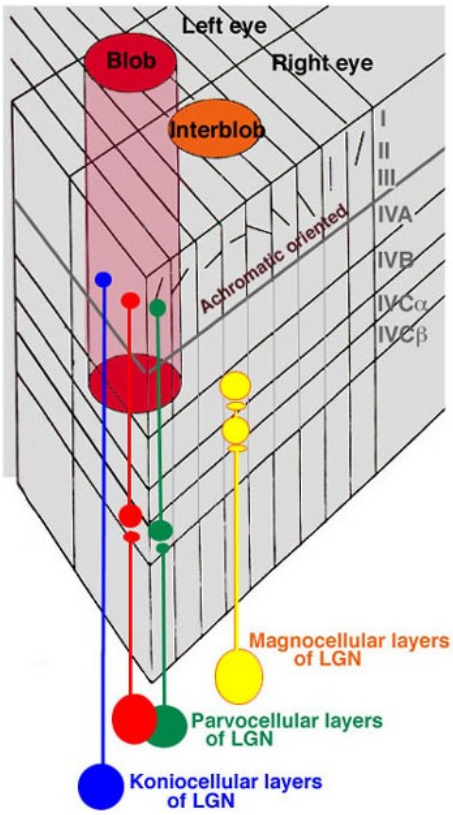


FIG. *2.27: Projection des différentes voies fonctionnelles issue du corps genouillé latéral sur le cortex visuel primaire.*

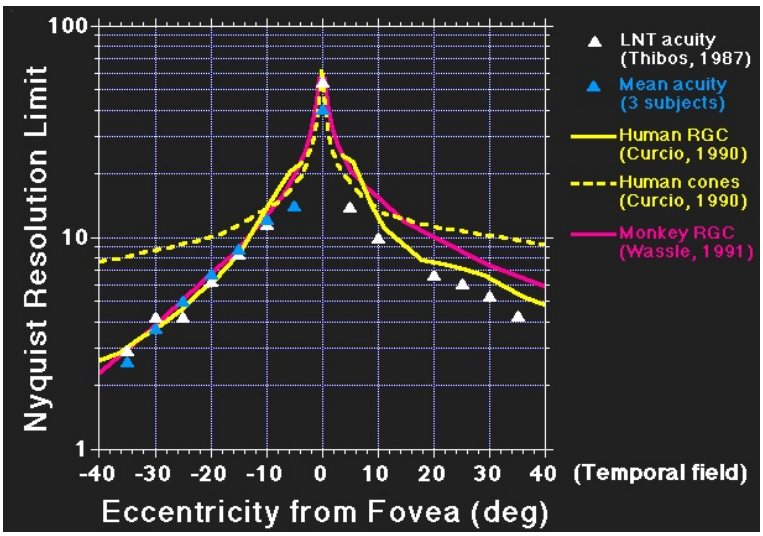


FIG. *2.28: Evolution avec l'excentricité de la fréquence de Nyquist théorique des cônes (en pointillés) et de la fréquence de Nyquist effective du système visuel obtenue par différentes études (autres courbes) [Thibos 87]. La résolution en périphérie est largement inférieure à ce qu'autoriserait la densité des cônes. Le facteur limitant est en fait la densité des cellules ganglionnaires.*

pourquoi l'intensité lumineuse peut avoir une largeur de bande spatiale bien supérieure aux fréquences de Nyquist marginales des sous-mosaïques. Dans cette thèse, nous étudions l'influence de l'irrégularité de l'arrangement des trois sous-mosaïques de cônes. Nous verrons que les ir-

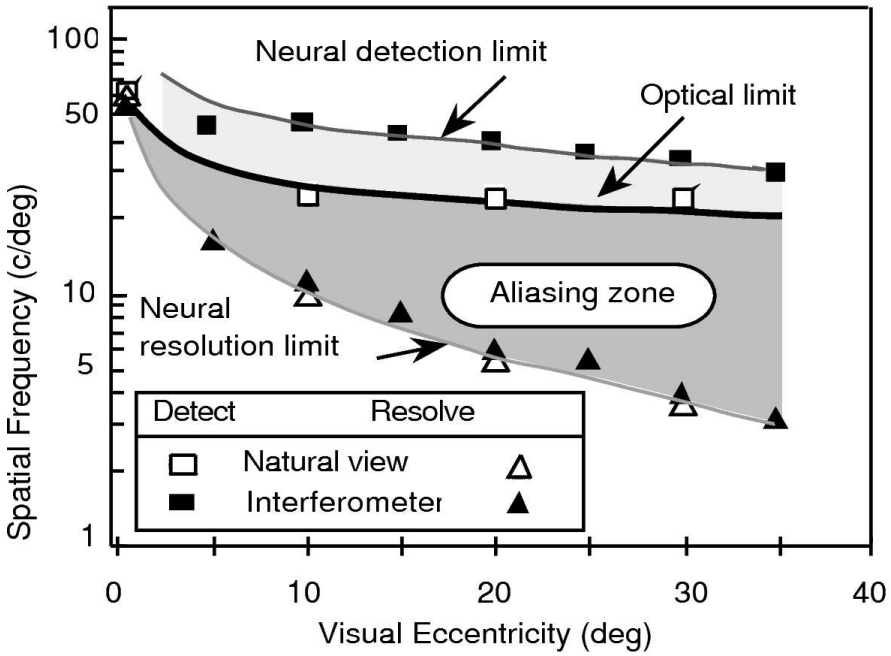


FIG. *2.29: Fréquences limites de résolution et de détection en fonction de l'excentricité chez l'humain [Thibos 91]. Des techniques d'interférométrie permettent d'atteindre des fréquences de détection plus élevée que la limite imposée par la fréquence de coupure optique.*

régularités permettent de briser les motifs réguliers induits par le repliement spectral, et donc d'en diminuer la visibilité.

En périphérie, où la densité de cônes est bien plus faible qu'en fovea, la fréquence de Nyquist est par exemple d'environ 9 cyc./deg. à 30° d'excentricité (Figure 2.30). La fréquence de coupure de l'optique décroît également avec l'excentricité, mais de manière beaucoup moins rapide. Elle serait d'environ 20 cyc/deg ou 50 cyc/deg à cette excentricité selon les études [Thibos 91, Anderson 91]. Ainsi en périphérie, le théorème de Shannon (p. 73) est largement violé : l'image peut contenir des fréquences au moins deux fois plus élevées que la fréquence de Nyquist. Des expériences montrent que l'on peut percevoir effectivement cet *aliasing* [Thibos 00]. Par ailleurs, comme nous l'avons vu, les cônes sont disposés de manière de plus en plus irrégulière en fonction de l'excentricité. Cette irrégularité permet-elle de réduire la perception du repliement spectral ? Contrairement au cas de la fovea, ce repliement spectral affecterait non seulement la composante chromatique mais aussi la composante d'intensité, car c'est la mosaïque globale qui ne respecte pas le théorème de Shannon.

La fréquence de Nyquist effective du système visuel humain a pu être mesurée par des expériences psychophysiques. En périphérie de la rétine, la fréquence de Nyquist perçue ne correspond en fait pas à la fréquence théorique associée à la mosaïque des cônes. Elle correspond en fait à la fréquence associée à la mosaïque des cellules ganglionnaires [Anderson 91]. En effet, il ne faut pas oublier qu'en périphérie, une cellule ganglionnaire intègre les réponses de plusieurs cônes. Le facteur limitant du système visuel n'est dont pas au niveau optique, ni réceptoral, mais post-réceptoral.

Mais attention, l'*aliasing* n'est pas facile à mettre en évidence expérimentalement, car souvent, la mise au point n'est pas faite en périphérie de la rétine [Thibos 00]. Par ailleurs, lorsque

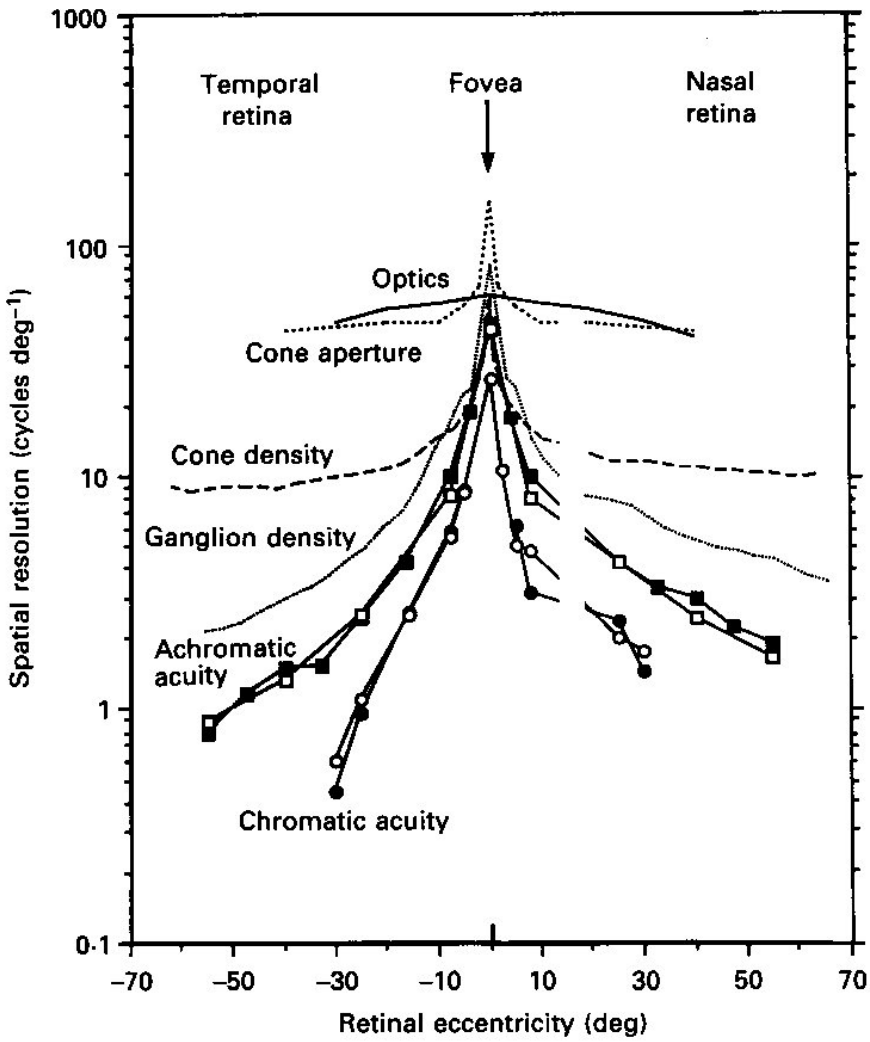


FIG. *2.30: Evolution des fréquences de résolution chromatique et achromatique en fonction de l'excentricité [Anderson 91]. L'acuité chromatique est inférieure à l'acuité achromatique, et décroît plus vite avec l'excentricité.*

des fréquences sub-Nyquist sont présentes en plus des fréquences supra-Nyquist, un phénomène de masquage ne permettrait pas de percevoir les fréquences repliées [Wang 97]. Williams *et al.* montrent un motif permettant de percevoir des fausses couleurs en périphérie du champ visuel, appelée couleurs de Brewster [Williams 91]. Ses fausses couleurs se présentent sous forme de bruit dynamique (dû aux micro-saccades oculaires), c'est à dire une structure très irrégulière qui varie dans le temps.

En ce qui concerne la limite fréquentielle en fovea pour la résolution d'un motif purement chromatique (isoluminant donc), celle-ci est d'environ 20 cyc./deg. [Kelly 83, Anderson 91], trois fois moins élevée que la fréquence limite pour l'intensité. La limite de Nyquist perceptuelle aux stimuli chromatiques décroit avec l'excentricité de manière plus rapide encore que ne le fait la limite de Nyquist pour les stimuli achromatiques, comme le montre la Figure 2.30. Cela suggère qu'il existe des facteurs limitants post-réceptoraux à la perception de la couleur en périphérie du champ visuel. L'augmentation de la taille du champ récepteur des cellules ganglionnaires associées à des connexions non spécifiques par rapport aux types de cône en est

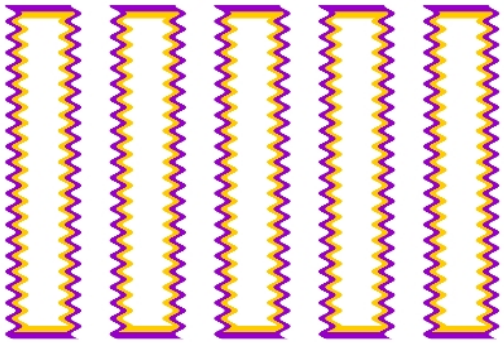


FIG. *2.31: Illustration de l'illusion aquarelle [Pinna 01]. L'intérieur des formes apparaît jaunâtre plutot que blanc.*

peut-être la raison, mais cela reste un sujet de controverse [Martin 01]. Par ailleurs, aucun phénomène d'*aliasing* provoqué par des stimuli isoluminants n'est rapporté, même si la limite perceptuelle est atteinte [Anderson 91] : les stimuli isoluminants sont perçus correctement, ou ne sont pas perçus du tout. On remarquera que les graphiques des Figures 2.28 et 2.30 ne donnent pas exactement les mêmes valeurs pour la résolution achromatique. Sur la Figure 2.30, les fréquences de résolution achromatique sont légèrement en-dessous de la limite théorique imposée par la densité de cellules ganglionnaires, tandis qu'elles suivent parfaitement cette limite dans le cas de la Figure 2.28. Ces différences sont imputables aux protocoles expérimentaux utilisés (correction ou non des aberrations chromatiques, par exemple). Dans des conditions normales de vision, la limite théorique fixée par les cellules ganglionnaires n'est pas atteinte.

Il est rapporté dans plusieurs études que l'intensité peut faciliter la discrimination de la couleur [Hilz 70, Hilz 74, Cole 90]. Il y aurait par exemple une facilitation de la perception de la couleur de petites cibles colorées sur fond coloré lorsqu'un fin contour d'intensité entoure la cible. De la même manière, la perception d'un motif sinusoïdal purement chromatique est facilité par la présence d'un motif cohérent en intensité. Ainsi les voies d'analyse achromatique et chromatique ne seraient pas totalement indépendantes dans le système visuel.

Enfin, une illusion visuelle récemment découverte met en évidence l'intéraction qui existe entre la structure spatiale d'une scène et le traitement de la couleur : ce sont les couleurs induites, ou l'illusion aquarelle [Pinna 01], Figure 2.31. Ce phénomène apparaît lorsqu'un contour fermé achromatique est flanqué du côté intérieur d'un coutour coloré.

4 Les modèles neurocomputationnels de vision des couleurs

Nous présentons ici les différents modèle neurocomputationnels de constructions des axes de perception chromatiques et achromatique à partir du signal issu des cônes. Deux courants se dégagent parmi ces modèles. L'un, le courant majoritaire au sein de la communauté, défend l'idée que la voie parvocellulaire véhicule à la fois l'information achromatique et l'opposition rouge-vert jusqu'au cortex visuel. L'autre soutient qu'il y aurait des voies spécifiques séparées pour véhiculer l'information d'intensité d'une part et l'information d'opposition de couleurs

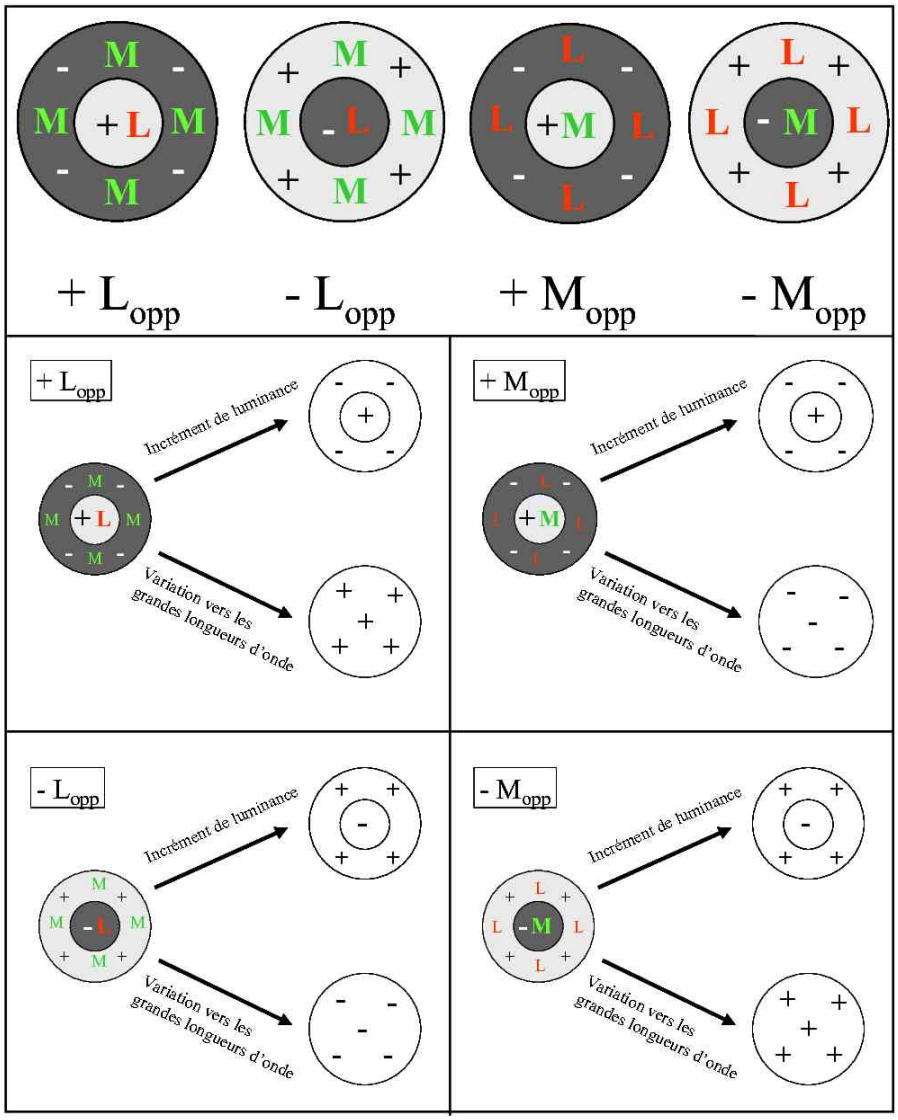


FIG. *2.32: Champs récepteurs du modèle de De Valois et De Valois. Ce sont leurs sommes deux à deux qui permettraient d'obtenir les signaux d'intensité et d'opposition de couleurs.*

d'autre part (Calkins *et al.*).

Modèle de De Valois et De Valois

De Valois et De Valois ont proposé en 1993 un modèle de couleur dit multi-étage [De Valois 93]. Une motivation de leur modèle basé sur plusieurs niveaux, émane de la constatation que les réponses en opposition de couleurs au niveau du corps genouillé latéral ne correspondent pas tout à fait aux oppositions de couleur réellement perçues. Une seconde motivation provient du fait que les cellules du corps genouillé latéral répondent à la fois à la couleur et à l'intensité.

Chaque cône est connecté à quatre cellules bipolaires : les bipolaires *midget* ON et OFF et les bipolaires diffuses ON et OFF. Les cellules horizontales et amacrines sont supposées constituer le pourtour inhibiteur du champ récepteur des ganglionnaires midget à la sortie de la rétine. Deux versions alternatives sont présentées qui diffèrent par la sélectivité ou non du

pourtour : $L_o = L - M$ ou $L_o = L - (LMS)$ (l'indice o pour "oppositions"). Cette sélectivité ou non n'a que peu d'influence sur le modèle, d'après les auteurs. Le champ récepteur de la voie magnocellulaire effectuerait une sommation de L et de M en son centre, et soustrairait l'intensité LMS dans le pourtour, créant ainsi un antagonisme spatial : $L + M - (LMS)$ (pour la voie ON), $-L - M + (LMS)$ pour la voie OFF.

Le champ récepteur est décomposé en deux parties : un antagonisme spatial, et une intégration spécifique à un cône. C'est à ce niveau que sont séparées les informations spatiales et chromatiques du modèle. Les oppositions de couleurs sont retrouvées par addition et soustraction des champs récepteurs, tel que l'illustre la Figure 2.32.

D'après le modèle, c'est l'addition des voies OFF et des voies ON qui permettent d'obtenir des oppositions de couleur. Il y a ici une grande confusion si l'on considère que la voie OFF véhicule l'information négative et que la voie ON véhicule l'information positive. Dans ce cas, effectuer la somme des deux n'a pas de sens en terme d'opposition de couleurs. La voie OFF n'étant pas l'opposée de la voie ON (mais seulement les valeurs négatives du signal), l'addition des signaux ON d'un couleur avec les signaux OFF d'une autre ne crée pas une opposition de couleurs mais bien une addition.

Mais, finalement, même si on considèrait que la rétine effectuait réellement des soustractions entre les champs récepteurs, l'objection la plus importante que l'on puisse faire est le fait de considérer que les signaux L, M et S sont disponible pour chaque position spatiale, alors que ce n'est pas le cas (les cônes sont spatialement multiplexés, il ne peuvent pas se superposer). En termes de traitement du signal, cela revient à faire une interpolation aux plus proches voisins de chaque sous-mosaïque, puis à effectuer la décomposition en intensité et en oppositions de couleurs. Cela donne de très mauvais résultats en pratique, on est loin des performances du système visuel humain, qui comme on l'a vu est proche de la fréquence de Nyquist de la mosaïque globale en fovea. Il faut donc un modèle qui soit plus précis, pour comprendre exactement comment se comportent les signaux.

4.0.1 Ingling et Martinez : le multiplexage fréquentiel de l'intensité passe-bande et de la chrominance

Ingling et Martinez [Ingling 83, Martinez-Uriegas 94] ont proposé un modèle de multiplexage des informations d'intensité et de chrominance en fovea pour le canal parvocellulaire, c'est-à-dire le canal issu des cellules ganglionnaires *midget*. Les auteurs partent d'une modélisation du champ récepteur de ces dernières. Comme nous l'avons vu dans la partie descriptive du système visuel, ces cellules présentent un antagonisme centre/pourtour à la fois spatial et spectral. La sélectivité spectrale de la périphérie n'a pas été démontrée, cependant, les auteurs en font l'hypothèse. Le centre et le pourtour du champ récepteur sont modélisés respectivement par les fonctions N et W. Ces fonctions sont toutes deux positives, d'intégrales égales, et N (comme "narrow") est à support plus étroit que W (pour "wide"), correspondant au fait que le pourtour est plus étendu que le centre. Du point de vue traitement du signal, N et W sont les réponses impulsionnelles de deux filtres passe-bas, la fréquence de coupure de N étant plus élevée que celle de W. Le signal issu des cônes est appelé L et M, selon qu'il vient d'un cône L ou d'un cône M. Ainsi, le champ récepteur L_{on} associé à une cellule ganlionnaire *midget* de

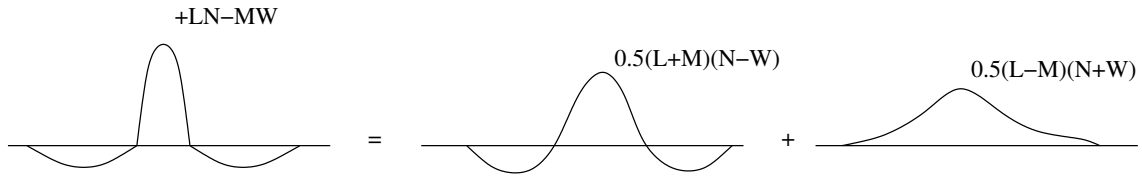


FIG. *2.33: Modélisation du champ récepteur d'une cellule ganglionnaire midget par Ingling et Martinez, et décomposition en somme d'un champ achromatique à comportement passe-bande et d'un champ chromatique à comportement passe-bas.*

type ON serait de la forme :

$$L_{on}(x, y) = L(x, y)N(x, y) - N(x, y)W(x, y)$$

Ce champ récepteur peut être réécrit de la manière suivante (pas souci de concision, les variables ne sont pas réécrites) :

$$L_{on} = 1/2\,\{(L + M)(N - W) + (L - M)(N + W)\}$$

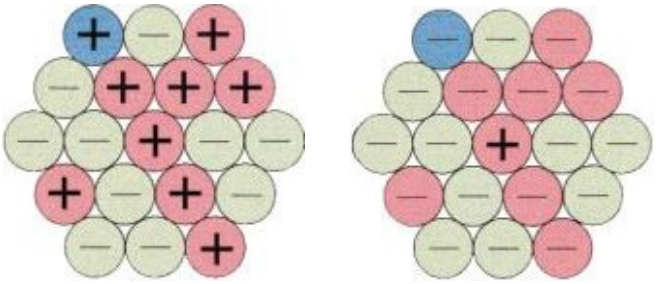
faisant apparaître une composante achromatique $(L + M)/2$ et une composante chromatique $(L - M)/2$. Il faut de plus observer le comportement fréquentiel des fonctions de pondération de ces deux composantes. $N - W$ est la différence entre deux filtres passe-bas, c'est donc un filtre passe-bande. L'intensité a donc un comportement passe-bande. $N + W$ est la somme de deux filtres passe-bas, c'est donc un filtre passe-bas. La chrominance a donc un comportement passe-bas. Les formes des champs récepteur sont illustrés à la Figure 2.33.

Les critiques que l'on peut faire de ce modèle sont les suivantes. Tout d'abord le choix du facteur $1/2$ dans la décomposition en composantes chromatique et achromatique n'est pas justifiée. Est-ce parce que l'auteur considère qu'il y a autant de cônes L que de cônes M ? Ensuite, comment justifier le fait que l'on ait un champ récepteur périphérique alimenté à la fois par des cônes L et des cônes M, tandis que l'hypothèse de départ est que la périphérie est discriminante vis-à-vis du type de cône ? Enfin, et c'est la critique la plus importante, ce modèle ne prend pas en compte le fait que la mosaïque des cônes effectue un multiplexage spatial des différentes informations L, M et S. On ne comprend alors pas quelles conséquences peut avoir ce multiplexage sur la nature de l'information véhiculée et sur les limites théoriques de séparation des informations chromatique et achromatique.

Ainsi la démarche de Ingling et Martinez qui part de considérations sur le champ récepteur supposé des cellules ganglionnaires pour inférer des propriétés fréquencielles sur les composantes d'intensité et d'oppositions de couleurs est intéressante et innovante. Cependant le modèle qui en découle n'est pas suffisamment précis. Il lui manque une base théorique solide.

4.0.2 Alleysson et Hérault : un modèle d'échantillonnage par une mosaïque

Le modèle proposé par Alleysson *et al.* prend en compte de manière précise le multiplexage spatial effectué par la mosaïque des cônes [Alleysson 05]. Ce modèle montre que le sous-échantillonnage et le multiplexage spatial effectué par les cônes correspond à un multiplexage fréquentiel des informations d'intensité et d'oppositions de couleurs. La contribution centrale de ces travaux provient du fait que l'on ait un multiplexage de l'intensité et des oppositions de



(a) Disposition *optimale* pour l'estimation d'oppositions de couleurs

(b) Disposition *optimale* pour l'estimation de l'information spatiale

FIG. *2.34: Modèle de Calkins. Les cellules ganglionnaires seraient spécialisées dans l'estimation de l'information spatiale et des oppositions de couleur en fonction de leurs champs récepteur.*

couleur dans le domaine de Fourier, sans qu'un cablage chromatique explicite ne soit nécessaire. Nous détaillons ce modèle, qui est à la base de ce travail de thèse, au chapitre suivant (p. 61).

Si ce modèle offre de nouvelles perspectives en matière de représentation du signal au niveau de la mosaïque des cônes, en particulier pour un arrangement régulier de la mosaïque des photorécepteurs, il est cependant insuffisant dans le cas d'un arrangement irrégulier.

4.0.3 Calkins *et al.* : la controverse du lieu de la séparation des informations spatiales et chromatiques dans le SVH

Contrairement aux trois modèles précédents, Calkins *et al.* considèrent que le système parvocellulaires ne véhicule que l'information d'intensité [Rodieck 91, Calkins 99]. La séparation entre intensité et oppositions rouge-vert aurait alors lieu dès les premières synapses de la rétine. L'opposition rouge-vert serait alors véhiculée par des ganglionnaires spécifiques, à l'image de la circuiterie dédiée aux oppositions bleu-jaune. Ces ganglionnaires auraient un antagonisme spectral, mais pas d'antagonisme spatial. Quelques études rapportent l'existence de telles cellules en faible nombre [de Monasterio 78, Reid 92], mais ces études sont incohérentes quant à la proportion de ces cellules et sont donc à prendre avec précaution. La voie parvocellulaire quand à elle ne conduirait que l'information d'intensité. Cela obligerait à avoir une voie fonctionnelle supplémentaire, ce qui serait certes coûteux d'un point de vue écologique, mais les auteurs affirment qu'il serait difficile de faire la séparation intensité/chrominance par filtrage ultérieurement car les deux informations seraient mélangées. On objectera qu'elles sont déjà mélangées au niveau de la mosaïque des cônes.

Les auteurs justifient leur modèle en considérant les structures des champs récepteurs et en leur associant des spécialisations fonctionnelles. Le champ récepteur présenté à la Figure 2.34(a), constitué d'une opposition spectrale coextensive serait spécialisé dans l'estimation des opposition de couleurs. A l'opposé, le champ récepteur de la voie parvocellulaire (Figure 2.34(b)), qui soustrait les valeurs du voisinage (L et M indiscriminés) au cône central uniquement, serait spécialisé dans l'estimation de l'information spatiale. S'il est compréhensible qu'une opposition spectrale sur l'ensemble du champ récepteur permette d'estimer une information chromatique

(l'intensité s'annule par soustraction), en revanche la seconde affirmation est bien moins évidente. Certes, la soustraction d'une somme pondérée des valeurs L et M sur une large région permet de supprimer les basses fréquences d'intensité (comme le montre le modèle d'Alleysson), et donc de mettre plus en évidence les contours de l'image. Mais le problème est que l'information chromatique reste entièrement présente dans ce signal, puisqu'elle n'est pas retranchée. Dans ce cas, cela signifierait soit que la présence de couleur ne perturbe en rien le signal d'intensité, soit qu'il y a un traitement post-rétinien pour supprimer ce *bruit chromatique* sur l'intensité. Dans le cas de la seconde hypothèse, on peut se demander alors pourquoi le cortex n'utilise pas ce mécanisme pour estimer les oppositions de couleur, puisque cela permettrait d'économiser des fibres au niveau du nerf optique. La première hypothèse est possible. Il n'est pas possible de trancher en l'état actuel des connaissances.

5 Conclusion

L'espace perceptif des couleurs est un espace à trois dimensions, induit par les trois types de photorécepteurs qui composent la surface sensible de la rétine. Le système visuel effectue un changement de base dans cet espace pour créer trois axes perceptifs : un axe achromatique d'intensité lumineuse de la scène imagée, et deux axes d'oppositions de couleurs, codant son contenu spectral.

Le système visuel, fruit d'une évolution sur des millions d'années, est une machinerie extrêmement complexe. L'état actuel des connaissances en anatomie et en électrophysiologie ne permet pas de comprendre en totalité les mécanismes mis en jeu dans la construction de ces trois axes perceptifs. Aucun des modèles de vision des couleurs de la littérature ne sont satisfaisants. A l'exception de celui proposé par Alleysson *et al.*, tous oublient de prendre en compte une donnée importante du problème : les échantillons chromatiques sont sous-échantillonnés et spatialement multiplexés. La conséquence est qu'à chaque position spatiale, une seule donnée chromatique est disponible, posant ainsi des problèmes d'acuités spatiale et chromatique.

Si le modèle d'Alleysson *et al.* prend effectivement en compte cette problématique, offrant une solution élégante dans le cas d'un arrangement régulier des photorécepteurs, il ne permet pas cependant de résoudre le problème lié à un arrangement irrégulier, comme c'est le cas dans la réalité. L'extension au cas irrégulier est l'objet même de ce travail de thèse.

Il est fort probable que les microsaccades oculaires facilitent la séparation de l'information issue des cônes en une information lumineuse et une information de couleur. L'amplitude de ces mouvements oculaires imperceptibles et involontaire peut couvrir plusieurs photorécepteurs. Cependant nous ne prenons en compte la dimension temporelle dans notre étude. La rétine est donc fixe, et la scène n'est pas en mouvement. En effet, si nous intégrons la dimension temporelle avant d'avoir explorer en profondeur le cas spatial, nous risquons d'aboutir à une solution non optimale en émettant des hypothèses trop simplificatrices sur les conditions de mouvement oculaire.

Enfin, en plus de la temporalité du problème, un autre facteur pourrait faciliter le mécanisme de construction des voies perceptives serait la binocularité de la vision. Il est en effet rapporté que les deux yeux ne sont pas indépendants pour la vision des couleurs [Medina 06]. Le temps

de réaction visuel serait diminué en vision binoculaire par rapport à la vision monoculaire dans des tâches de discrimination de variations d'intensité achromatique et dans le cas de variations isoluminantes chromatiques. Il y aurait plus de facilitation dans le cas de variations d'intensité que dans le cas de variations isoluminantes [Jiménez 02]. Cependant, cette facilitation est minime, et on peut légitimement considérer que le système visuel peut séparer l'information d'intensité et d'oppositions de couleurs en conditions monoculaires.

Dans le chapitre suivant, nous étudions le mécanisme de vision des couleurs d'une espèce dont la démographie a connu une forte croissance ces dernières années, il s'agit des appareils photographiques numériques.

Chapitre 3

L'acquisition d'une image couleur par un appareil photographique numérique

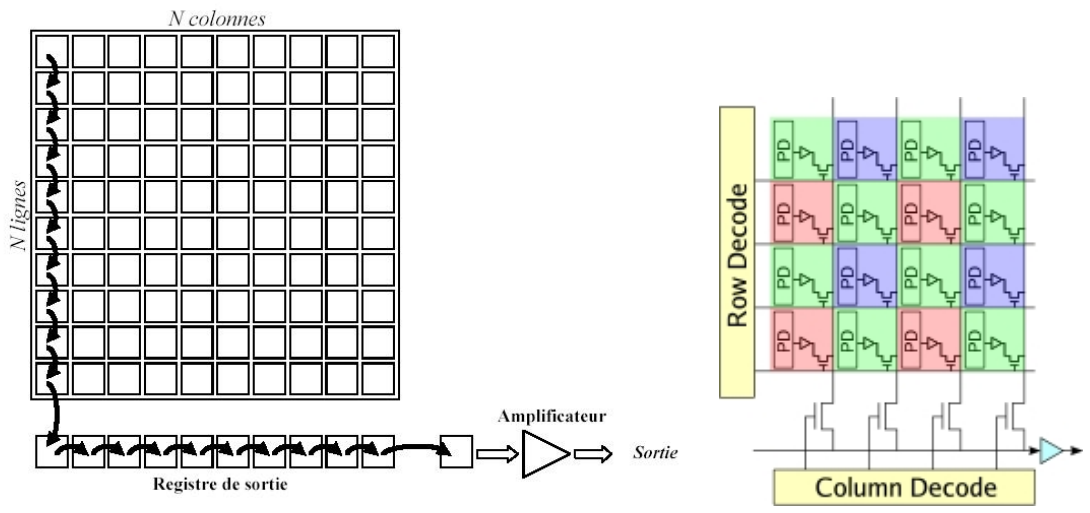
Il existe une forte analogie entre la rétine humaine et les capteurs de caméras numériques dites "mono-capteur"[1], en ce sens que la scène est échantillonnée à l'aide de trois classes de photorécepteurs multiplexées spatialement et sensibles chacune à une gamme de longueur d'onde. Les images échantillonnées par les caméras numériques nécessitent avant affichage une étape d'interpolation, que nous appellerons démosaïçage. Cette étape permet de retrouver une image à trois composantes chromatiques par pixel, c'est-à-dire une image en couleur.

Nous considérons que l'étape de démosaïçage est un équivalent de ce qui est accompli par le système visuel humain lors de la construction du canal de luminance et des deux canaux d'oppositions de couleurs. La connaissance de la luminance et des deux oppositions de couleur est en effet équivalente à la connaissance des trois informations rouge, verte et bleue, à un changement de base près.

Le bénéfice que peuvent apporter les caméras numériques à la compréhension de la vision des couleurs est double. D'une part, les algorithmes disponibles dans la littérature du démosaïçage peuvent être une source d'inspiration pour élaborer des modèles de vision des couleurs. D'autre part, c'est un formidable domaine d'expérimentation qui permet de confronter les modèles théoriques à la réalité du traitement d'image. La communauté de la vision des couleurs étant majoritairement constituée de physiologistes, de psychologues et d'anatomistes, une démarche d'ingénieur peut permettre de corroborer ou d'infirmer certaines hypothèses. Par exemple, l'implémentation du modèle des De Valois [De Valois 93] sous forme d'interpolation par copie de pixel en montre bien les limites.

Ce chapitre comporte deux parties. Dans un premier temps, nous décrivons l'*anatomie* d'une caméra numérique. Une caméra est proche du système naturel que constituent l'oeil et le cortex visuel, avec une lentille, un capteur arrangé sous forme de mosaïque et la présence d'une unité de calcul pour traiter les données mesurées. Dans un second temps, nous établissons un état de l'art du domaine du démosaïçage. Nous dégageons les principales caractéristiques des

[1] les caméras basées sur trois capteurs ne font pas l'objet de notre attention. Ces caméras permettent d'obtenir une très bonne qualité d'image, mais leur coût de fabrication et leur colume en font des dispositifs réservés aux professionnels.



(a) Matrice CCD (b) Matrice CMOS

FIG. *3.1: (a) Transferts de charges dans un capteur CCD. (b) Architecture de base d'un capteur CMOS (PD=photodiode).*

algorithmes de la littérature. D'une manière générale, deux propriétés sont mises à profit par ces algorithmes : la corrélation spectrale entre les plans couleurs, et la corrélation spatiale des pixels. L'algorithme qui nous intéressera le plus, dans le cadre de ce travail de thèse, sera celui d'Alleysson *et al.* [Alleysson 05]. Cet algorithme est en effet dérivé d'un modèle d'échantillonnage inspiré par le système visuel humain. Il montre que la représentation de la luminance et des oppositions de couleurs reste valide pour une image en mosaïque et qu'elle dépend de la proportion de chaque classe de photorécepteurs dans la mosaïque.

1 Anatomie d'un appareil photographique numérique

Nous décrivons ici l'anatomie d'un système artificiel de vision, à savoir une caméra ou un appareil photographique numériques.

1.1 Le capteur photosensible

L'élément de base d'une caméra numérique est son capteur photosensible, qui permet de convertir des photons en signal électrique. Deux technologies coexistent sur le marché actuel des capteurs d'images [Fossum 93, Magnan 03] : la technologie CCD ("*Coupled Charge Device*") et la technologie CMOS ("*Complementary Metal Oxyde Semi-conductor*"). Les capteurs CCD ont été les premiers à être exploités commercialement. La lumière incidente traverse une grille en polysilicium d'une capacité MOS et une charge est accumulée sous la grille. La charge est ensuite *transférée* d'un pixel à l'autre en jouant sur les potentiels des grilles des pixels (les charges suivent les puits de potentiel). Ainsi, la lecture d'une image se fait par propagation de proche en proche dans le sens des colonnes, puis dans le sens des lignes, pour être finalement amplifiée (Figure 3.1(a)). La lecture d'une image est donc séquentielle, ce qui peut être pénalisant si l'on souhaite traiter uniquement une région d'intérêt. Le transfert se faisant de proche en

proche, le taux de transfert avec succès d'un pixel à l'autre est primordial, puisqu'un électron non transféré se retrouvera comptabilisé sur un pixel voisin, générant ainsi des intérférences. Le fait que la détection et le transfert des charges soient faits par le même élément oblige, par ailleurs, à effectuer une obturation le temps du transfert des charges. Afin de pouvoir collecter une image pendant le transfert, deux dispositifs exploitent des registres de stockage non exposés à la lumière : le CCD à transfert de trame, où l'image est transférée vers une matrice de même taille non exposée ; et le CCD interligne, où des registres de transport protégés sont intercalés entre les lignes. Dans le cas du transfert de trame, la taille du capteur est doublée. Dans le du CCD interligne, c'est le facteur de remplissage qui passe de 100% à 20-30%, le facteur de remplissage étant le ratio de la surface photosensible d'un pixel ramené à sa surface totale. Ce facteur de remplissage peut être virtuellement augmenté par l'ajout d'une microlentille sur chaque pixel, faisant converger les rayons lumineux verts la zone photosensible du pixel. Enfin, le capteur CCD peut souffrir d'éblouissement. Ce phénomène apparaît lorsqu'un collecteur est saturé en électrons. Les électrons se diffusent alors vers les pixels voisins de la même colonne. Des dispositifs appelés "*Lateral Overflow Drain*" permettent de contenir l'éblouissement, au détriment du facteur de remplissage.

Les capteurs CMOS, contrairement aux capteurs CCD qui eux sont basés sur une technologie dédiée, tirent partie des avancées technologiques régulières du secteur des semi-conducteur CMOS. Un capteur CMOS est organisé comme une mémoire, c'est-à-dire que les pixels sont adressables en ligne et en colonne (Figure 3.1(b)). Ainsi, l'adressage aléatoire est possible, on peut charger uniquement une région d'intérêt. Les capteurs actuels sont dits "à pixels actifs" du fait de la présence d'un amplificateur au sein même de chaque pixel, qui délivre le signal à un bus vertical. Du fait de la présence des transistors et des bus, le facteur de remplissage est réduit à un ordre de grandeur de 40-60% et la zone photosensible n'est pas à la surface du capteur.

Pour résumer, on peut dire que le capteur CCD a l'avantage d'avoir :

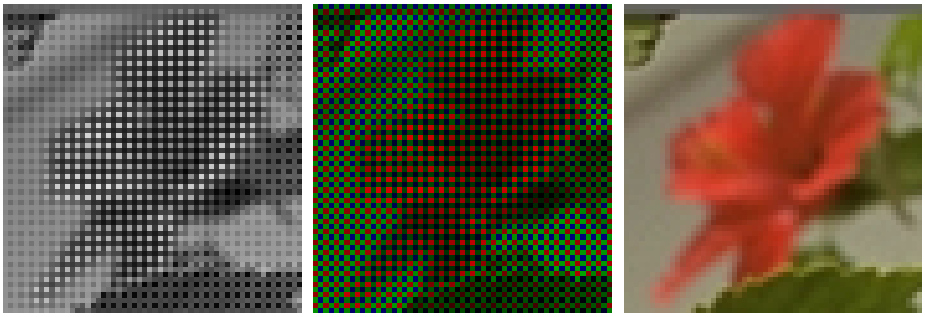
- un meilleur rendement quantique (rapport entre le nombre d'électrons collectés et le nombre de photons incidents),
- un meilleur facteur de remplissage,
- un courant d'obscurité (bruit apparaissant dans l'obscurité) plus faible. Les prises de vue en basse luminosité sont donc moins bruitées.

En revanche, le capteur CMOS a les avantages suivants par rapport au capteur CCD :

- il coûte moins cher à la fabrication (c'est une technologie standard),
- la lecture est plus rapide,
- il consomme moins de puissance,
- il n'est pas sujet à l'éblouissement.

1.2 La matrice de filtres couleur

Un capteur photosensible tel quel n'est pas sensible à la couleur. Il est uniquement sensible à une intensité lumineuse, mais pas aux variations spectrales de la lumière. Pour rendre le capteur sensible à la couleur, un filtre de couleur est placé sur chaque pixel. Le filtre est soit rouge, soit vert, soit bleu (les primaires de la synthèse additive), et permet de mesurer l'intensité lumineuse dans une gamme de longueurs d'onde. Ces filtres sont multiplexés spatia-



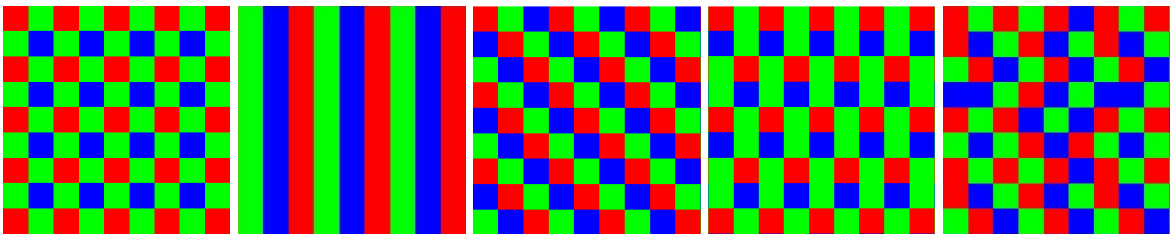
(a) (b) (c)

FIG. *3.2: (a) Image en niveaux de gris, mesurée par le capteur ; (b) Image mesurée démultiplexée, à une composante couleur par pixel ; (c) Image couleur reconstruite, à trois composantes couleur par pixel.*

lement, on parle de matrice de filtres de couleur, ou de mosaïque, par analogie à la mosaïque des cônes. L'image ainsi échantillonnée est une image en niveaux de gris (Figure 3.2(a)), dont chaque pixel donne l'information d'intensité lumineuse dans la gamme de longueurs d'onde associée au filtre qui le recouvre. Cette image, contenant une information chromatique spatialement multiplexée, peut être démultiplexée (Figure 3.2(b)). On obtient alors une image couleur dont chaque pixel contient une seule information chromatique sur les trois nécessaires à la représentation d'une image couleur. Cette image couleur est donc sous-échantillonnée. Une étape algorithmique d'interpolation est nécessaire pour obtenir une image couleur pleine résolution (Figure 3.2(c)). Cette étape s'appelle le démosaïçage. Comme dans tout problème d'échantillonnage d'interpolation, elle peut générer des artefacts de reconstruction. Ces artefacts seront décrits plus tard dans ce chapitre, au paragraphe présentant le modèle d'échantillonnage d'Alleysson *et al.* [Alleysson 05].

L'arrangement spatial des filtres de couleur à un impact direct sur la qualité des images reconstruites, comme nous le verrons au Chapitre 5. La Figure 3.3 montre différents arrangements possibles de la mosaïque de filtres. La mosaïque la plus utilisée dans l'industrie est celle proposée par Bayer en 1976 [Bayer 76] (Figure 3.3(a)). Sa suprématie est telle, que tous les algorithmes paraissant dans la littérature lui sont dédiés. Elle a la particularité de contenir deux fois plus de filtres verts que de filtre rouges ou bleus. La justification de cette différence de proportions réside dans le fait que la sensibilité à la lumière atteint un maximum dans les longueurs d'onde associées au vert. Une autre particularité est qu'à chaque colonne et à chaque ligne, seulement deux des trois types de filtres sont présents. Cela a des incidences sur les résolutions verticale et horizontale de l'image, et sur la manière de calculer les valeurs manquantes, comme nous le verrons au cours de ce chapitre.

Les fonctions de sensibilités des filtres jouent un rôle important dans la qualité des images [Alleysson 03]. Des sensibilités très proches permettront un démosaïçage plus aisé, du fait de la plus grande corrélation entre les plans couleur, au détriment du rendu des couleurs. Cependant, la conversion vers l'espace de couleur d'un dispositif d'affichage, dont les composantes sont moins corrélées, amplifie le bruit. D'après l'étude menée dans [Alleysson 03], les défauts de démosaïçage et de rendu des couleurs seraient minimisés lorsque les trois fonctions de sensibilités sont équi-réparties dans le spectre de la lumière visible. L'influence des fonctions de



(a) Bayer (b) Bandes verticales (c) Bandes diagonales (d) [Lukac 05b] (e) Pseudo-aléatoire

FIG. *3.3: Exemples de mosaïques de filtres couleur.*

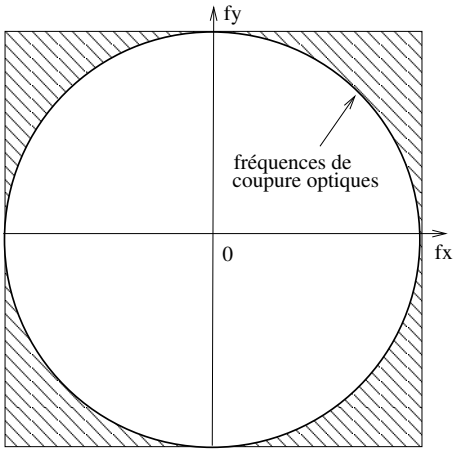


FIG. *3.4: Cercle représentant la fréquence de coupure d'une optique à ouverture circulaire, dans le domaine de Fourier. Les fréquences de la zone hachurée sont coupées.*

sensibilité spectrale ne sera pas étudiée dans cette thèse.

1.3 Le système optique

Le système optique d'une caméra se décompose généralement en deux parties. La première est constituée d'une lentille se trouvant au point d'entrée du flux lumineux. Son rôle est de faire converger les rayons lumineux vers le capteur photosensible. Cette lentille est l'équivalent du cristallin chez l'homme. L'optique est généralement circulaire, et joue également un rôle de filtre passe-bas anti-repliement. Une optique circulaire associée à un capteur dont les pixels sont sur une grille cartésienne implique que certaines fréquences ne peuvent pas être atteintes (si l'on souhaite éviter tout repliement spectral). La Figure 3.4 illustre cela, en montrant le cercle représentant la fréquence de coupure optique de la lentille, tangent au carré représentant les fréquences maximales autorisées par la grille du capteur. Les fréquences de la zone hachurée sont coupées par le filtre optique. Nous verrons au Chapitre 5, que cette zone peut-être mise à profit pour optimiser l'arrangement de la mosaïque de filtres de couleur.

Un défaut optique courant - survenant généralement lorsque l'ouverture de la pupille est maximale - s'appelle le "vignettage". Il a pour effet d'assombrir les bords de l'image, en particulier les coins. Sa cause vient du fait que la quantité de lumière atteignant chaque cellule photosensible dépend de l'inclinaison du flux incident, et donc de l'eccentricité (Figure 3.5). La quantité de lumière recueillie est maximale au centre du capteur, et diminue à mesure que l'on

s'approche des bords. Ce défaut peut-être facilement corrigé dans un système numérique, par une fonction compensant la diminution d'intensité.

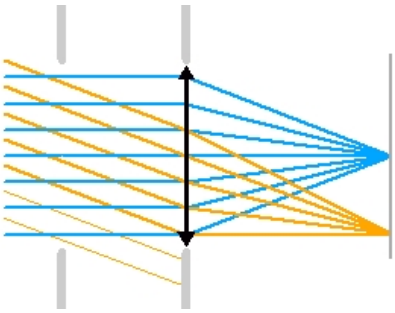


FIG. *3.5: Illustration du phénomène de vignettage. Selon son inclinsaison, une partie du faisceau de lumière incident est bloqué. Il en résulte une diminution de l'éclairement avec l'eccentricité.* *(Crédit : "Observatoire de Paris/U.F.E.")*

Le système optique d'une caméra est également constitué - généralement - d'un réseau de microlentilles. C'est le cas pour les capteurs dont le facteur de remplissage est faible et/ou dont la zone sensible n'est pas directement à la surface du capteur. A l'image de la lumière éclairant un puits, la zone sensible n'est atteinte que par les rayons dont l'angle est proche de la verticale (Figure 3.6). Une microlentille placée à la surface du pixel permet d'augmenter la quantité de lumière atteignant la zone sensible, et donc d'augmenter sa sensibilité.

1.4 Autres traitements algorithmiques

En plus du démosaïçage et de l'anti-vignettage, de nombreuses fonctions interviennent dans la chaîne de traitement d'une image acquise par une caméra. On peut citer, entre autres, le débruitage de l'image acquise (gestion du bruit intrinsèque au capteur, mais aussi du bruit poivre et sel dû aux pixels défectueux), la balance des blancs (afin de régler le niveau d'intensité du blanc), ou encore le *tone mapping* (afin d'avoir un rendu de grande dynamique).

Nous nous focalisons dans notre travail sur le problème du démosaïçage, à la fois sur l'algorithme de reconstruction et sur l'arrangement spatial des filtres de couleur. Nous supposons, dans la suite de ce rapport, qu'une image acquises au travers d'une mosaïque de filtres de couleur peut être simulée par un simple sous-échantillonnage d'une image couleur sans défaut, suivi d'un multiplexage spatial. Nous supposons donc que les problèmes liés au bruit, à la dynamique ou encore au vignettage sont résolus. Nous décrivons dans la suite du chapitre quelques algorithmes s'appliquant à la mosaïque de Bayer.

2 Etat de l'art du démosaïçage dans les caméras numériques

Le démosaïçage a vu un engouement grandissant depuis le début des années 80 grâce à l'avènement des appareils photographiques numériques. De nouveaux algorithmes paraissent presque tous les mois dans la littérature scientifique. Le graphique[2] de la Figure3.7 montre

[2]Le script permettant d'obtenir ce graphique est disponible sur Internet à l'adresse http ://ats.cs.ut.ee/u/kt/stuff/scholartrend/

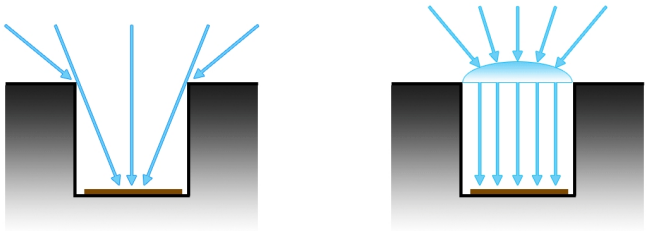


FIG. *3.6: Une microlentille recouvrant le pixel permet de diriger le flux lumineux au fond du puits vers la zone photosensible.*

l'évolution de la probabilité de trouver l'expression "*color filter array*" dans les articles scientifiques contenant l'expression"*image processing*", en fonction des années. Comme on le voit sur ce graphique, le nombre d'articles traitant de matrice de filtres couleur explose depuis le début des années 2000.

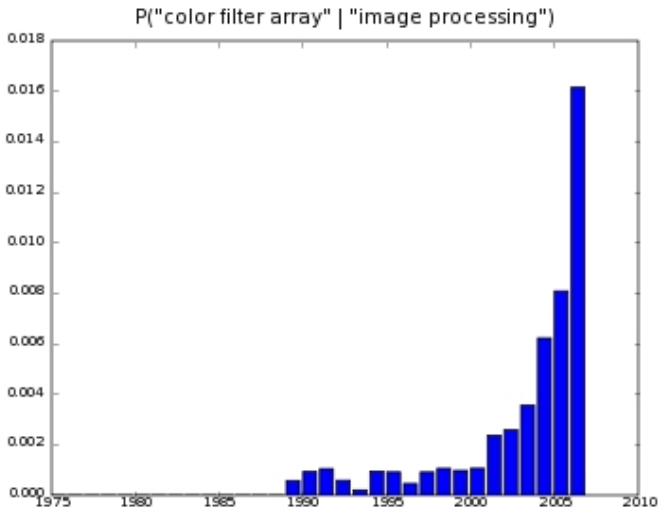


FIG. *3.7: Evolution de la probabilité conditionnelle de trouver l'expression "color filter array" dans les publications scientifiques contenant "image processing".*

Les nouveaux algorithmes publiés sont le plus souvent des petites variantes ou améliorations de quelques principes fondamentaux. Nous allons essayer de dégager dans la suite ces grands principes, comparer les méthodes existantes entre elles. La description se veut non exhaustive.

D'une manière générale, un algorithme de démosaïçage est basé sur un modèle exploitant les propriétés de corrélation spatiale des images (*i.e.* il vaut mieux interpoler le long d'un contour plutôt qu'à travers celui-ci) et sur un modèle exploitant les propriétés de corrélation spectrale des plans couleur (forte corrélation des hautes fréquences des trois plans) :

$$C_i(x,y) = f^{-1}(\sum_{u,v} w(u,v) f(CFA(u,v))) \tag{3.1}$$

où les C_i sont les canaux couleur (avec $i \in \{R, V, B\}$), w est le modèle spatial prenant en compte la corrélation intraplan, et f le modèle spectral prenant en compte la corrélation interplan. Sauf mention contraire, il est question dans la suite de ce chapitre de la mosaïque de Bayer. Les indices des pixels correspondent à la numérotation de la Figure 3.8.

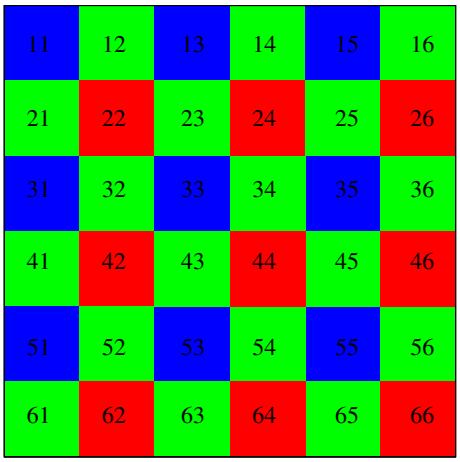


FIG. *3.8: Numérotation de la position des pixels dans la matrice de Bayer.*

2.1 Méthodes simples

Les méthodes de reconstruction les plus simples consistent à interpoler les trois plans couleur de manière indépendante (f est alors l'identité, dans l'Equation (3.1)), sans *a priori* sur le contenu de l'image. On peut citer la méthode par copie de pixel et la méthode par interpolation bilinéaire. Si ces méthodes sont très rapides en temps de calcul, elles présentent l'inconvénient de générer des artefacts de reconstruction gênants, sous forme de fausses couleurs (Figure 3.9).

(a) (b)

FIG. *3.9: Reconstruction par copie de pixel (a) et par interpolation bilinéaire (b). Ces méthodes génèrent beaucoup de fausses couleurs, du fait du repliement spectral.*

La méthode par copie de pixel est une simple interpolation au plus proche voisin. C'est cette interpolation que l'on utiliserait si l'on implémentait le modèle de couleur de De Valois *et al* [De Valois 93], décrit au chapitre précédent. En effet, ce modèle ne tient pas compte du sous-échantillonnage effectué par la mosaïque des cônes.

L'interpolation bilinéaire des pixels bleus et verts s'obtient de la manière suivante ; le pixel

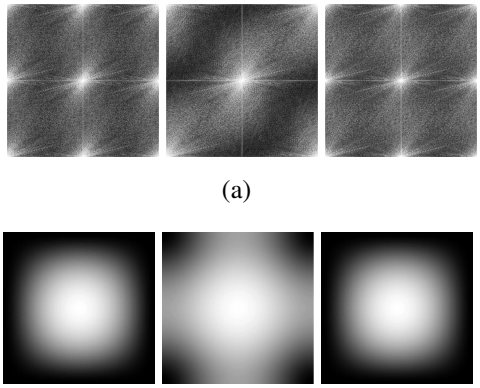
bleu et le pixel rouge jouant des rôles symétriques :

$$B_{22} = \frac{B_{11} + B_{13} + B_{31} + B_{33}}{4}$$
$$B_{12} = \frac{B_{11} + B_{13}}{2}$$
$$V_{22} = \frac{V_{12} + V_{21} + V_{23} + V_{32}}{4}$$

Les noyaux d'interpolation bilinéaires s'appliquant à la mosaïque de Bayer sont les suivants :

$$f_{R,B} = \frac{1}{4}\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix} \qquad f_V = \frac{1}{4}\begin{bmatrix} 0 & 1 & 0 \\ 1 & 4 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

Comme le pas d'échantillonnage est régulier, il est possible d'appliquer ces filtres par convolution. L'opération de filtrage effectuée permet de supprimer les réplicats apparaissant en hautes fréquences du fait du sous-échantillonnage. La position fréquentielle des réplicats dépend de l'arrangement spatial des filtres de couleur. Dans le cas de Bayer, les réplicats se positionnent à la fréquence $(0,5; 0,5)$ dans le plan de Fourier pour le plan couleur vert, et aux fréquences $\{(0,5; 0,5), (0; 0,5), (0,5; 0)\}$ pour les plans rouge et bleu (Figure 3.10(a)). Si l'on prend la transformée de Fourier des noyaux bilinéaires (Figure 3.10(b)), on remarque qu'ils atténuent effectivement ces fréquences. Un noyau d'interpolation dont le bagarit serait plus contraint, avec un gain plus proche de 1 dans la bande passante et plus proche de zéro en bande coupée permettrait d'améliorer dans une certaine mesure la qualité de l'image reconstruite. Cependant on est loin des performances que l'on peut obtenir en utilisant un modèle spectral et/ou spatial.



(a)

(b)

FIG. *3.10: Spectre d'amplitude de chacun des plans couleurs d'une image couleur échantillonnée selon le motif de Bayer (a) et réponse fréquentielle des noyaux interpolateurs bilinéaires correspondants (b).*

2.2 Les modèles spectraux

La qualité des images reconstruites peut être augmentée en s'appuyant sur la corrélation existant entre les trois plans couleur. La Figure 3.11 donne le schéma général de ce genre d'algorithme, avec dans cet exemple le canal vert pris comme canal d'intensité.

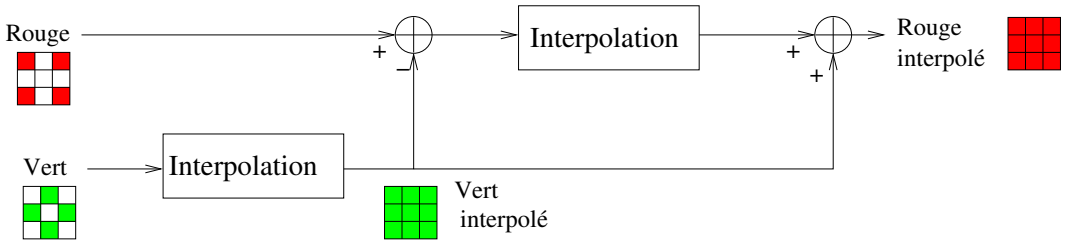


FIG. *3.11: Illustration d'une interpolation classique du canal rouge (l'interpolation du canal bleu est similaire). Le canal vert est interpolé de manière marginale. Puis il est utilisé par soustraction ou division pour l'interpolation du canal rouge.*

2.2.1 Méthodes basées sur la constance des teintes

Cok propose d'interpoler la "teinte", plutôt que chacune des couleurs prises séparément [Cok 87]. L'idée intuitive est que la teinte ne varie pas à la surface d'un objet. La méthode devrait donc diminuer l'apparition de fausses couleurs. La teinte est calculée comme le rapport entre Rouge et Vert (R/V) et Bleu et Vert (B/V). L'algorithme est alors le suivant. On commence par interpoler les pixels verts manquants par la méthode d'interpolation bilinéaire proposée précédemment. Puis on calcule pour les pixels rouges et bleus existants les teintes R/V et B/V. On interpole ensuite de manière bilinéaire ces teintes, puis on multiplie par les valeurs de V pour obtenir les valeurs de R et de B correspondantes. Par exemple, pour l'interpolation du pixel bleu à la position (2,2) on a :

$$B_{22} = \frac{1}{4} V_{22} \left(\frac{B_{11}}{V_{11}} + \frac{B_{13}}{V_{13}} + \frac{B_{31}}{V_{31}} + \frac{B_{33}}{V_{33}} \right)$$

Cet algorithme est fondamental à plusieurs titres. D'une part, il est réutilisé dans de nombreuses autres propositions ([Laroche 94, Hibbard 95, Kimmel 99]...). D'autre part, à partir de cette proposition, la plupart des algorithmes proposés ne vont plus utiliser de manière indépendante les canaux de couleurs, mais vont exploiter l'intercorrélation entre les canaux.

Pei et Tam [Pei 00, Pei 03] proposent d'interpoler non pas le rapport des couleurs, mais plutôt la différence. Ils justifient ce choix par deux raisons. La première est une raison de coût de calcul, la soustraction étant exécutée de manière plus rapide que la division par un processeur de calcul. La seconde est que ce concept se rapproche de la représentation intensité/teinte/saturation, utilisée par exemple dans le codage de télévision NTSC. Pei et Tam n'utilisent pas la corrélation entre canaux seulement pour le calcul de R et de B, mais également mais également pour l'interpolation du plan V en faisant une première estimation bilinéaire des plans R et B. Cette première estimation est ensuite remise à jour une fois le plan V calculé. Cette manière itérative de procéder, par estimation et mise à jour, est utilisée dans de nombreux algorithmes. L'algorithmes de Pei et Tam se déroule donc comme suit :

- initialisation bilinéaire de R et de B aux positions correspondants aux pixels V mesurés,
- interpolation bilinéaire du plan V en passant par la différence entre les valeurs de V connues et celles de R et de B initialisées,
- mise à jour de R et de B par interpolation bilinéaire en passant par la différence entre les valeurs V calculées et celles de R et de B mesurées.

Cet algorithme offre un très bon rapport entre la rapidité en temps de calcul - il n'utilise que

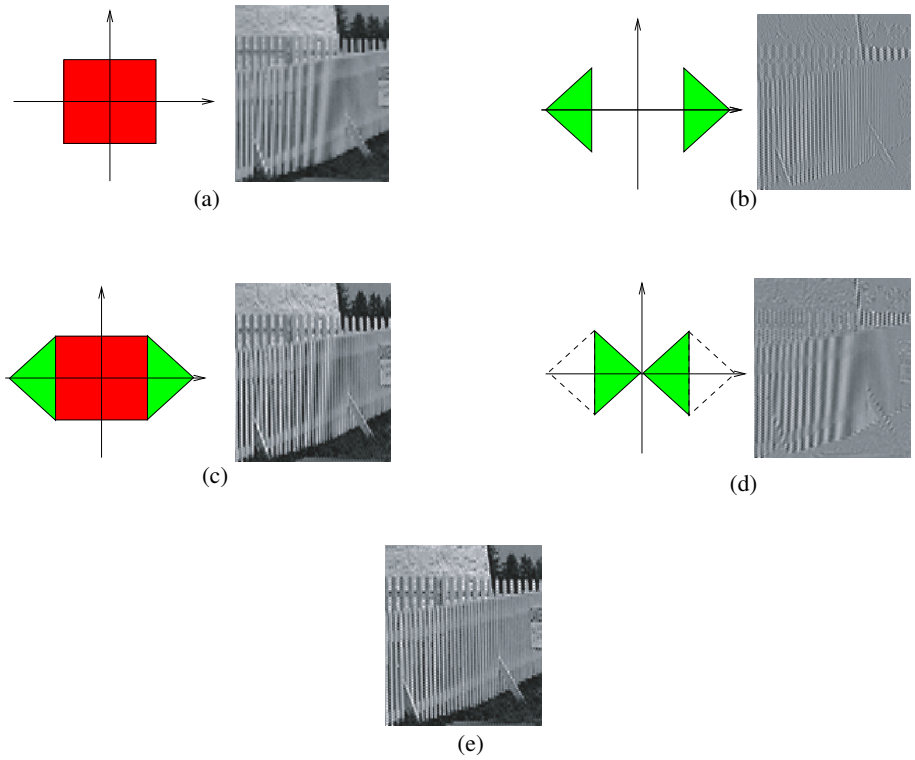


FIG. *3.12: Illustration de l'algorithme de Glotzbach et al. [Gunturk 05]. (a) Basses fréquences du canal rouge. (b) Hautes fréquences du canal vert. (c) Les hautes fréquences du canal vert sont recopiées sur le canal rouge. (d) Les hautes fréquences du canal vert sont modulées pour simuler le repliement spectral. (e) L'estimation du repliement spectral est soustraite au canal rouge.*

des interpolations bilinéaires - et la qualité des images obtenues. Il illustre le fait que faire des estimations/actualisations permet d'augmenter substantiellement le PSNR des images reconstruites.

Malvar propose également un algorithme linéaire basé sur un modèle de constance des teintes [Malvar 04]. Son hypothèse est que les contours d'une image contiennent davantage d'intensité que de chrominance. L'auteur traduit cette hypothèse ainsi : si un gradient est détecté sur les valeurs mesurées R, alors il y a de fortes chances que ce gradient soit achromatique et qu'il se retrouve, par conséquent, également sur les valeurs V et B. Ainsi, à une position R, le gradient calculé sur des valeurs R mesurées, sera ajouté au résultat de l'interpolation bilinéaire de V et de B. Voici par exemple le noyau utilisé pour calculer une valeur V sur un pixel R :

$$\frac{1}{8}\begin{bmatrix} & & -1 & & \\ & & 2 & & \\ -1 & 2 & 4 & 2 & -1 \\ & & 2 & & \\ & & -1 & & \end{bmatrix}$$

Cet algorithme est du même ordre de complexité que celui proposé par Pei et Tam, mais donne toutefois de moins bons résultats en terme de PSNR.

2.2.2 Algorithmes par substitution de fréquences

Les algorithmes de Glotzbach *et al.* et Gunturk *et al.* sont deux stratégies d'implémentation d'une même idée [Glotzbach 01, Gunturk 02]. Ils sont basés sur la constatation que les trois canaux couleurs sont très corrélés au niveau des hautes fréquences spatiales. Cette hypothèse

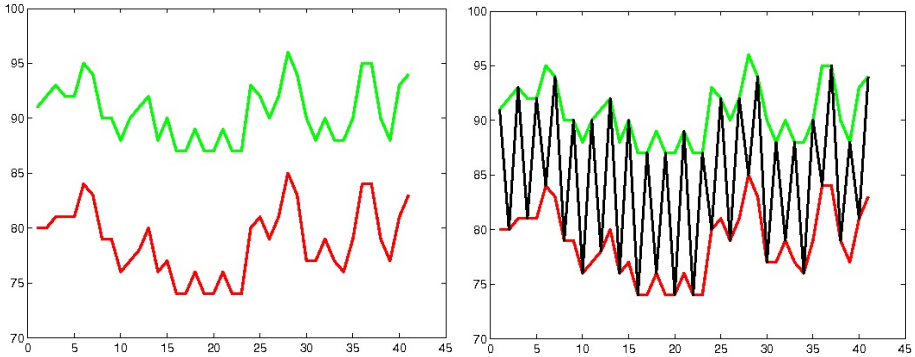


FIG. *3.13: Coupes des plans R et V (en rouge et vert, respectivement)d'une image couleur et de la même image spatialement multiplexée (en noir), telle qu'acquise à travers la mosaïque de Bayer. Le multiplexage spatial fait apparaître des hautes fréquences véhiculant l'information chromatique.*

est équivalente à celle de la constance des teintes, puisque cela revient à dire que la différence entre canaux couleur est un signal basse-fréquence. Comme les hautes fréquences du canal vert sont plus faciles à reconstruire (les échantillons verts représentent 50% de la mosaïque de Bayer), l'idée est de substituer les hautes fréquences du canal V à celles des canaux R et B, et de les utiliser pour annuler le repliement spectral.

Dans les deux cas, l'algorithme commence par une première interpolation des trois canaux. Puis les hautes fréquences du canal vert sont projetées sur celles des canaux rouge et bleu. Vient ensuite l'étape d'annulation du repliement spectral. C'est à cette étape que les deux algorithmes diffèrent. Glotzbach *et al.* procèdent par modulation des hautes fréquences du canal vert, afin d'estimer la composante repliée, et par soustraction de cette dernière aux canaux rouge et bleu (Figure3.12). Gunturk *et al.*, quant à eux, régularisent l'image obtenue, par projections alternées sur des ensembles convexes. Par dualité entre le domaine fréquentiel et le domaine spatial, la modification d'une composante fréquentielle d'une image a une incidence sur l'ensemble des pixels de cette image, et *vice versa*. Ainsi, les auteurs définissent deux ensembles convexes :

- l'ensemble des images en niveaux de gris dont les hautes fréquences sont égales à celles du canal vert,
- l'ensemble des images en niveaux de gris possédant les mêmes échantillons mesurés que le canal rouge (resp. bleu)

Les projections alternées sur ces deux ensembles permettent de converger vers la solution minimisant la distance entre une image et son projeté.

L'algorithme par projections alternées a longtemps été considéré comme l'algorithme de référence du point de vue de la qualité des images reconstruites. Aujourd'hui, d'autres algorithmes obtiennent une meilleur qualité visuelle pour une moindre complexité.

2.2.3 Algorithmes par sélection de fréquences

L'algorithme par "sélection de fréquence" proposé par Alleysson *et al.* est basée sur un modèle d'échantillonnage spatio-chromatique. Nous détaillons ce modèle, avant de parler de l'algorithme propement dit et de ses artefacts de reconstruction.

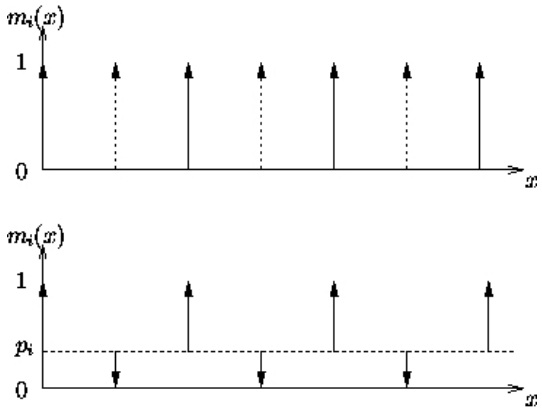


FIG. *3.14: Le sous-échantillonnage peut être décomposé en une partie constante définie pour chaque pixel et une partie fluctuante.*

Le modèle intensité et oppositions de couleurs pour les images en mosaïque

Comme nous l'avons dit plus haut, le processus d'acquisition d'une image au travers d'une mosaïque de filtres couleur peut être vu comme un sous-échantillonnage suivi d'un multiplexage spatial d'une image couleur. Comme on peut le voir sur la Figure 3.13 représentant la coupe d'une image avant et après le passage à travers la mosaïque de Bayer, ce multiplexage spatial introduit une composante haute fréquence. L'amplitude de cette composante dépend uniquement de la différence entre les signaux des canaux couleur. Elle code donc une information purement chromatique, autrement dit, la chrominance.

De manière plus formelle, l'échantillonnage à une seule couleur par pixel peut s'écrire comme la somme des produits respectifs des trois plans couleurs avec des fonctions d'échantillonnage m_i :

$$I_m(x,y) = \sum_{i \in \{R,V,B\}} C_i(x,y) m_i(x,y)$$

où la fonction $m_i(x,y)$ vaut 1 si le filtre de type i est présent sur le pixel (x,y), et vaut 0 dans le cas contraire. Ces fonctions sont mutuellement exclusives car un seul filtre de couleur peut être présent à chaque position spatiale. Dans le cas de la matrice de Bayer, ces fonctions sont [Alleysson 05] :

$$\begin{cases} m_R(x,y) = \frac{1}{4}(1+\cos(\pi x))(1+\cos(\pi y)) \\ m_V(x,y) = \frac{1}{2}(1-\cos(\pi x)\cos(\pi y)) \\ m_B(x,y) = \frac{1}{4}(1-\cos(\pi x))(1-\cos(\pi y)) \end{cases} \tag{3.2}$$

Décomposons chacune d'entre elles en une partie constante p_i correspondant à la probabilité d'avoir un récepteur de type i, et une partie fluctuante $\tilde{m}_i$ (Figure 3.14) :

$$m_i(x,y) = p_i + \tilde{m}_i(x,y) \tag{3.3}$$

Alors l'image mesurée par le capteur devient :

$$I_m(x,y) = \underbrace{\sum_i p_i C_i(x,y)}_{\varphi(x,y)} + \underbrace{\sum_i C_i(x,y)\tilde{m}_i(x,y)}_{\psi(x,y)} \tag{3.4}$$

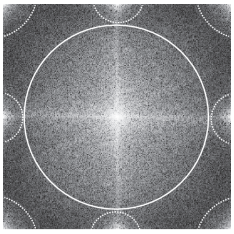


FIG. *3.15: Spectre d'amplitude d'une image en mosaïque. L'information d'intensité est localisée au centre, tandis que l'information de chrominance est modulée sur les côtés du spectre.*

Le premier terme est une combinaison positive des trois canaux couleur, c'est donc un terme d'intensité, que nous appellerons $\varphi(x, y)$. Comme les p_i sont constants, cette intensité n'est pas affectée pour le sous-échantillonnage. C'est exactement celle que l'on aurait obtenue par combinaison linéaire des plans couleur C_i pondérés par les p_i, *avant le sous-échantillonnage*. Si on fixe un point (x, y), correspondant au filtre de couleur j, le second terme s'écrit (en remplaçant $\tilde{m}_i$ par $1 - p_i$) :

$$\psi(x, y) = (1 - p_j)C_j(x, y) - \sum_{i \neq j} p_i C_i(x, y) \tag{3.5}$$

$\psi(x, y)$ est donc un signal d'opposition de couleurs, appelé aussi signal de chrominance. Comme il y a trois filtres de couleur, cela définit trois signaux d'oppositions de couleurs. Cependant, un seul signal d'opposition de couleurs est effectivement défini à chaque position spatiale : celui correspondant au type de filtre de couleur présent à la position spatiale considérée. Il y a donc un sous-échantillonnage des oppositions de couleurs dans une image en mosaïque, tandis que l'intensité n'est pas affectée.

Une manière de comprendre l'effet des fonctions $\tilde{m}_i$ sur l'information de chrominance de l'image couleur est de considérer que ces fonctions modulent en amplitude l'information de chrominance. Or ces fonctions de modulation ont une moyenne spatiale nulle. Cela signifie que ce sont des porteuses de fréquences non nulles. Par voie de conséquence, l'information de chrominance ψ est un signal modulé à des fréquences non nulles.

Ainsi, et c'est le concept fondamental qui découle de ce modèle d'échantillonnage, le *multiplexage spatial* des composantes couleur d'une image se traduit par un *multiplexage fréquentiel* de la composante d'intensité et des composantes d'oppositions de couleurs.

On peut calculer explicitement à quelles fréquences les fonctions m_i modulent le signal. Dans l'espace de Fourier, les fonctions d'échantillonnage deviennent [Alleysson 05] :

$$\begin{cases} M_R(x, y) = \dfrac{1}{4} \displaystyle\sum_{r=-1}^{1} \dfrac{1}{1 + |r|} \delta(f_x - r/2) \sum_{s=-1}^{1} \dfrac{1}{1 + |s|} \delta(f_y - s/2) \\ M_V(x, y) = \dfrac{1}{2} \delta(f_x)\delta(f_y) \displaystyle\sum_{\substack{r=-1 \\ r \neq 0}}^{1} \delta(f_x - r/2) \sum_{\substack{s=-1 \\ s \neq 0}}^{1} \delta(f_y - s/2) \\ M_B(x, y) = \dfrac{1}{4} \displaystyle\sum_{r=-1}^{1} \dfrac{(-1)^r}{1 + |r|} \delta(f_x - r/2) \sum_{s=-1}^{1} \dfrac{(-1)^s}{1 + |s|} \delta(f_y - s/2) S \end{cases}$$

Il y a donc modulation de l'information de chrominance aux fréquences (0,0.5), (0.5,0) et (0.5,0.5), c'est-à-dire dans les coins et au milieu des bords du spectre, comme le montre la

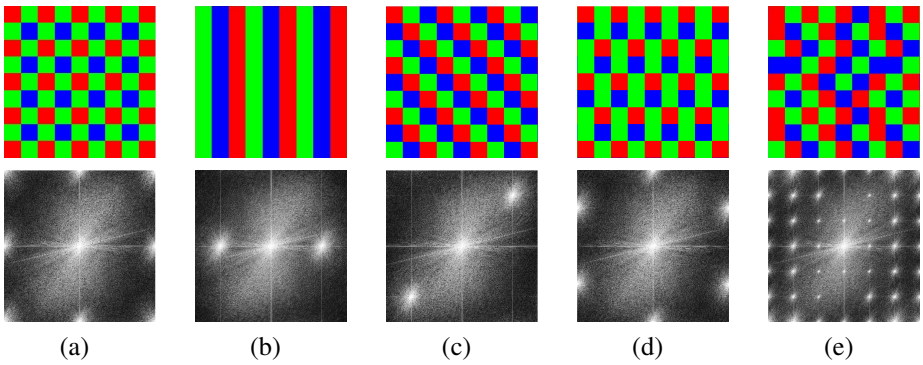


FIG. *3.16: Mosaïques et spectres d'amplitude d'une image correspondante : (a) mosaïque de Bayer ; (b-c) mosaïques à bandes verticales et diagonales ; (d) mosaïque proposée par Lukac [Lukac 05b] ; (e) mosaïque irrégulière.*

Figure 3.15. La Figure 3.16 montre les spectres d'amplitude de plusieurs types de mosaïques. On peut voir que la position des porteuses dans le plan de Fourier dépend de la configuration spatiale de la mosaïque.

Ce modèle d'échantillonnage permet certes de se représenter les informations d'intensité et de chrominance contenues dans une image à une seule couleur par pixel. Mais quel est l'intérêt de cette représentation pour le démosaïçage ?

Pour saisir l'intérêt d'une telle représentation, il faut comprendre quelles sont les limites du démosaïçage par interpolation marginale de chaque canal couleur (cas des "méthodes simples", paragraphe 2.1). Ses limites sont imposées par le théorème de Shannon appliqué à chaque plan couleur. Par exemple, pour les plans rouge et bleu, la mosaïque de Bayer impose une fréquence de Nyquist à 0.25 dans les directions horizontale et verticale. Cela provient du fait que le signal est modulé aux fréquences 0.5 dans ces deux directions.

Si on considère la représentation intensité-chrominance d'une image en mosaïque, le signal modulé n'est pas le canal couleur lui-même, mais un signal de chrominance. Or, un signal de chrominance possède une largeur de bande bien inférieure à celle d'un canal couleur. Cela a pour conséquence de repousser la fréquence de Nyquist vers une fréquence plus élevée, allouant ainsi une plus grande largeur de bande à l'intensité que celle qui aurait été imposée par les fréquences de Nyquist marginales de chacun des plans couleur sous-échantillonnés.

Il est remarquable de constater que la mosaïque de Bayer module l'information aux plus hautes fréquences autorisées par le pas d'échantillonnage, maximisant ainsi la largeur de bande allouée à l'intensité. Nous verrons au Chapitre 5 que cet arrangement n'est en fait *pas* optimal. Certains arrangements permettent d'éviter la modulation aux fréquences de Nyquist horizontale et verticale, fréquences plus critiques que celles se trouvant sur les coins du spectre.

Enfin, cette représentation de l'image en mosaïque, comme la somme d'une composante d'intensité en bande de base et de composantes de chrominances modulées en hautes fréquences, n'est pas sans rappeler les systèmes de télévision, le format NTSC par exemple [Coulombe 96]. C'est en fait le même principe, à la différence près qu'il est possible, dans le cas d'un système de télévision, d'appliquer des filtres anti-repliement à l'intensité et à la chrominance avant le codage.

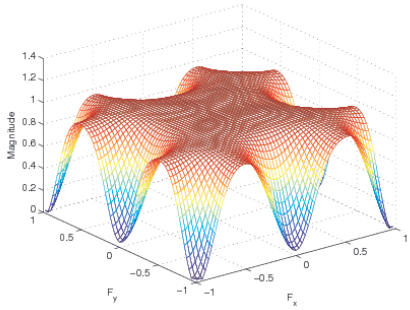


FIG. *3.17: Réponse fréquentielle linéaire du filtre utilisé pour l'estimation de l'intensité.*

Algorithme inspiré par le modèle

Alleysson *et al.* dérivent un algorithme de leur modèle d'échantillonnage. Le principe est de faire une interpolation basée sur le principe de la constance des teintes calculées par différence, mais en remplaçant le canal vert par le signal d'intensité défini comme la combinaison linéaire des plans couleur. L'avantage est que l'intensité est disponible à chaque pixel par filtrage passe-bas, tandis que le signal vert doit être interpolé. Le noyau du filtre utilisé est le suivant :

$$\frac{1}{128}\begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & -2 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 2 & 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & -1 & 1 & -5 & 3 & -9 & 3 & -5 & 1 & -1 & 0 \\ 1 & 0 & 2 & 3 & 1 & 7 & 1 & 3 & 2 & 0 & 1 \\ 0 & -2 & 1 & -9 & 7 & 104 & 7 & -9 & 1 & -2 & 0 \\ 1 & 0 & 2 & 3 & 1 & 7 & 1 & 3 & 2 & 0 & 1 \\ 0 & -1 & 1 & -5 & 3 & -9 & 3 & -5 & 1 & -1 & 0 \\ 0 & 0 & 1 & 1 & 2 & 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & -2 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Sa réponse fréquentielle est donnée à la Figure 3.17. Comme on peut le voir, ce filtre atténue les régions du spectre où se trouvent les porteuses de chrominance. Ce filtre a été construit à l'aide de gaussiennes dans le domaine fréquentiel, dont les écarts-type ont été choisi de manière à optimiser le PSNR obtenu sur une base d'image.

Une fois l'intensité estimée, la chrominance multiplexée s'obtient en retranchant l'intensité à l'image en mosaïque. Elle est ensuite démultiplexée et interpolée par filtrage bilinéaire.

Alternativement, on peut estimer la chrominance multiplexée par filtrage passe-haut. Utiliser deux filtrages pour l'estimation de l'intensité et de la chrominance multiplexée permet de préserver une zone de garde fréquentielle entre les deux signaux, afin d'éviter les zones où ces signaux se recouvrent.

Artefacts de reconstruction

Les images reconstruites par démosaïçage peuvent souffrir de quatre types d'artefacts que sont : le lissage excessif, l'effet de grille, l'effet d'aquarelle et les fausses couleurs. Ces artefacts sont illustrés à la Figure 3.18. Le modèle d'intensité-chrominance permet d'expliquer précisément l'origine de ces artefacts.

Lorsque l'intensité possède des composantes fréquentielles plus hautes que la fréquence de coupure du filtre d'intensité, alors ces composantes sont perdues. Il en résulte un lissage de l'image (Figure 3.18(a)). Ces composantes hautes fréquences peuvent être interprétées comme étant de la chrominance. Il en résulte alors des fausses couleurs (Figure 3.18(d)).

De manière duale, lorsque la chrominance multiplexée possède des composantes fréquentielles plus basses que la fréquence de coupure du filtre de chrominance (c'est le cas au niveau d'un contour chromatique franc, par exemple), alors ces composantes sont perdues. Il en résulte un effet d'aquarelle (Figure 3.18(c)). Ces composantes de chrominance peuvent être interprétées comme étant de l'intensité haute fréquence. Il en résulte alors un effet de grille (Figure 3.18(b)).

Dans le cas où la chrominance est calculée par différence entre l'image en mosaïque et l'image d'intensité, le lissage excessif est alors systématiquement associé à des fausses couleurs. De même, l'effet de grille est alors systématiquement associé à l'effet d'aquarelle. Ce dernier artefact est cependant moins perceptible que les autres.

Remarquons que l'on peut retrouver le même résultat que celui obtenu par interpolation bilinéaire en choisissant un filtre d'intensité suffisamment passe-bas. Il y a dans ce cas beaucoup de fausses couleurs et de lissage, mais très peu d'effet de grille et d'aquarelle.

Dubois et Lian *et al.* proposent chacun un algorithme adaptatif basé sur ce modèle d'échantillonnage. Dubois utilise la redondance des porteuses de chrominance dans le domaine fréquentiel, la même information étant modulée aux fréquences (0,0.5) et (0.5,0). Il sélectionne, à chaque pixel, la composante dont le repliement spectral avec l'intensité est minimal [Dubois 05]. Cette méthode permet de réduire significativement les fausses couleurs. Lian *et al.* utilisent un modèle spatial basé sur les gradients lors de l'interpolation de la chrominance [Lian 05]. Cela permet de réduire l'effet de grille.

2.3 Les modèles spatiaux

Nous venons de décrire différentes manières de mettre à profit la corrélation spectrale entre les trois canaux couleur. En plus d'utiliser un modèle de corrélation spectrale, certains algorithmes emploient également un modèle spatial d'interpolation, c'est-à-dire qu'ils prennent en compte la structure locale de l'image par estimation des contours. Ces méthodes peuvent être divisées en deux groupes, selon que la direction d'interpolation est choisie de manière binaire ou de manière pondérée.

2.3.1 Interpolations directionnelles par *prise de décision* sur les contours

Hibbard [Hibbard 95] utilise le gradient horizontal et vertical autour du pixel manquant, calculé à partir des pixels voisins existants. L'idée est de réaliser une interpolation le long des contours et d'éviter une interpolation à travers eux. Par exemple, l'interpolation du pixel vert à la position (4,4), est la suivante :

1. Calcul du gradient horizontal et vertical $\Delta H = |G_{43} - G_{45}|$ et $\Delta V = |G_{34} - G_{54}|$
2. Si $\Delta H > \Delta V$

$$G_{44} = (G_{34} + G_{54})/2$$

(a) Lissage excessif (b) Effet de grille

(c) Aquarelle (d) Fausses couleurs

FIG. *3.18: Illustrations des quatre artefacts de reconstruction possibles.*

3. Sinon, si $\Delta H < \Delta V$

$$G_{44} = (G_{43} + G_{45})/2$$

4. Sinon,

$$G_{44} = \frac{1}{4}(G_{43} + G_{45} + G_{34} + G_{54})$$

Il est important de noter que les méthodes utilisant le gradient sont très utilisées dans le démosaïçage. Ces méthodes sont judicieuses puisqu'à chaque pixel manquant dans la matrice de Bayer, le gradient suivant les deux directions est parfaitement connu à partir des pixels existants. Une extension de cette méthode est proposée dans [Laroche 94], où le gradient au pixel vert manquant est estimé à partir d'un voisinage plus large de pixels rouges ou bleus suivant la couleur du pixel présent à la position verte manquante. Par exemple, pour l'interpolation du pixel vert à la position (4,4) il suffit de changer la façon dont le gradient est calculé :
$\Delta H = |(R_{42} + R_{46})/2 - R_{44}|$ et $\Delta V = |(R_{24} + R_{64})/2 - R_{44}|$

Hamilton et Adams [Hamilton 97] proposent d'utiliser d'autres estimateurs pour la classification et l'interpolation des pixels. Dans cette implémentation, les auteurs utilisent un calcul de gradients sur les pixels verts et de laplaciens sur les pixels rouges ou bleus. Leur méthode pour l'interpolation du pixel vert G_{44} est basée sur l'algorithme suivant :

1. Calculer $\Delta H = |G_{43} - G_{45}| + |2R_{44} - R_{42} - R_{46}|$ et $\Delta V = |G_{34} - G_{54}| + |2R_{44} - R_{24} - R_{64}|$
2. Si $\Delta H > \Delta V$
$$G_{44} = \frac{G_{34} + G_{54}}{2} + \frac{2R_{44} - R_{24} - R_{64}}{4}$$
3. Sinon, si $\Delta H < \Delta V$
$$G_{44} = \frac{G_{43} + G_{45}}{2} + \frac{2R_{44} - R_{42} - R_{46}}{4}$$
4. Sinon,
$$G_{44} = \frac{G_{43} + G_{45} + G_{34} + G_{54}}{4} + \frac{4R_{44} - R_{42} - R_{46} - R_{24} - R_{64}}{8}$$

Les auteurs proposent également une méthode directionnelle pour l'interpolation des pixels rouges et bleus. Cette méthode d'interpolation est utilisée pour l'initialisation du canal vert dans l'algorithme par projections alternées [Gunturk 02] présentée dans le paragraphe "Algorithmes par substitution de fréquences".

Adams et Hamilton [Adams 96, Adams 97] ont ensuite proposé plusieurs méthodes plus sophistiquées, à partir de celle-ci, proposant alors plusieurs autres types de classifications et d'interpolateurs. Ces méthodes sont basées sur l'intégration des caractéristiques de l'algorithme de constance de teintes.

L'algorithme proposé par Hirakawa *et al.* [Hirakawa 05] est un autre exemple où la direction d'interpolation est choisie sur un critère basé sur contenu spatial d'une image. Le critère retenu ici est un critère d'homogénéité entre l'intensité et la chrominance sur un voisinage.

Ces interpolations directionnelles permettent de réduire significativement l'apparition de fausses couleurs (lorsque la prise de décisions est correcte).

Lorsque la direction d'interpolation est basée sur une prise de décision, il se peut que la décision soit mauvaise. L'interpolation se fait alors, par exemple, en travers d'un contour, générant ainsi des fausses couleurs voyantes. Une étape de post-traitements est alors souvent utilisée. Par exemple, le filtrage médian se révèle être efficace dans ce cas [Hirakawa 05].

2.3.2 Interpolation linéaire *pondérée* par les contours

Plutôt que de choisir une direction d'interpolation, une méthode alternative consiste à effectuer une interpolation par combinaison linéaire des pixels voisins, pondérée par un critère de vraisemblance de la présence d'un contour à chaque pixel [Kimmel 99, Lian 05, Lukac 05b] :

$$G_{xy} = \sum_{i,j \in \xi} w'_{ij} I_m(x, y)$$

où ξ est le voisinage du pixel (x, y) et où w'_{ij} est une normalisation du poids $w(i, j)$ indiquant la vraisemblance d'un contour au pixel (i, j) :

$$w'_{xy} = \frac{w_{ij}}{\sum_{j,l \in \xi} w_{kl}}$$

Par exemple, pour l'interpolation de l'intensité à la position $(4, 4)$, Lian *et al.* calculent le poids du pixel d'intensité immédiatement à sa gauche (position comme l'inverse d'un gradient :

$$w_{43} = \frac{1}{1 + |R_{44} - R_{42}| + |L_{45} - L_{43}|}$$

les valeurs d'intensité étant calculées sur les positions verte en utilisant le modèle d'Alleysson. Si le gradient est élevé, cela signifie que le pixel $(4, 3)$ est situé sur un contour, alors $w_{43})$ est faible et ce pixel sera peu pris en compte pour l'interpolation.

Lukac *et al.* calculent ce poids comme étant fonction de la somme de tous les gradients possible entre le pixel (i, j) et ses voisins de même couleur [Lukac 05b]. Dans la littérature, cet algorithme se distingue des autres du fait que ce soit l'un des rares, à notre connaissance, qui puisse être appliqué à n'importe quel arrangement de la mosaïque (à condition que cet arrangement soit connu). Cependant, sa complexité est bien plus élevée dans le cas d'une mosaïque irrégulière que dans le cas d'une mosaïque régulière telle que celle de Bayer. En effet, un pixel est interpolé seulement si son voisinage comporte suffisamment d'échantillons disponibles. Dans le cas contraire, il sera interpolé lors d'une prochaine passe, lorsque la condition sera remplie. Par conséquent, l'algorithme nécessite un plus grand nombre de passes si la densité en échantillons est irrégulière.

Ces méthodes par pondération permettent de réduire significativement l'effet de grille. Elles ont aussi l'avantage de se passer d'instructions conditionnelles, alors que celles-ci sont nécessaires pour les algorithmes basés sur un choix de direction d'interpolation.

La mosaïque de Bayer implique nécessairement un calcul itératif dès lors que l'on souhaite faire une interpolation orientée contour des canaux rouge et bleu. Pour justifier cela, considérons le cas de l'interpolation d'un pixel rouge R_{33} à la position d'un pixel de couleur bleue. Imaginons que l'on souhaite effectuer une interpolation directionnelle. Le problème est qu'il n'y a pas de pixels rouges dans la direction horizontale, ni dans la direction verticale. Les plus proches voisins sont uniquement en diagonale. Il faut donc faire une première estimation des quatre plus proches voisins. Là aussi, ces voisins posent problème, car ils ont chacun deux voisins rouges disponibles, mais dans une seule direction (soit horizontale, soit verticale). Ainsi

methode	R	G	B	$\Delta E_{S-CIELAB}$	Complexité
bilineaire	29.22 (±3.32)	33,04 (±3.25)	29.27 (±3.33)	1.79 (±0.81)	6
Rapport de teintes	31.57 (±2.88)	33.23 (±3.16)	32.39 (±3.23)	1.81 (0.76)	12
Diff. de teintes	32.85 (±3.30)	33.23 (±3.16)	32.57 (±3.21)	1.77 (±0.83)	9
[Pei 03]	37.21 (±3.12)	38.87 (±2.99)	36.06 (±2.99)	1.12 (±0.44)	19
[Malvar 04]	35.36 (±3.34)	38.87 (±2.99)	34.15 (±3.12)	1.25 (±0.53)	22
[Alleysson 05]	37.83 (±2.44)	40.74 (±2.26)	36.48 (±2.38)	1.10 (±0.34)	77
[Gunturk 02] 3 iter.	38.40 (±2.73)	41.37 (±2.40)	37.46 (±2.59)	0.99 (±0.35)	405
[Gunturk 02] 8 iter.	39.29 (±2.54)	41.37 (±2.40)	37.81 (±2.47)	0.96 (±0.33)	885
[Lian 05]	38.77 (±2.59)	42.12 (±2.79)	38.62 (±2.88)	0.82 (±0.31)	63
[Dubois 05]	38.81 (±2.50)	42.82 (±2.50)	38.62 (±2.69)	0.83 (±0.30)	2274
[Li 05] 1 thr.	38.37 (±2.45)	41.77 (±2.40)	38.44 (±2.67)	0.93 (±0.33)	77
[Li 05] 2 thr.	38.07 (±2.45)	41.37 (±2.40)	38.12 (±2.65)	0.96 (±0.35)	117
[Hirakawa 05]	38.02 (±3.24)	39.59 (±3.21)	36.76 (±3.06)	0.90 (±0.37)	161

TAB. *3.1: PSNR moyens (dB) et écarts-type pour les plans R, V B ; ΔE dans l'espace S-CIELAB ; et complexité algorithmique en terme de nombre d'opérations par pixel, pour différents algorithmes de la littérature.*

l'estimation des plus proches voisins ne peut pas être adaptative, elle est forcément linéaire. Cela signifie qu'il faudra effectuer une (ou plusieurs) étape(s) de raffinement de ces voisins. Une possibilité est donc de suivre l'algorithme suivant :

1. estimation des 4 plus proches voisins, par interpolation linéaire. Par exemple, pour le pixel R_{32} :
$$R_{32} = \frac{R_{22} + R_{42}}{2}$$
2. interpolation adaptative du pixel R_{33} en utilisant les voisins horizontaux et verticaux estimés à l'étape 1 :
$$R_{33} = w'_{23}R_{23} + w'_{34}R_{34} + w'_{43}R_{43} + w'_{32}R_{32}$$
3. mise-à-jour des quatre voisins, par interpolation adaptative. Par exemple, pour le pixel R_{32} :
$$R_{32} = w'_{22}R_{22} + w'_{33}R_{33} + w'_{42}R_{42} + w'_{31}R_{31}$$
4. retourner à l'étape 2.

2.4 Evaluation des performances des algorithmes

Les valeurs de PSNR moyennées sur 23 images de la base Kodak (présentée en annexe), calculées entre les images reconstruites et les image servant à la simulation de l'acquisition à travers la mosaïque de Bayer, sont consignées dans la Table 3.1, pour différents algorithmes de la littérature. Ces algorithmes ont été implémentés à l'aide des articles cités dans la table. Sont consignées également dans la table, l'erreur ΔE calculée dans l'espace S-CIELab, ainsi que

la complexité algorithmique. La complexité algorithmique est calculée en nombre d'opérations par pixel. Chaque instruction (multiplication, division...) est comptée comme une opération, cependant la multiplication par un coefficient et l'addition du résultat sont comptées comme une seule opération, à l'image de l'opération MAC (pour Multiplication/Accumulation) sur les processeurs de signaux numériques. Ainsi le coût d'une convolution correspond au nombre de coefficients du noyau. L'erreur ΔE S-CIELab est une extension spatiale de la métrique ΔE CIEL*a*b, métrique incluant des propriétés du système visuel humain.

Il est clair que ces critères mathématiques sont de moins bons évaluateurs de la qualité que l'oeil humain. Cependant une campagne d'évaluation subjective, à l'aide de sujets humain demande beaucoup de temps. Nous avons donc opté pour ces critères mathématiques, en gardant à l'esprit que la qualité visuelle est importante, comme il est classique de procéder dans la communauté de demosaïçage.

La moitié supérieure de la Table 3.1 (jusqu'à la ligne [Alleysson 05]) renseigne des algorithmes qui, au plus, utilise un modèle spectral. La moitié inférieure renseigne des algorithme prenant en compte à la fois les corrélations spectrale et spatiale. Nous pouvons tirer de cette étude les enseignements suivant :

- les algorithmes adaptatifs donnent les meilleurs performances. Ils sont cependant globalement plus coûteux en temps de calcul. L'algorithme de Lian *et al.*, basé sur le modèle spectral d'Alleysson *et al.*, réalise un très bon compromis qualité/efficacité.
- L'algorithme d'Alleysson *et al.* donne les meilleurs résultats parmi les algorithmes non adaptatifs. Malgré sa simplicité, il est toutefois le plus coûteux. Sa complexité est due à la taille du filtre utilisé (un noyaux 9×9). Son point faible est l'effet de grille présent le long des contours.

3 Conclusion

Les caméras numériques mono-capteur sont confrontées à la même problématique que le système visuel humain : elles ne disposent que d'une composante couleur par position spatiale. Une étape d'interpolation des trois canaux couleurs, le démosaïçage, est nécessaire à la visualisation de l'image acquise par le capteur. On peut faire là encore une analogie entre la reconstruction d'une image couleur dans une caméra numérique et la construction des trois voies perceptuelles du système visuel.

Les algorithmes de démosaïçage mettent à profit deux propriétés des images numériques couleur pour améliorer la qualité des images reconstruites. L'une est la corrélation spatiale entre les pixels. Cette propriété est exploitée par des interpolations directionnelles déterminées par des calculs de gradient. La seconde est la corrélation spectrale des plans couleur. L'idée est que les signaux d'oppositions de couleurs sont de largeur de bande plus faible que les plans couleur eux-même. Ils sont donc plus "faciles" à interpoler. La raison pour laquelle le système visuel passe par l'espace intensité-chrominance est donc sans doute une raison écologique, dictée à la fois par la théorie de l'échantillonnage et par la statistique des images naturelles.

Tous les algorithmes, hormis celui proposé par Alleysson *et al*, utilisent le canal vert comme canal d'intensité pour calculer les oppositions de couleur. Les filtres de couleur verte étant

majoritaires dans la matrice de Bayer, ce choix est défendable. Cependant, que faire si l'on considère une matrice dans laquelle le vert n'est pas sur-représenté par rapport au bleu et au rouge ? C'est le cas par exemple de la mosaïque à bandes. Le canal vert constitue-t-il toujours un choix judicieux de canal d'intensité ? Il est clair que non, étant donné que sa résolution est alors très basse. En revanche, le choix d'une intensité définie comme une moyenne pondérée des trois canaux permet d'avoir un signal de bien meilleure résolution, quelles que soient les proportions des photorécepteurs.

Il ressort donc de ce chapitre que dans le domaine des caméras numériques, l'espace d'intensité et d'oppositions de couleurs est utilisé pour le traitement du signal sous forme de mosaïque. Cependant, à la différence du système visuel humain où l'intensité est définie comme une combinaison des informations issues des cônes L, M et peut-être S, les traiteurs d'image du domaine du démosaïçage définissent l'intensité comme étant le canal vert. Ainsi, la méthode d'Alleysson *et al.* est la plus proche de la biologie de la couleur. Réciproquement, l'implémentation du modèle d'Alleysson sur des caméras numériques permet de valider la démarche et d'en cerner les limites.

Néanmoins, l'algorithme d'Alleysson *et al.* n'est valable, en l'état, que pour des mosaïques régulières, car les porteuses de chrominance sont alors bien localisées dans le spectre de Fourier. Nous souhaitons étendre cet algorithme à une mosaïque quelconque, en particulier à une mosaïque irrégulière, comme c'est le cas dans la rétine. Nous faisons donc l'étude, au chapitre suivant, des propriétés et des méthodes d'interpolation des signaux échantillonnés irrégulièrement.

Chapitre 4

Echantillonnage irrégulier et interpolation

L'OBJECTIF de cette thèse étant de faire un modèle d'interpolation des couleurs adapté à toute forme de mosaïque, et en particulier aux mosaïques irrégulières, nous nous intéressons dans ce chapitre aux propriétés de l'échantillonnage irrégulier et à certains algorithmes de reconstruction de signaux échantillonnés irrégulièrement.

Nous verrons que l'échantillonnage irrégulier possède la propriété de convertir le phénomène de Moiré en bruit. Le phénomène de Moiré est une manifestation cohérente, très visible, du repliement spectral dans le cas d'un échantillonnage régulier. Dans le cas irrégulier, le repliement spectral se manifeste sous une forme déstructurée, qui ne véhicule pas d'information potentiellement trompeuse.

Remarque préliminaire : nous considérons des grilles d'échantillonnage irrégulières, ou *pseudo-aléatoires*, par analogie avec les fonctions de tirage $rand()$ d'un nombre pseudo-aléatoire en informatique. Dans ce domaine, une fonction $rand()$ prend en argument un entier, appelé *graine*, et calcule à partir de celle-ci une séquence de nombres. Cette séquence de nombre peut paraître aléatoire si on ne connaît pas la graine. Cependant elle est totalement déterministe, et une même graine donnera toujours la même séquence de nombres. Dans notre cas, une mosaïque est une suite de $N \times N$ dans $\{0, 1\}$. Si nous considérons le même pixel initial (i, j) et observons répétitivement le même parcours à partir de ce pixel, alors nous rencontrerons toujours la même suite de nombres. En revanche si nous partons d'un pixel quelconque, nous ne savons pas quelle trajectoire nous allons obtenir.

1 Théorèmes d'échantillonnage

1.1 Echantillonnage régulier

1.1.1 Théorème de Shannon

Soit f un signal continu à une dimension, de $\mathbb{R}$ dans $\mathbb{R}$. La version régulièrement échantillonnée f_{ech} de ce signal s'écrit comme le produit du signal par un peigne de Dirac :

$$f_{ech}(t) = f(t)\,\text{Ш}_T(t) = f(t) \sum_{k=-\infty}^{+\infty} \delta(t - kT)$$

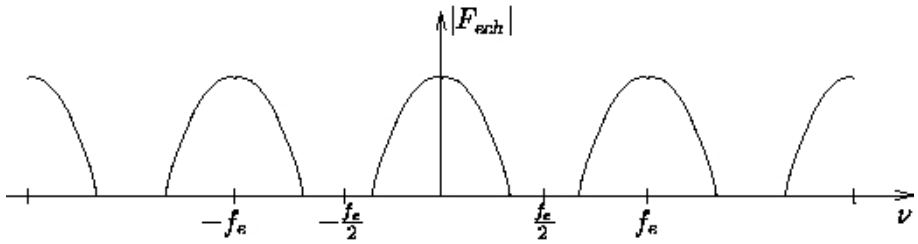


FIG. *4.1: Schéma du spectre d'amplitude de f_{ech}. Le spectre de f est périodisé.*

où T est l'intervalle entre deux échantillons. La transformée de Fourier F_{ech} du signal échantillonné est égale, par propriété de la transformée de Fourier (TF), au produit de convolution entre la TF de f et celle du peigne de Dirac $\amalg(t)$:

$$F_{ech}(\nu) = F(\nu) * \frac{1}{T} \sum_{k=-\infty}^{+\infty} \delta(\nu - \frac{k}{T}) = \frac{1}{T} \sum_{k=-\infty}^{+\infty} F(\nu - \frac{k}{T})$$

Dans le cas d'un échantillonnage régulier, le spectre de f est donc périodisé. Il est répliqué tous les $\nu_e = \frac{1}{T}$ (Figure 4.1).

Ainsi, si le signal f est à bande limitée (c'est-à-dire si sa TF est à support borné), alors f peut être reconstruit sans erreur à partir de ses échantillons par l'intermédiaire d'un filtre passe-bas idéal, à condition que la fréquence d'échantillonnage ν_e soit supérieure à deux fois la fréquence maximale de f. Et nous avons dans ce cas :

$$F(\nu) = F_{ech}(\nu) \cdot rect_{\nu_e}(\nu)$$

où $rect_{\nu_e}$ est une porte centrée sur la fréquence nulle et de largeur ν_e et d'amplitude T. La

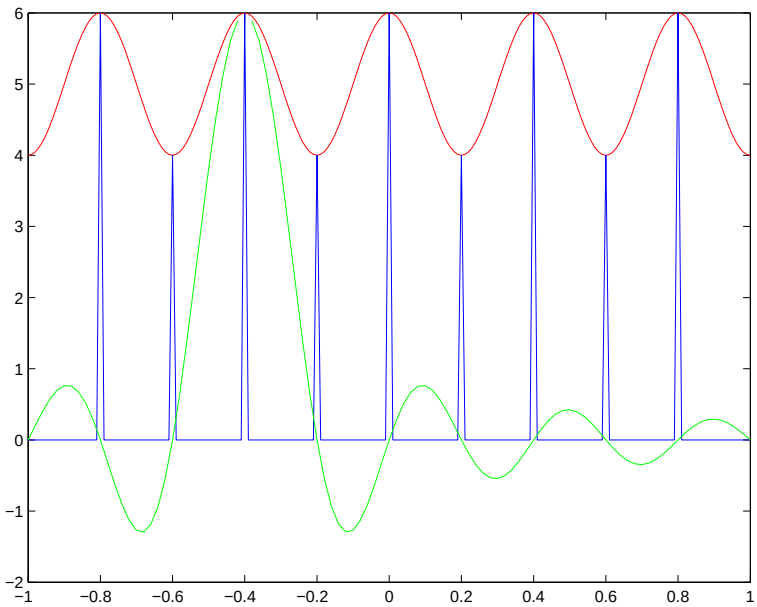


FIG. *4.2: Sinus cardinal (en vert) pondéré par un échantillon. Il participe à la reconstruction du signal original (en rouge). Les échantillons sont représentés en bleu. Une condition nécessaire à la reconstruction parfaite est que l'interpolateur prenne la valeur 1 à l'instant d'un échantillon et la valeur 0 aux instants de tous les autres échantillons.*

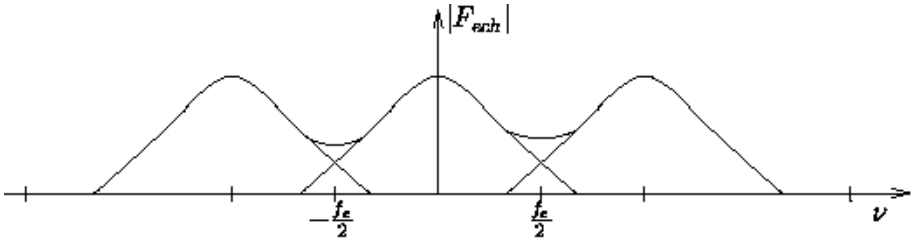


FIG. *4.3: Spectre d'amplitude de f_{ech}. Le spectre de f et ses réplicats se chevauchent. On ne peut pas récupérer f par filtrage passe-bas.*

transformée inverse d'une porte étant le sinus cardinal, le signal reconstruit s'écrit :

$$\boxed{f(t) = \sum_{k=-\infty}^{+\infty} f(kT) sinc(\pi\nu_e(t - kT))} \tag{4.1}$$

Ceci constitue le théorème de Shannon. La fréquenc $\frac{\nu_e}{2}$ est appelée fréquence de Nyquist. Elle correspond à la fréquence maximale que peut contenir le signal f si on souhaite pouvoir le reconstruire par filtrage linéaire sans erreur.

Un signal analogique à bande limitée est donc égal à une somme pondérée de sinus cardinaux décalés et dont les coefficients de pondérations sont les échantillons mesurés à intervalles réguliers. Il est à noter que si l'on veut une reconstruction sans erreur du signal, alors la fonction d'interpolation utilisée doit valoir 1 en 0 (pour que l'on retrouve bien les échantillons mesurés) et doit s'annuler tous les nT (afin de ne pas perturber la valeur des échantillons mesurés). Cette condition est remplie par la fonction sinus cardinal (Figure 4.2).

Dans la pratique, la plupart des signaux ne sont pas à bande limitée, ne serait-ce que parce que la fenêtre d'observation est nécessairement de durée limitée. Ainsi une telle reconstruction souffre d'artefacts de reconstruction se manifestant par des oscillations le long des contours. Cet artefact est appelé phénomène de Gibbs.

Par ailleurs le support infini de la fonction sinus cardinal rend l'implémentation exacte irréalisable, puisque la somme dans l'Equation (4.1) est alors infinie. La reconstruction de signaux à l'aide de fonctions à supports bornés, telles que les splines [Unser 99], est nettement plus efficace. De plus la décroissance lente du sinus cardinal rend l'interpolation très sensible au bruit.

1.1.2 Repliement spectral

Dans le cas où la condition de Shannon n'est pas repectée, c'est-à-dire lorsque ν_{max}, la fréquence maximale du signal, est supérieure à la fréquence de Nyquist, alors le spectre répliqué à la fréquence ν_e vient recouvrir une partie du spectre centré en zéro (Figure 4.3). C'est ce que l'on appelle le repliement spectral, ou *aliasing* en anglais. Le signal reconstruit par filtrage passe-bas n'est plus égal à f.

Les composantes fréquentielles situées au-dessus de la fréquence de Nyquist sont translatées en fréquences et se retrouvent alors à des fréquences inférieures à la fréquence de Nyquist. Ces composantes hautes fréquences sont alors perçues comme une information basse fréquence. Cette information n'a alors pas de sens, elle est erronée. La Figure 4.4 illustre ce phénomène.

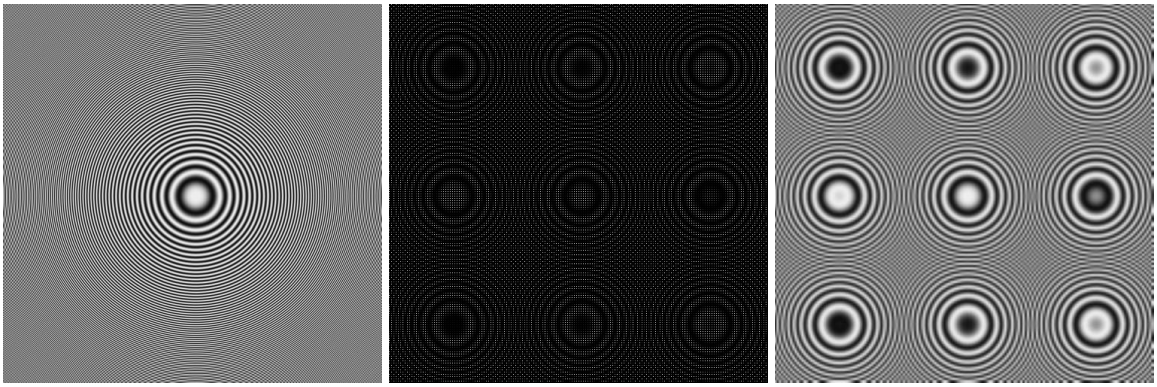


FIG. *4.4: Illustration du phénomène de Moiré dû au repliement spectral. A gauche : signal correctement échantillonné ; Au centre : on ne garde qu'un échantillon sur neuf, régulièrement espacés ; A droite : image reconstruite à partir de la version sous-échantillonnée. Des motifs structurés apparaissent dans les régions de hautes fréquences de l'image.*

L'image de gauche, constituée de cercles concentriques respecte le théorème de Shannon, sa fréquence maximale est à la limite de la fréquence de Nyquist. L'image de droite est la reconstruction de la même image dont on a gardé un échantillon sur neuf. L'image sous-échantillonnée viole donc largement largement le théorème de Shannon. Le résultat de l'interpolation nous montre une image dont le contenu est totalement différent de l'image originale. Quelle est la cible sur les neufs qui apparaissent qui correspond à la réalité ? On ne peut pas le savoir si on ne connaît pas l'image originale.

Par ailleurs, dans le cas bi-dimensionnel, non seulement un signal de fréquence différente peut apparaître, mais il peut également être d'orientation différente. C'est ce qui se produit dans l'exemple de la Figure 4.4, puisque l'image reconstruite n'est pas consituée de cercles concentriques dont les fréquences auraient été simplement démodulées. Il y a bien eu également des changements d'orientation des motifs contenus dans l'image, aboutissant à plusieurs motifs de cercles concentriques. Le résultat est donc une information d'autant plus trompeuse pour l'observateur, puisque ni la fréquence ni l'orientation ne correspondent à la scène originale.

Le cas des signaux à deux dimensions a l'avantage par rapport aux signaux à une dimension de permettre des techniques adaptatives pour contourner ce repliement. Lorsqu'il y a de l'*aliasing* dans une direction, d'autres directions peuvent en être exemptes.

L'*aliasing* peut également se manifester dans le cas d'un échantillonnage temporel sur un objet en mouvement, comme dans l'exemple classique des roues tournant à l'envers dans les films.

1.2 Transposition au cas irrégulier : le théorème de Yen

Le théorème de Shannon peut être transposé au cas d'un échantillonnage irrégulier, c'est-à-dire un échantillonnage dont le pas entre deux échantillons voisins n'est pas constant. S'il existe une fonction γ continue et inversible de $\mathbb{R}$ dans $\mathbb{R}$ telle que $\gamma(t_n) = nT$ et telle que $f(\gamma^{-1}(.))$ soit à bande limitée, alors f peut être parfaitement reconstruite à partir de ses échantillons de la

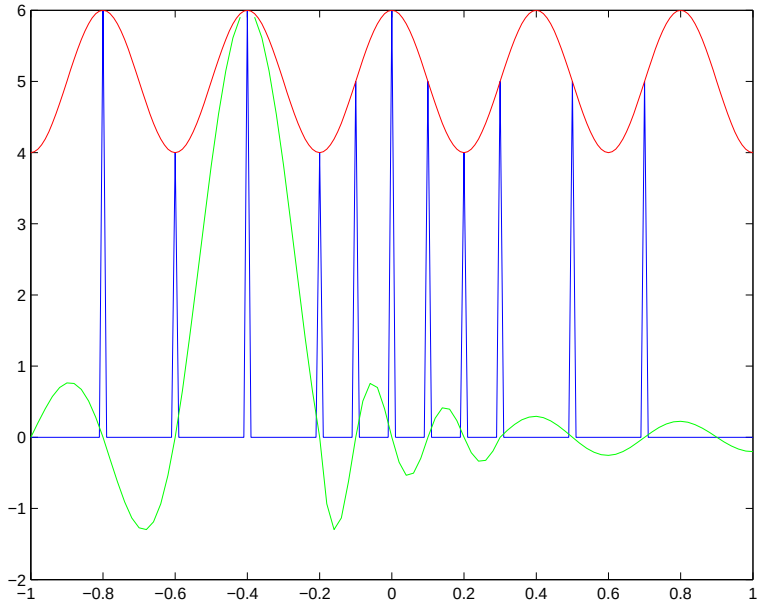


FIG. *4.5: Même exemple que Figure 4.2, cependant quelques échantillons ont été ajoutés. La fréquence d'échantillonnage est doublée sur une portion du signal. L'interpolateur idéal "sinus cardinal" doit s'adapter à la position des échantillons afin que ses zéros suivent la même irrégularité.*

manière suivante [Clark 85] :

$$f(t) = \sum_{n=-\infty}^{+\infty} f(t_n) sinc(\omega_0(\gamma(t) - nT))$$

Pour démontrer ce résultat, supposons qu'une telle fonction γ, c'est à dire que la fonction $f(\gamma^{-1}(.))$ respecte le critère de Shannon. Alors $f(\gamma^{-1}(.))$ peut s'écrire :

$$\begin{aligned} f(\gamma^{-1}(t)) &= \sum_{n=-\infty}^{+\infty} f(\gamma^{-1}(nT)) sinc(\omega_0(t - nT)) \\ &= \sum_{n=-\infty}^{+\infty} f(t_n) sinc(\omega_0(t - nT)) \end{aligned}$$

Et, par changement de variable $\tau = \gamma^{-1}(t)$, on obtient :

$$f(\tau)) = \sum_{n=-\infty}^{+\infty} f(t_n) sinc(\omega_0(\gamma(\tau) - nT))$$

Ce qui démontre le résulat.

Remarquons que dans le cas où γ est la fonction identité, c'est-à-dire lorsque l'échantillonnage est régulier, on retombe bien sur le théorème de Shannon énoncé dans le précédent paragraphe.

Dans le cas régulier, la même fonction interpolatrice (le sinus cardinal) est appliquée à chaque échantillon. L'interpolation (ou le filtrage) est alors dite "uniforme" ou spatialement invariante. Au contraire dans le cas irrégulier chaque échantillon pondère une fonction interpolatrice différente, du fait de la dépendance en $\gamma(t)$. L'interpolation est alors spatialement variante.

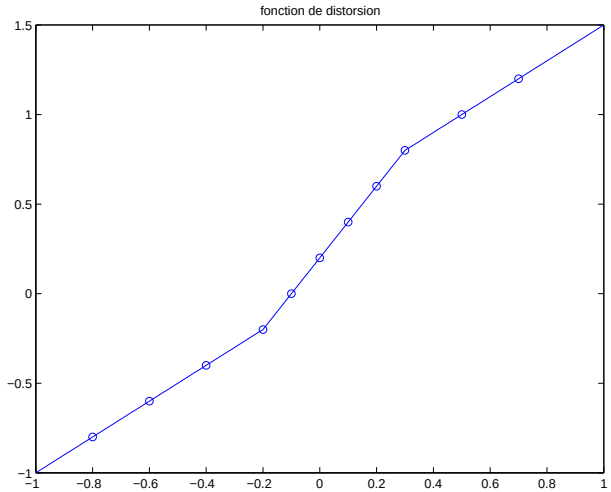


FIG. *4.6: Fonction de distorsion associée à l'exemple de la Figure 4.5. Cette fonction effectue une transformation sur le temps afin de répartir les échantillons régulièrement. Elle a été obtenu par interpolation linéaire des points* $nT = \gamma(t_n)$ *(cercles). L'abscisse représente les instants successifs des échantillons irréguliers ; l'ordonnée représente les instants successifs des échantillons réguliers obtenus par transformation.*

Cette formule d'interpolation peut s'interpréter comme la modification locale de la fréquence du sinus cardinal afin que celui-ci *s'adapte* aux instants d'échantillonnage, c'est-à-dire qu'on l'étire ou le contracte afin qu'il s'annule aux instants d'échantillonnage. Cela se voit sur la Figure 4.5, Figure qui reprend l'exemple de la Figure 4.2 en doublant le nombre d'échantillons sur un certain segment.

Une autre manière d'interpréter cette formule est de voir cela comme une distortion (par dilatations et contractions) du signal échantillonné afin que les échantillons soient régulièrement répartis. Cette transformation s'effectue par le biais de la fonction γ, appelée fonction de distorsion. Celle associée à l'exemple de la Figure 4.2 est représentée Figure 4.6.

Appelons S_{ω_0} l'ensemble des fonctions dont le spectre est contenu dans la bande $[-\omega_0, +\omega_0]$, donc échantillonnable à la fréquence $2\omega_0$. Appelons $S^{\gamma}_{\omega_0}$ l'ensemble des images des S_{ω_0} par $\gamma()$. Il est intéressant de noter que les éléments de $S^{\gamma}_{\omega_0}$ ne sont pas nécessairement à bande limitée, et ce, même si le nombre moyen d'échantillons respecte le critère de Nyquist. Si $\gamma()$ réalise un zoom uniforme (ie : $\gamma(t) = at$), alors la largeur de bande de f est diminuée d'un facteur $1/a$.

$$f(at) \xleftrightarrow{TF} \frac{1}{a} F(\frac{\nu}{a})$$

Comme, en première approximation, γ peut être considérée comme localement linéaire, on réalise localement des contractions/dilatations, et ainsi la largeur de bande du signal se trouve modifiée.

Le fait de modifier la fréquence maximale d'un signal grâce à une fonction de distorsion trouve une application directe dans la compression de données. Dans le cas de l'échantillonnage régulier, la fréquence d'échantillonnage se fait sur la base de la fréquence maximale globale du signal. Or cette fréquence maximale peut n'apparaître que de manière très localisée dans le signal, et ainsi on sur-échantillonne certaines portions du signal. Ce sur-échantillonnage n'a aucune conséquence sur la reconstruction mais il est inutilement coûteux en mémoire. Au

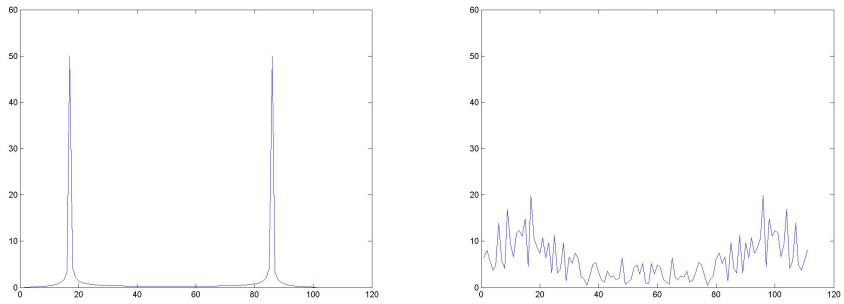


FIG. *4.7: Spectre d'une sinusoïde sous-échantillonnée au-delà du critère de Nyquist : échantillonnage régulier (à gauche), aléatoire (à droite).*

contraire il est possible d'estimer la bande passante instantanée du signal à échantillonner et ainsi de faire varier la densité locale en échantillons nécessaire à la reconstruction parfaite du signal [Clark 85, Le Floch 97]. Le nombre d'échantillons utilisés dans ce cas est le nombre minimal (optimal du point de vue de la compression des données).

L'interpolation de Yen souffre des mêmes défauts que l'interpolation par sinus cardinal dans le cas régulier, à savoir la complexité algorithmique, la décroissance lente et le ringing. Elle est donc très peu utilisée en pratique. Une application directe du théorème de Yen est faite dans [Ahumada 89], par le biais de réseaux de neurones.

1.3 Comparaison d'une interpolation à partir d'un sous-échantillonnage régulier vs. irrégulier

Un exemple de conséquence spectrale du sous-échantillonnage d'un signal est donnée Figure 4.7. Ces courbes ont été obtenue en sous-échantillonnant une sinusoïde de manière régulière dans un cas, et aléatoire dans l'autre. Dans le cas régulier l'énergie est concentrée sur deux fréquences. C'est la conséquence de l'*aliasing*. Dans le cas aléatoire, cette énergie est répartie sur une large bande de fréquences. Le signal résultant n'a pas de cohérence, il ne véhicule pas d'information.

La Figure 4.8 montre le résultat de l'interpolation de l'image CZP dans les cas régulier et irrégulier. Comme on peut le voir, les cercles concentriques résultants du repliement spectral sont transformés en bruit. Ainsi, l'échantillonnage irrégulier permet de supprimer une information erronée.

Du point de vue perceptif, il est rapporté que les symétries et la répétition d'un signal basse-fréquence sont hautement perceptibles, et sont difficiles à éliminer perceptuellement [Oliva 06, Barstow 53, Huang 65].

2 Analyse spectrale

Comme l'illustre la Figure 4.8, l'échantillonnage aléatoire a pour propriété de convertir l'effet de Moiré, inhérent à l'échantillonnage régulier, en bruit. Il s'agit dans les deux cas de la conséquence du repliement spectral. La littérature dans le domaine a tendance à parler de

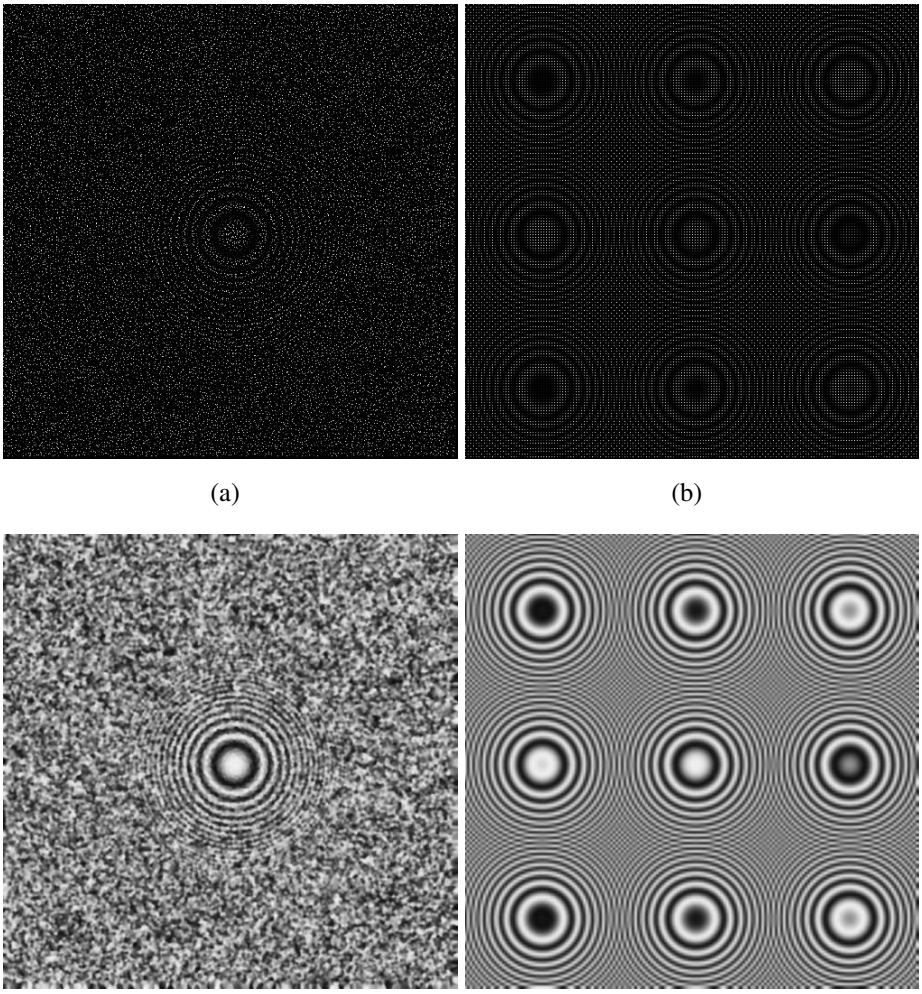


FIG. *4.8: (a) Image CZP sous-échantillonnée irrégulièrement. (b) Image CZP sous-échantillonnée régulièrement. 1 échantillon sur 9 a été conservé dans les deux cas. (c) image irrégulière interpolée,* $psnr = 9,44dB$*. (d) image régulière interpolée,* $psnr = 9.36dB$*. La méthode d'interpolation utilisée dans les deux cas est la convolution normalisée basée sur un noyau gaussien d'écart-type 2 pixels. Les deux images ont la même qualité objective, mais véhiculent des informations très différentes.*

propriété d'*anti-aliasing* à propos de l'échantillonnage irrégulier. Il me semble qu'il faudrait parler plutôt d'anti-Moiré. Car dans les deux cas, le repliement spectral intervient. C'est sa manifestation qui diffère, et ainsi la perception que l'on en a.

C'est l'aspect fréquentiel qui permet dans le cas régulier de comprendre les conséquences du sous-échantillonnage sur le signal. C'est pourquoi nous nous intéressons au contenu fréquentiel des signaux échantillonnés aléatoirement.

2.1 Approche intuitive

Pour comprendre la propriété anti-Moiré de l'échantillonnage irrégulier, considérons l'exemple de l'échantillonnage d'une sinusoïde. Si la fréquence de la sinusoïde est basse par rapport à la fréquence d'échantillonnage, alors une pertubation aléatoire de l'échantillonage régulier revient

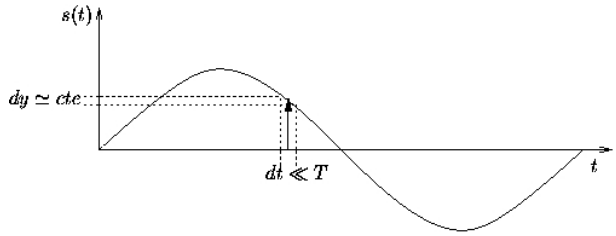


FIG. *4.9: Sinusoïde échantillonnée par un échantillonnage régulier perturbé. La période T de la sinusoïde est grande par rapport à l'amplitude de la perturbation des échantillons dt. La valeur des échantillons n'est par conséquent que légèrement perturbée.*

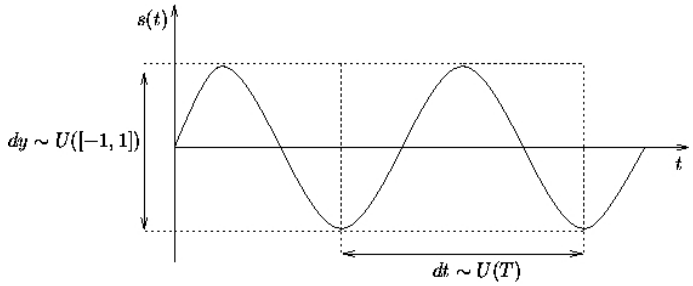


FIG. *4.10: Sinusoïde échantillonnée par un échantillonnage régulier perturbé. La période T de la sinusoïde est du même ordre de grandeur que l'ampitude de la perturbation des échantillons dt. Si l'amplitude des perturbations suit une loi uniforme, alors les mesures des échantillons peuvent prendre n'importe quelle valeur entre -1 et 1, ils suivent une loi uniforme. On a un bruit blanc.*

à une perturbation aléatoire de l'amplitude de chaque échantillon (Figure 4.9). La fréquence de la sinusoïde reste la même, cependant la perturbation de l'amplitude des échantillons se traduit par un peu de bruit. Ce bruit est d'autant plus faible que la fréquence de la sinusoïde est basse, puisque les perturbations sont plus petites.

Maintenant, considérons une sinusoïde dont la fréquence est largement supérieure à la fréquence de Nyquist. Dans le cas régulier, on obtiendrait après reconstruction une sinusoïde de basse fréquence, c'est-à-dire un signal contenant une information perceptible erronée. Supposons que les échantillons soient perturbés aléatoirement de manière uniforme. Alors l'échantillon peut se trouver n'importe où sur plusieurs périodes de la sinusoïde avec la même probabilité. Cela signifie que l'échantillon peut prendre n'importe quelle valeur de la sinusoide entre -1 et 1 avec la même probabilité (cf Figure 4.10). De plus les échantillons étant perturbés de manière indépendante, il en résulte que la valeur des échantillons est elle aussi décorrélée. Les échantillons ne représentent donc plus une information erronnée comme c'est le cas dans l'échantillonnage régulier avec l'*aliasing*, mais plutôt du bruit, un signal dépourvu de sens.

Il existe trois grandes classes de distributions aléatoires d'échantillons. Les distributions dites de "Poisson" et de "Poisson à disque" sont définies par la loi de probabilité régissant l'intervalle entre deux échantillons consécutifs. La troisième s'obtient par la perturbation d'un échantillonnage régulier. Dans la suite de ce chapitre nous nous proposons d'étudier les propriétés de ces trois échantillonnages. Dans chaque cas, l'espérance du module carré de la TF ainsi que la Densité Spectrale de Puissance moyenne (DSPm) seront données.

2.2 Echantillonnage de Poisson

L'échantillonnage de Poisson est un échantillonnage dont la distance entre deux échantillons successifs suit une loi de probabilité exponentielle $p(t)$ dont le paramètre $\lambda = N/T$ est le taux d'échantillonnage moyen :

$$p(t) = \begin{cases} \lambda e^{-\lambda t} & \text{si } t \geq 0 \\ 0 & \text{si } t < 0 \end{cases}$$

Le nombre d'échantillons entre deux points suit quant à lui une loi de Poisson $\frac{(\lambda t)^i}{i!} e^{-\lambda t}$, donnant ainsi son nom au processus. Un tel échantillonnage peut également être vu comme le résultat de N tirages indépendants x_n suivant une loi uniforme sur un intervalle $[-\frac{T}{2}, \frac{T}{2}]$. Dans ce cas, la distance entre chaque tirage et le début de l'intervalle peut être vu comme la réalisation d'une variable suivant une loi exponentielle de paramètre $\lambda_i = 1/T$. Par théorème, la variable $X = min\{x_n\}$ suit une loi exponentielle de paramètre $\sum_n \lambda_i = N/T = \lambda$. Par conséquent un échantillonnage composé de N_T échantillons dont chacun d'eux est le résultat d'un tirage uniforme sur un intervalle de longueur T, est effectivement un échantillonnage de Poisson de paramètre $\lambda = N_T/T$.

Sur un intervalle de longueur T, on montre que la valeur moyenne du module carré de la TF vaut :

$$TF\{\underset{T}{\text{Ш}}\}(\nu) = N_T^2 \delta(\nu) + N_T$$

Sur un intervalle de longueur infinie, notre échantillonneur $\text{Ш}(t)$ est un signal de puissance moyenne finie. La DSP $\Gamma(\nu)$ d'un tel signal est définie comme la TF de son autocorrélation $\gamma(\tau)$ définie comme :

$$\gamma(t) = \lim_{T \to \infty} \frac{\int_{-T/2}^{T/2} \text{Ш}(t)\, \text{Ш}(t-\tau) dt}{T}$$

La DSP moyenne vaut alors :

$$E[\Gamma_{\text{Ш}}(\nu)] = \lambda^2 \delta(\nu) + \lambda$$

Finalement, la DSP moyenne d'un signal f sera :

$$E[\Gamma_{f.\,\text{Ш}}(\nu)] = \Gamma_f(\nu) * E[\Gamma_{\text{Ш}}(\nu)] = \lambda^2 \Gamma_f(\nu) + \lambda \underbrace{\int_{-\infty}^{+\infty} \Gamma(\nu) d\nu}_{\text{puissance de } f}$$

On retrouve donc bien la DSP du signal analogique, cependant un bruit blanc y est ajouté.

2.3 Echantillonnage à gigue

L'échantillonnage à gigue (*jitter sampling* en anglais) s'obtient par la perturbation de chaque échantillon d'un échantillonnage régulier. Le pas d'échantillonnage régulier (qui est l'inverse de la densité en échantillons) vaut $1/\lambda$. Chaque échantillon est perturbé de manière indépendante d'une valeur ε_n, suivant une loi de probabilité $p(\varepsilon)$. Le signal échantillonné s'écrit :

$$x_{ech}(t) = x(t) \sum_{n=-\infty}^{+\infty} \delta(t - n/\lambda + \varepsilon_n)$$

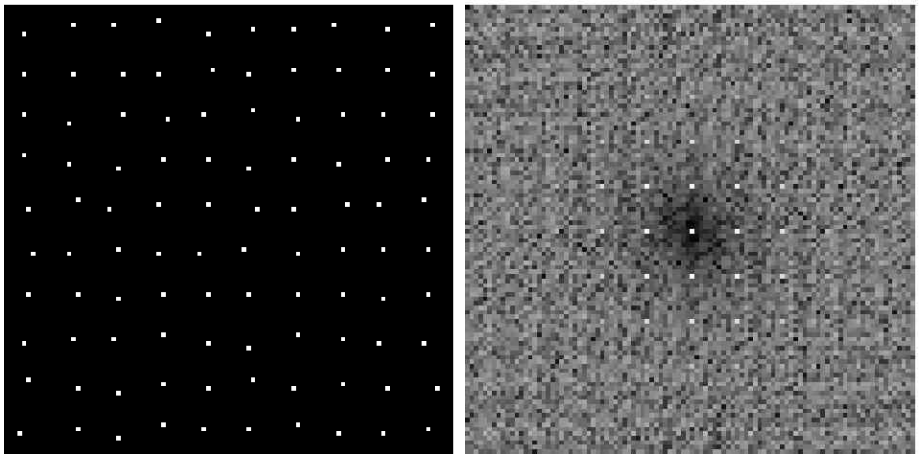


FIG. *4.11: Exemple d'échantillonnage à gigue et spectre d'amplitude correspondant. Les impulsions de Dirac sont progressivement atténuées à mesure que l'on s'éloigne de la fréquence centrale. Le bruit basse fréquence est également atténué.*

On montre, en passant par l'espérance de l'autocorrélation, que la DSPm de l'échantillonneur vaut dans ce cas [Dippé 85] :

$$DSP_m(\nu) = \lambda^2|\Pi(\nu)|^2 \sum_{n=-\infty}^{+\infty} \delta(\nu - n\lambda) + \lambda(1 - |\Pi(\nu)|^2)$$

Les spectres répliqués sont donc pondérés par la fonction caractéristique de la loi de répartition. Ainsi les réplicats sont atténués, et la limite théorique de shannon est repoussée. Le bruit large-bande qui apparaît est atténué au niveau des basses fréquences, c'est un bruit bleu. Un exemple d'un tel échantillonneur et de son spectre d'amplitude est donné à la Figure 4.11.

Dans le cas particulier où $p(\varepsilon)$ suit une loi uniforme sur $[\frac{\Delta}{2}, \frac{\Delta}{2}]$, alors :

$$\begin{aligned} DSP_m(\nu) &= \lambda^2 sinc^2(\pi\nu/\lambda) \sum_{n=-\infty}^{+\infty} \delta(\nu - n\lambda) + \lambda(1 - sinc^2(\pi\nu/\lambda)) \\ &= \lambda^2\delta(\nu) + \lambda(1 - sinc^2(\pi\nu/\lambda)) \end{aligned}$$

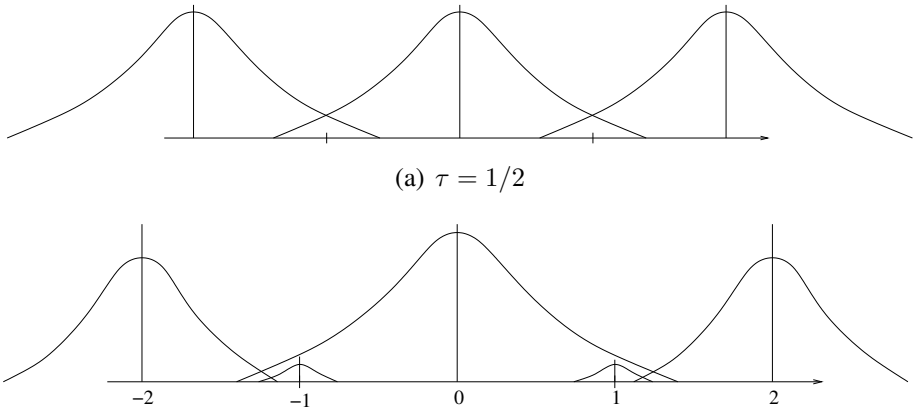
Dans ce cas, les réplicats sont totalement éliminés. On retrouve bien le spectre orginal (en calculant la moyenne sur une infinité de cas).

Remarque : dans le cas où on ne connait pas la position exacte des échantillons et que l'erreur commise est une loi uniforme entre -B et +B (c'est le cas d'une gigue que l'on ne contrôle pas), Bland *et al.* montrent que le spectre moyen est atténué par un facteur [Bland 97] :

$$E[F_{ech}(\nu)] = F(\nu)sinc(2\pi\nu B)$$

L'estimation est alors biaisée. Ce dernier résultat est valable aussi bien pour l'échantillonnage irrégulier que pour l'échantillonnage régulier.

Une manière de comprendre intuitivement comment le transfert se fait entre des porteuses fixes dans le cas réguliers et des porteuses large bande en passant du cas régulier au cas irrégulier consiste à considérer le cas simple de la somme de deux peignes de Dirac Ш_1 et Ш_2 de même



(a) $\tau = 1/2$

(b) $\tau \neq 1/2$

FIG. *4.12: Illustration du spectre d'un signal échantillonné par la somme de deux peignes de Dirac de période 1 et décalés de τ. (a) $\tau = 1/2$, cas de l'échantillonnage régulier, le spectre est périodisé toutes les fréquences 2n, avec une amplitude de deux. (b) $\tau \neq 1/2$, l'échantillonnage est irrégulier. Il y a un transfert d'énergie de la fréquence 2 vers la fréquence 1, tandis que le spectre en bande de base garde la même amplitude. Les conditions de repliement spectral sont donc changées.*

période $T = 1$ et décalés d'un temps τ :

$$\begin{aligned} \text{Ш}_1(t) + \text{Ш}_2(t) &= \sum_{n=-\infty}^{+\infty} \delta(t-n) + \sum_{n=-\infty}^{+\infty} \delta(t-n-\tau) \\ &= \sum_n \delta(t-n) * (\delta(t) + \delta(t-\tau)) \end{aligned}$$

Lorsque τ vaut $1/2$ (modulo 1) on retombe alors sur un échantillonnage régulier. Sinon, l'échantillonnage est irrégulier, la distance entre deux échantillons voisins étant soit τ, soit $1-\tau$. Prenons la TF de cet échantillonneur :

$$TF\{\text{Ш}_1 + \text{Ш}_2\}(\nu) = \sum_n \delta(\nu - n)(1 + e^{-2j\pi\nu\tau})$$

Si $\tau = 1/2$, l'exponentielle complexe vaut -1 pour les fréquences impaires et vaut $+1$ aux fréquences paires. Ainsi on retrouve le spectre d'un échantillonneur régulier de période $1/2$:

$$\tau = 1/2 \quad \Rightarrow \quad TF\{\text{Ш}_1 + \text{Ш}_2\}(\nu) = 2\sum_n \delta(\nu - 2n)$$

Maintenant, si $\tau \neq 1/2$, alors l'exponentielle complexe n'atteint jamais -1 aux fréquences impaires. Ainsi les impulsions de Dirac aux fréquences impaires ne sont pas annulées. L'amplitude de ces porteuses est minimale (elle tend vers zéro) pour τ proche de $1/2$, et maximale pour $\tau = 0$ (l'échantillonneur est alors le peigne régulier de période 1). Par ailleurs, les fréquences paires $(2k)$ sont quant à elles atténuées car le facteur $(1 + e^{-4j\pi k\tau})$ est strictement inférieur à deux, et est minimum pour $\tau = 1/4$. Ainsi :

$$TF\{\underset{1}{\text{Ш}}(t) + \underset{2}{\text{Ш}}(t)\} = 2\delta(\nu) + \underbrace{\alpha_i \sum_n \delta(\nu - 2n - 1)}_{\text{fréquences impaires}} + \underbrace{\alpha_p \sum_n \delta(\nu - 4n - 2)}_{\text{fréquences paires}}$$

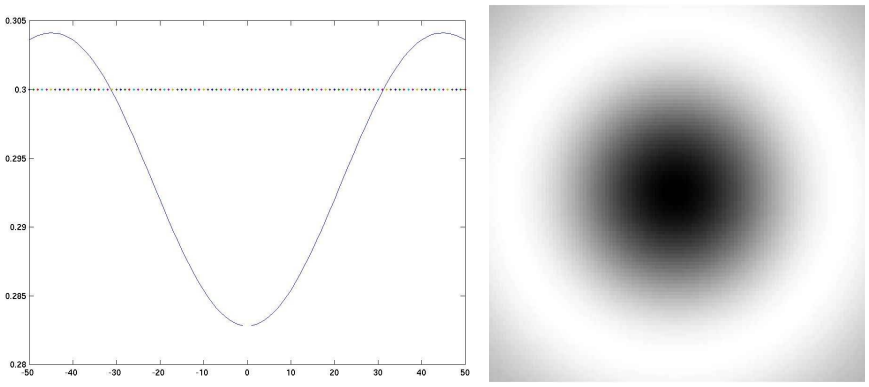


FIG. *4.13: Spectre de l'échantillonnage de Poisson à disque dans les cas à une dimension et à deux dimensions pour* $d_{min} = 0.1$ *et* $\lambda = 0.3$*. La composante continue n'est pas représentée. Les pointillés donnent le niveau de* λ*, c'est-à-dire ce qui aurait été obtenu dans le cas d'un échantillonnage de Poisson.*

Si τ est inférieur à $1/2$, alors $|\alpha_i = \sqrt{-2 + 2cos(2\pi\tau)}|$ est strictement positif, et $\alpha_p = \sqrt{-2 + 2cos(4\pi\tau)}$ strictement inférieur à 2. Cela modifie les conditions de repliement spectral (Figure 4.12).

2.4 Echantillonnage de Poisson à disque

Cet échantillonnage reprend l'échantillonnage de Poisson en y ajoutant une contrainte de distance minimale entre deux échantillons. L'intervalle inter-occurence suit alors une loi exponentielle de paramètre λ' retardée de t_{min} :

$$p(t) = \{\lambda' e^{-\lambda' t} * \delta(t - t_{min})\}_{[t_{min},\infty]}$$

L'espérance de l'intervalle inter-occurrence est alors $(d_{min} + 1/\lambda')$, donc le taux d'échantillonnage λ est :

$$\lambda = \frac{1}{d_{min} + 1/\lambda'} = \frac{\lambda'}{1 + d_{min}\lambda'}$$

On montre alors que la densité spectrale de puissance moyenne vaut [Dippé 85] :

$$\boxed{\Gamma_s(\nu) = \begin{cases} \lambda^2\delta(\nu) & \text{si } \nu = 0 \\ \lambda\left(1 - \dfrac{2\lambda'\nu\sin(d_{min}\nu) + 2\lambda'^2\cos(d_{min}\nu) - 2\lambda'^2}{2\lambda'\nu\sin(d_{min}\nu) - 2\lambda'^2\cos(d_{min}\nu) + \nu^2 + 2\lambda'^2}\right) & \text{sinon} \end{cases}}$$

Le spectre de Fourier de cet échantillonnage comprend une impulsion de Dirac à la fréquence nulle ainsi que du bruit large bande, mais dont l'énergie basse fréquence est atténuée (Figures 4.13 et 4.14). Ce type de bruit est un bruit bleu. L'absence de basses fréquences traduit le fait qu'il n'y ait pas de région plus dense que d'autres en échantillons. On évite ainsi les amas d'échantillons survenant dans le cas de l'échantillonnage de Poisson.

Remarque : l'échantillonnage de Poisson et l'échantillonnage régulier sont deux cas limites de l'échantillonnage de Poisson à disque. L'échantillonnage de Poisson est obtenu dans le cas où la distance minimale entre deux échantillons voisins est nulle. L'échantillonnage régulier est obtenu dans les cas où λ' tend vers l'infini, c'est-à-dire lorsque la probabilité d'avoir un échantillon à la distance d_{min} est égale à un.

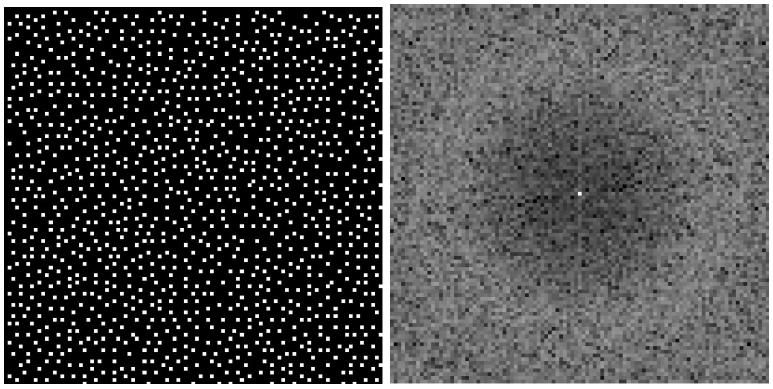


FIG. *4.14: Exemple d'échantillonnage de Poisson à disque, et spectre correspondant. Les basses fréquences sont atténuées.*

L'échantillonnage aléatoire permet donc de retrouver en moyenne le spectre d'un signal sans être limité en théorie par l'*aliasing*. Ainsi l'échantillonnage aléatoire peut sembler être une solution "miracle" au problème de repliement spectral. L'échantillonnage irrégulier est en fait particulièrement bien adapté pour faire l'analyse spectrale de signaux stationnaires, car on peut alors calculer des moyennes (utilisation par exemple d'un périodogramme moyenné). Dans ce cas on s'affranchit de la limite de Nyquist, on ne récupère pas de réplicats, mais on augmente le niveau de bruit.

Bien que dans le cadre de notre étude l'échantillonneur ne varie pas (la rétine d'un individu est la même pour la vie), l'analyse que nous venons de faire nous donne tout de même des indications sur les positions fréquentielles des réplicats dans le cas aléatoire. Car le bruit qui est ajouté à la $DSPm$ correspond en fait à des porteuses qui modulent le signal. Ainsi nous savons par exemple que plus la contrainte de distance de garde est forte entre deux échantillons successifs, moins il y aura de réplicats dans les basses fréquences.

2.5 Conséquence de la contrainte de grille régulière

Dans nos simulations, nous faisons un échantillonnage irrégulier dont les échantillons sont disposés sur une grille cartésienne. Ainsi les positions des échantillons sont des multiples du pas T de la grille. Ceci introduit de l'*aliasing* dans le sens classique du terme. On peut voir cet échantillonnage comme un peigne de Dirac régulier dont chaque impulsion de Dirac est multipliée par un facteur α_k prenant les valeurs 0 ou 1 selon que l'échantillon est présent ou non :

$$\text{Ш}(t) = \sum_{k=-\infty}^{+\infty} \alpha_k \delta(t - k\Delta)$$

α_k est une variable aléatoire de Bernoulli, d'espérance λ. La TF moyenne de cet échantillonneur est :

$$\begin{aligned} E[TF\{\text{Ш}(t)\}] &= E[\sum_{k=-\infty}^{+\infty} \alpha_k e^{(-2j\pi\nu k\Delta)}] \\ &= \sum_{k=-\infty}^{+\infty} E[\alpha_k] e^{(-2j\pi\nu k\Delta)} \\ &= \lambda.\frac{1}{\Delta} \sum_{k=-\infty}^{+\infty} \delta(\nu - \frac{k}{\Delta}) \end{aligned}$$

On constate donc qu'il y a bien des réplicats aux fréquences multiples de l'inverse du pas de la grille, comme dans le cas d'un échantillonnage régulier. Il y a donc potentiellement du repliement spectral. En revanche il est intéressant de remarquer que si on sous-échantillonne aléatoirement une grille, la fréquence de Nyquist est alors celle de la grille prise dans sa globalité. Il ne faut bien sûr pas omettre qu'un bruit large bande s'ajoute. Il peut donc potentiellement y avoir du repliement spectral sous les deux formes associées respectivement à l'échantillonnage régulier et à l'échantillonnage irrégulier.

Quelle est la conséquence de cette contrainte pour nos mosaïques de couleur aléatoires ? l'intensité peut être en proie au repliement spectral au sens classique, tandis que la chrominance sera sous forme d'un bruit large-bande. Reprenons le modèle d'échantillonnage d'Alleysson et Hérault décrit au chapitre précédent, dans lequel chaque sous-grille d'échantillonnage m_i, avec $i \in \{R, V, B\}$, est décomposée en une composante constante et une composante fluctuante ($m_i = p_i + \tilde{m}_i$). Une image mesurée par une telle mosaïque s'écrit alors :

$$I_m(x, y) = \sum_i p_i C_i(x, y) + \sum_i C_i(x, y)\tilde{m}_i(x, y)$$

Cette équation fait apparaître que le terme d'intensité n'est pas sous-échantillonné par rapport à la grille d'échantillonnage globale (car les p_i sont spatialement invariants), tandis que les composantes de chrominance sont modulées par les fonctions $\tilde{m}_i$ possédant le caractère irrégulier des sous-mosaïques m_i. Ainsi, l'intensité est échantillonnée par la grille globale qui est une grille régulière. Elle est donc soumise au théorème de Shannon pour les signaux échantillonnés régulièrement, c'est-à-dire qu'il peut y avoir du repliement spectral entre le spectre d'intensité et ses réplicats. Les caractéristiques d'échantillonnage induites par les sous-mosaïques seront, quant à elles, totalement transférées aux composantes de chrominance. Si, par exemple, les filtres de couleur rouges, verts et bleus sont déterminés par un tirage aléatoire uniforme sur l'ensemble de l'image, alors le spectre d'une image couleur échantillonnée par une telle matrice sera composé d'un bruit blanc modulant l'information de chrominance et du spectre d'intensité périodisé. Le schéma à une dimension de la Figure 4.15 illustre un tel spectre.

2.6 L'échantillonnage irrégulier ou le théorème de Shannon local

La Figure 4.16 montre la TF à court terme (fenêtre de taille 9×9) pour 9 pixels successifs d'une grille d'échantillonnage irrégulière de type Poisson. Avant chaque transformation, la moyenne locale a été retranchée, ainsi ces figures montrent uniquement les positions des réplicats, il n'y a pas de signal en bande de base. On constate que la position des réplicats varie (de

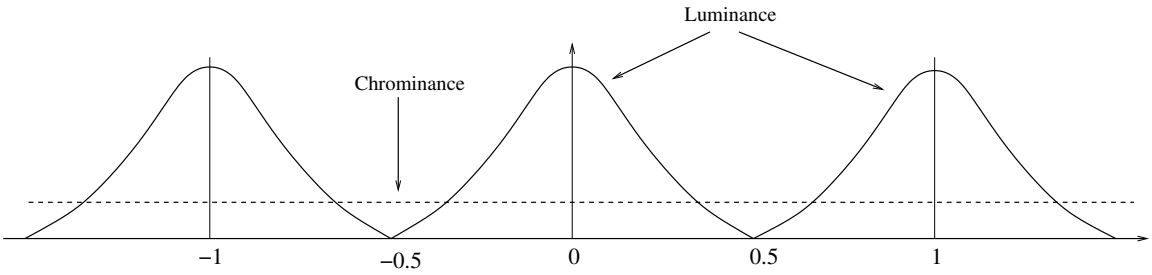


FIG. *4.15: Schéma de la DSPm 1D d'une image couleur échantillonnée par une mosaïque aléatoire à coordonnées entières. l'intensité est périodisée tandis que la chrominance est sous forme de bruit.*

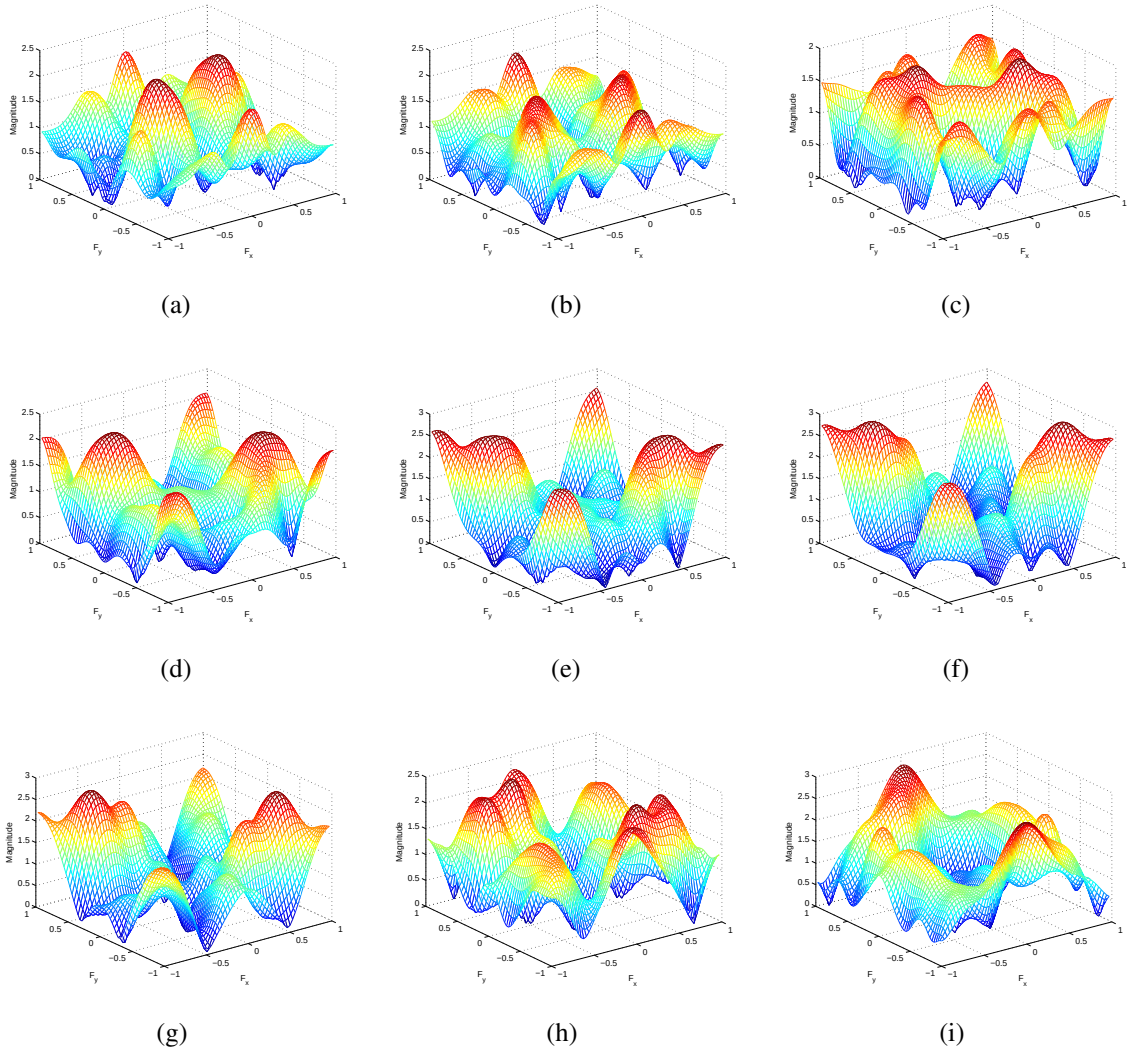
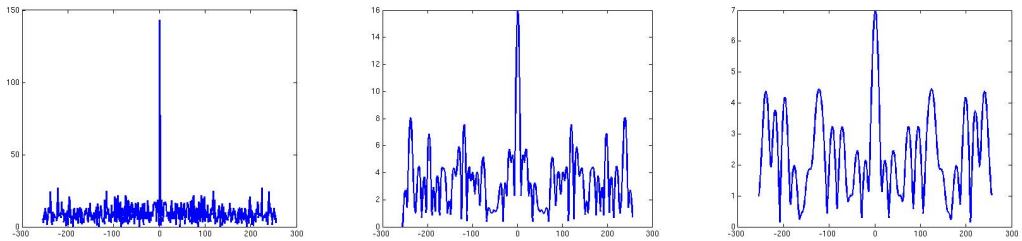


FIG. *4.16: Spectres d'amplitudes locaux d'un échantillonnage irrégulier. Une fenêtre de* 9×9 *pixels centrée successivement sur 9 pixels consécutifs de la grille globale (un pixel contient la valeur 1 ou 0 selon qu'une échantillon est présent ou non). Les valeurs moyennes ont été retranchées avant le calcul de la TF, ainsi les spectres visualisés montrent uniquement les positions des réplicats.*

manière continue) avec la position spatiale. Si la grille avait été régulière, le spectre n'aurait pas varié en fonction de la position spatiale. On remarquera la variété des fréquences occupées par les réplicats. En particulier, comme on peut le voir sur la sous-figure 4.16(c), les porteuses



(a) (b) (c)

FIG. *4.17: Spectre d'amplitude d'un peigne Dirac aléatoire (processus de Poisson) sur une fenêtre de 512 points (a), 64 points (b) et 32 points (c). Tandis que les réplicats semblent de très faible amplitude lorsque la fênetre est de 512 points, leur amplitude augmente si on réduit la taille de la fenêtre d'observation. Localement, les réplicats sont en réalité de même amplitude que le signal en bande de base.*

peuvent descendre dans les basses fréquences, recouvrant ainsi le spectre en bande de base. Comme nous l'avons vu dans la section précédente sur l'analyse spectrale, c'est le type d'aléatoire qui permet de contrôler si des réplicats vont être modulés à des basses fréquences. En particulier, une distance de garde entre deux voisins (échantillonnage de type Poisson à disque) permet de réduire les modulations en basse fréquence, puisque ce sont les amas d'échantillons et les zones vides d'échantillons qui sont responsables de ces basses fréquences.

Nous avons représenté sur la Figure 4.17 le spectre d'amplitude d'un peigne de Dirac irrégulier (échantillonnage de type Poisson) pour une fenêtre d'observation respectivement de 512, 64 et 32 points. Le spectre réalisé sur 512 points nous montre clairement que les répliquats ont une amplitude bien inférieure au signal en bande de base (7 fois plus petite dans cet exemple). On peut penser alors que l'échantillonnage irrégulier permet effectivement de réduire le repliement spectral, du moins pour faire de l'analyse spectrale, puisqu'alors les raies du signal analysé seront bien distinguées du bruit des réplicats. En revanche, comme on peut le voir sur les spectres réalisés sur des fenêtre de 64 et 32 points, à mesure que l'on réduit la taille de la fenêtre d'observation, l'amplitude des réplicats se rapproche de celle du signal en bande de base. Les réplicats ont en réalité une amplitude du même ordre de grandeur que le signal en bande de base, si on fait une analyse locale. Or c'est l'analyse locale qui nous intéresse dans le cadre de notre étude, car nous souhaitons pouvoir interpoler les échantillons manquants en utilisant un noyau le plus compact possible.

Ainsi, lorsque l'on applique la TF à court terme à un signal échantillonné aléatoirement, on se rend compte que les réplicats possèdent la même amplitude que le signal en bande de base, exactement comme dans le cas de la TF appliquée globalement à un échantillonnage régulier. Alors pourquoi lorsque l'on applique la TF de manière globale à un échantillonnage irrégulier, les réplicats ne semblent avoir qu'une faible amplitude par rapport au signal en bande de base ? Cela provient de la localité, dans le domaine spatial, des fréquences qui interviennent dans le spectre de l'échantillonneur. Pour comprendre cela, il est intéressant de considérer la TF globale d'un signal comme la somme de ses TF à court terme. Chaque TF à court terme est composée d'une valeur moyenne - la même valeur pour tous les voisinages du signal - et de fréquences propres au voisinage considéré, du fait de la nature aléatoire de l'échantillonneur. Ainsi lorsque l'on somme ces TF à court terme, les valeurs moyennes de chaque voisinage vont s'additionner

pour donner la valeur moyenne globale, tandis que les autres fréquences ont une plus faible probabilité d'être commune à plusieurs tranches. Ainsi la valeur moyenne aura globalement une plus forte énergie que les autres composantes fréquentielles.

Finalement, on ne peut pas concevoir une méthode d'interpolation en se basant sur des considérations fréquentielles de la TF d'un signal échantillonné irrégulièrement, contrairement au cas régulier. La TF est bien adaptée pour des signaux périodiques (puisque la base de Fourier est constituée de signaux périodiques, les sinusoïdes), ce qui n'est pas le cas, par nature, d'un échantillonneur aléatoire.

Quelle quantité d'information peut potentiellement souffrir de repliement spectral suivant la nature de l'échantillonnage ? Deux cas se présentent selon la nature du signal échantillonné :

1. Le signal présente beaucoup de périodicités. Cela signifie que certaines de ses composantes fréquentielles sont très présentes. Alors, dans le cas régulier, comme les porteuses sont présentes à chaque position spatiale, **si ces composantes violent le théorème de Shannon**, il y a repliement spectral à coup sûr. Au contraire, dans le cas irrégulier, comme la position fréquentielle des porteuses varie avec la position spatiale, il y aura moins de pixels pour lesquels les porteuses provoqueront du repliement spectral. Il y a donc moins de repliement spectral. **Si ces composantes ne violent pas le théorème de Shannon**, l'échantillonnage régulier permet de reconstruire le signal sans erreur. Au contraire, l'échantillonnage irrégulier peut introduire du repliement spectral, puisque les porteuses modulent plus bas en fréquence. L'échantillonnage irrégulier est donc avantageux dans le cas où le théorème de Shannon est violé. Sinon il risque de donner de plus faibles performances que l'échantillonnage régulier.
2. Le signal est de nature plus aléatoire. Cela signifie qu'un nombre plus grand de composantes fréquentielles sont présentes dans son spectre. Dans ce cas la probabilité d'avoir du repliement spectral est la même quelque soit la nature de l'échantillonnage. En effet, à chaque position spatiale, il n'y a pas de probabilité plus élevée pour que signal présente des composantes à telle ou telle fréquence. Ainsi, peu importe à quelles fréquences se trouvent les porteuses.

3 Algorithmes d'interpolation de signaux échantillonnés irrégulièrement

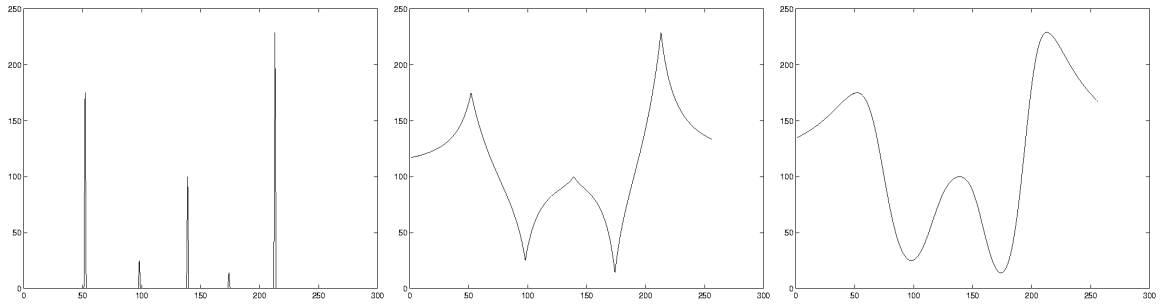
Nous décrivons ici divers algorithmes de la littérature.

3.1 Interpolation de Shepard

L'interpolation de Shepard est également appelée interpolation "pondérée par l'inverse des distances".

Soient $\{f_i\}_{\{i=1...N\}}$ l'ensemble des échantillons connus du signal f et $d_i(x, y)$ la distance du point (x,y) à l'échantillon f_i de coordonnées (x_i, y_i) définie par

$$d_i(x, y) = \sqrt{(x - x_i)^2 + (y - y_i)^2}$$



(a) (b) (c)

FIG. *4.18: Exemple d'interpolation de Shepard. (a) Cinq échantillons à interpoler. (b-c) Interpolations pour* $\alpha = 1$ *et* $\alpha = 2$*, respectivement.*

Alors la fonction d'interpolation de Shepard s'écrit :

$$f(x,y) = \frac{\displaystyle\sum_{i=1}^{N} \frac{f_i}{(d_i(x,y))^\alpha}}{\displaystyle\sum_{i=1}^{N} \frac{1}{(d_i(x,y))^\alpha}}$$

La valeur $f(x,y)$ est une somme pondérée des échantillons mesurés. Le coefficient de pondération est égal à l'inverse de la distance entre l'échantillon et le point (x,y), élevé d'un exposant α. Ainsi plus un échantillon est proche du pixel à interpoler, plus son influence est grande. Au contraire, les échantillons éloignés ont une influence moindre.

La fonction interpolée est de classe C^0 en tout point. Pour $\alpha = 1$ elle est indéfiniment dérivable en tout point autre qu'un échantillon (cf fig.4.18). Une valeur plus élevée de α permet de modifier l'ordre de continuité de la fonction reconstruite. Plus α est grand, plus le poids du plus proche voisin sera prépondérant par rapport aux autres échantillons. Lorsque α tend vers l'infini on obtient l'interpolation par "plus proche voisin".

3.2 Interpolation affine par triangles

Cette interpolation s'effectue en deux étapes. Tout d'abord une triangulation de Delaunay des échantillons mesurés est effectuée. Elle est obtenue par dualité à partir du diagramme de Voronoï (Figure 4.19). Dans un second temps les pixels manquants sont interpolés. Leur valeur est égale à l'intersection entre la verticale du pixel et le plan affine passant par les sommets du triangle auquel il appartient. Les deux inconvénients de cette méthode sont le temps de calcul élevé, et le manque de régularité de la fonction reconstruite (elle est affine par morceaux).

3.3 Exploitation de la contrainte de bande limitée

L'algorithme Papoulis-Gerchberg [Ferreira 94] s'applique aux fonctions à bande limitée, c'est-à-dire celles dont le support du spectre de Fourier est borné. Son principe est proche de celui de l'algorithme "alterning projection" présenté dans le chapitre précédent. Il repose lui aussi sur deux contraintes à respecter :

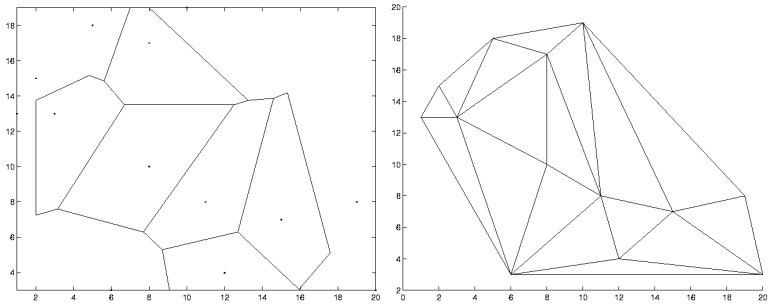


FIG. *4.19: Diagramme de Voronoï et triangulation de Delaunay associée.*

– dans le domaine fréquentiel, la contrainte de bande limitée,
– dans le domaine spatial, la correspondance avec les échantillons mesurés.

L'idée est que la présence d'échantillons à zéros induit des variations brutales dans le signal, donc des hautes fréquence (Figure 4.20(a) et 4.20(b)). Toutes les fréquences supérieures à la fréquence maximale du signal original sont dues aux zéros du sous-échantillonnage. Cet algorithme supprime les hautes fréquences afin d'éliminer ces discontinuités, tout en s'assurant de la cohérence avec les valeurs mesurées.

Soit f_n le signal obtenu à la n^{ime} itération, f_0 étant le signal mesuré, et f le signal original. Soit D l'opérateur d'échantillonnage, tel que $f_0 = Df$. Soit B un opérateur de limitation de bande (un filtre passe-bas), tel que le signal original vérifie $f = Bf$. Le signal f_{n+1} est obtenu par :

$$f_{n+1} = (I - D)Bf_n + f_0$$

La convergence est assurée par le fait que les ensembles de projection sont des ensembles convexes.

Examinons une itération sur l'exemple de la Figure 4.20 qui est un signal à une dimension sous-échantillonné de manière aléatoire (échantillonnage de Poisson). La Figure 4.20(c) montre le signal sous-échantillonné dont on a supprimé les fréquences supérieures à la fréquence maximale (connue) du signal (projection dans le domaine fréquentiel). Les basses fréquences restant inchangées, le signal possède toujours la même valeur moyenne et la même allure globale mais peine à s'écarter de cette allure globale pour rejoindre les échantillons à zéros et les échantillons initiaux.

La projection sur les échantillons observés, dans le domaine spatial (Figure 4.20(d), droite), assure quant à elle la cohérence du signal avec les échantillons mesurés. L'itération de ces deux étapes permet de converger vers le signal original.

Le choix de la fréquence de coupure, c'est-à-dire la fréquence maximale que l'on autorise au signal reconstruit est un élément décisif. Une fréquence de coupure trop basse aura pour effet de supprimer les détails de l'image. Si au contraire la fréquence de coupure est trop haute, alors le signal aura le "temps" de revenir vers zéro dans les zones où il y a peu d'échantillons, laissant ainsi des taches sombres dans ces zones.

Dans le cas d'un échantillonnage de Poisson, bien que l'on ait une fréquence d'échantillonnage moyenne donnée, la fréquence d'échantillonnage "locale" n'est pas toujours égale à la fréquence d'échantillonnage moyenne. Certaines régions ont une densité en échantillons supérieure à la moyenne, et d'autres ont au contraire une densité inférieure à la moyenne. Ainsi la

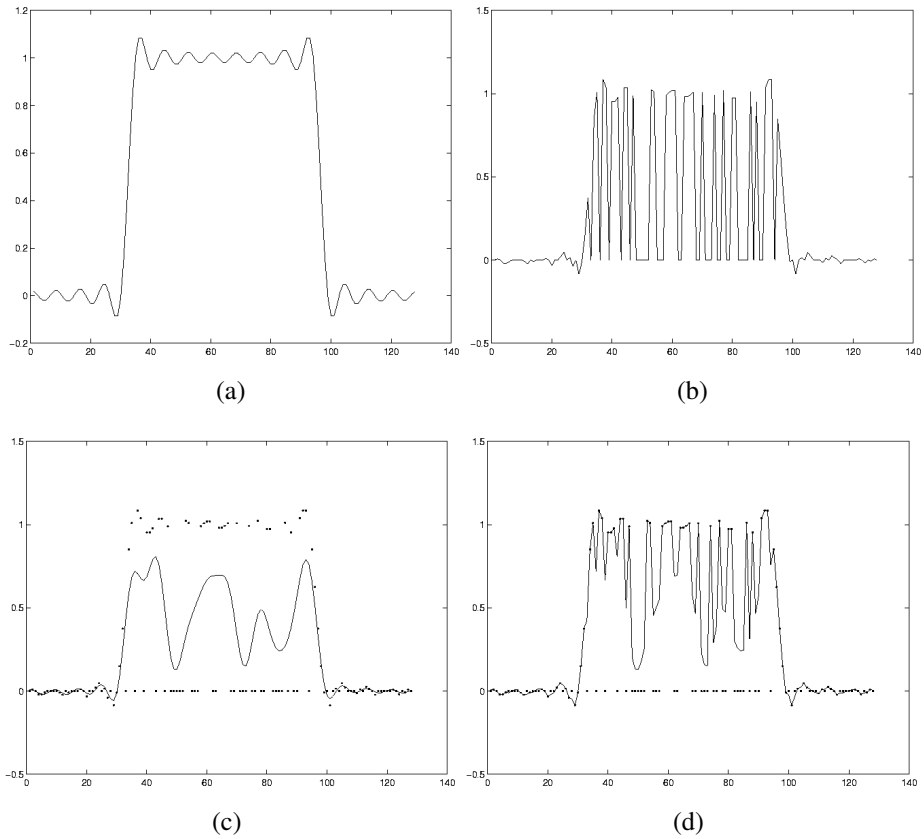


FIG. *4.20: Illustration de l'algorithme de Papoulis-Gerchberg. (a) Signal à bande limitée. (b) Le sous-échantillonnage, en mettant à zéro certains échantillons, introduit de fortes variations. (c) Résultat de la projection dans le domaine fréquentiel. (d) Résultat de la projection dans le domaine spatial sur les échantillons observés. Les itérations se font sur les étapes (c) et (d).*

fréquence d'échantillonnage "locale" varie autour de la fréquence d'échantillonnage moyenne. Ces variations de la densité locale ont pour conséquence d'introduire des taches sombres (celles évoquées plus haut) dans les zones où la densité locale est plus faible (Figure 4.21(a)). Il faudrait donc adapter la fréquence de coupure à la densité locale en échantillons. Si on considère maintenant un échantillonnage de type Poisson à disque, alors la fréquence d'échantillonnage locale est davantage homogène sur l'ensemble de l'image, c'est-à-dire plus proche de la fréquence d'échantillonnage moyenne. Ces taches disparaissent (Figure 4.21(b)). Il n'est dans ce cas pas la peine d'adapter localement la fréquence de coupure.

3.4 Convolution normalisée

La convolution normalisée est une méthode de reconstruction qui prend en compte l'*incertitude* du signal. Considérons une image $f(x)$ et un filtre passe-bas $g(x)$. Supposons que f soit échantillonnée aléatoirement. On peut associer à cette image une "carte de certitude" (*certainty map* en anglais) $c(x)$ qui vaut 1 si le pixel x correspond à la position d'un échantillon, et 0 dans le cas contraire [Knutsson 93]. Cette carte de certitude n'est autre que ce que nous appelons une "sous-mosaïque" dans le cadre du démosaïçage. La reconstruction de l'image à partir de ses échantillons se fait alors en trois étapes :

(a) (b)

FIG. *4.21: Interpolations par l'algorithme Papoulis-Gerchberg à même fréquence moyenne de sous-échantillonnage, à même fréquence de coupure du filtre de reconstruction, mais dans le cas d'un sous-échantillonnage de Poisson (à gauche, PSNR=26.7dB), de Poisson à disque (à droite, PSNR=28.0dB). La fréquence de coupure est manifestement trop haute dans le cas poissonnien, elle est correcte dans le cas poissonnien à disque.*

1. on calcule le résultat de la convolution entre l'image échantillonnée et le filtre :

$$N(x) = (f(x)c(x)) * g(x)$$

2. on calcule le résultat de la convolution entre la carte de certitude et le filtre :

$$D(x) = c(x) * g(x)$$

3. on effectue la division point à point entre les deux images obtenues, cela donne l'image reconstruite :

$$\tilde{f}(x) = \frac{N(x)}{D(x)} \tag{4.2}$$

La Figure 4.23 montre un exemple d'image échantillonnée irrégulièrement, filtrée passe-bas par un filtrage uniforme et filtrée passe-bas par une convolution normalisée. Cette méthode n'est pas une interpolation au sens strict, car les échantillons mesurés peuvent être modifiés.
Note : l'interpolation de Shepard (présentée dans un paragraphe précédent) est une convolution normalisée, dont les coefficients du filtre sont déterminés par leur distance au centre du filtre.

FIG. *4.22: A gauche : sous-échantillonnage régulier ; A droite : sous-échantillonnage irrégulier (Poisson). Dans les deux cas on a sous-échantillonné autant (1 échantillon sur 9 pour le rouge et le bleu) et la même technique d'interpolation a été appliquée. Les deux images ont un PSNR identique de 31.8dB. L'aliasing dans le cas régulier se retrouve sous forme de bruit dans le cas irrégulier.*

(a) (b) (c)

FIG. *4.23: Illustration de l'interpolation par convolution normalisée. (a) Image échantillonnée irrégulièrement. (b) Résultat d'un filtrage passe-bas uniforme. (c) Normalisation de (b) par la densité locale des échantillons.*

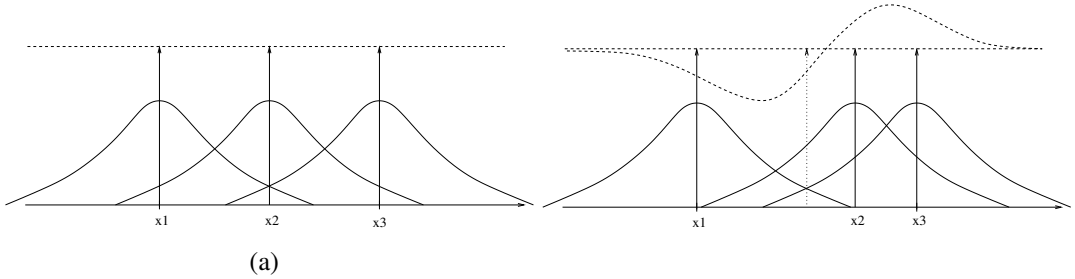


FIG. *4.24: (a) Echantillonnage régulier d'un signal constant. La somme des réponses impulsionnelles permet de retrouver le signal constant. (b) Echantillonnage irrégulier d'un signal constant. Les réponses impulsionnelles étant irrégulièrement espacées, leur somme n'est pas constante. La convolution normalisée consiste à normaliser localement ces réponses impulsionnelles en fonction de la densité locale en échantillons, afin de contrebalancer les fluctuations dans le signal reconstruit.*

Explication intuitive de la convolution normalisée

Considérons un signal f constant égal à un, unidimensionnel, et qui serait échantillonné de manière régulière. Le filtrage par convolution avec un filtre g peut s'interpréter comme la somme des réponses impulsionnelles centrées sur chaque échantillon mesuré :

$$\tilde{f}(x) = f(x) * g(x) = \sum_{i=-\infty}^{+\infty} f(x_i) \underbrace{g(x - x_i)}_{\text{noyau centré sur } x_i}$$

Considérons 3 échantillons successifs x_1, x_2 et x_3, et imaginons que l'échantillon x_2 subisse une perturbation si bien qu'il se retrouve décalé vers la droite. Alors la distance de x_1 à x_2 se trouve augmentée, et la distance de x_2 à x_3 se trouve diminuée, on est en présence d'un échantillonnage irrégulier. De la même manière qu'en régulier, un filtrage par convolution peut être appliqué. Cependant, les distances entre les échantillons étant augmentées et diminuées, la somme des réponses impulsionnelles se trouve modifiée. Elle ne sera pas égale à un signal constant, mais aura tendance à se creuser sur le segment $[x_1 x_2]$, et à se bomber sur le segment $[x_2 x_3]$ (Figure 4.24). Il convient donc d'apporter une correction à cette interpolation qui permette d'amplifier le signal si la distance aux échantillons les plus proches est grande (c'est-à-dire si la densité en échantillons est faible) ou au contraire de contenir le signal si la distance aux échantillons voisins est petite (c'est-à-dire si la densité est forte). Ceci peut se faire par un coefficient de normalisation qui dépende de la position dans le signal et qui tienne compte de la densité locale en échantillons :

$$\tilde{f}(x) = \sum_{i=-\infty}^{+\infty} f(x_i) \frac{g(x - x_i)}{p(x)}$$

où $p(x)$ est l'expression de la densité en échantillons. Si on prend comme définition de la densité, le filtrage passe-bas de la grille d'échantillonnage $p(x) = f(x) * c(x)$, alors on retrouve l'Equation (4.2).

Le noyau d'interpolation n'est plus le même pour chaque interpolation, il dépend de la densité locale. Mais attention, pour passer d'un noyau à un autre, il ne suffit pas de normaliser

globalement le noyau avec la densité en x_i, mais il faut normaliser chaque point x du noyau par la densité en x. Ceci rend impossible une normalisation des échantillons *a priori* (c'est-à-dire avant la convolution), elle doit obligatoirement se faire *a posteriori* ou pendant l'interpolation.

Un autre point de vue

On peut voir la convolution non pas comme un noyau qui se positionnerait sur chaque échantillon mesuré, mais comme un noyau qui se centre sur **l'échantillon à calculer** et qui effectue l'intégrale du produit entre le noyau et le signal échantillonné :

$$\tilde{f}(x) = \int_{\mathbb{R}} f(t)c(t)g(x-t)du$$

Les échantillons du noyau interpolateur que je qualifierai d'*utiles* sont les échantillons du noyau qui sont *en face* des positions des échantillons mesurés, c'est-à-dire là où $c(x) \neq 0$. Appelons $a(x)$ la somme de ces échantillons utiles à la postion x :

$$a(x) = \sum c(x_i)g(x-x_i) \tag{4.3}$$

$a(x)$ correspond en fait à la convolution entre g et la fonction de certitude $c(x)$. La valeur de $a(x)$ dépend de la position spatiale, puisque $c(x)$ est par hypothèse irrégulière. Cela signifie que la contribution des coefficients du filtre g va varier d'un échantillon à calculer, à l'autre. Il convient donc de normaliser le filtre à chaque position spatiale, afin que la somme de ses coefficients dits *utiles* soit constante :

$$\tilde{f}(x) = \frac{f(x) * g(x)}{a(x)} = \frac{f(x) * g(x)}{c(x) * g(x)}$$

ce qui correspond à la définition de la convolution normalisée donnée Equation (4.2).

Une différence entre ces deux points de vue vient du fait que dans un cas la normalisation de la réponse impulsionnelle est spatialement variante (chaque réponse impulsionnelle est normalisée point à point par une fonction non constante correspondant à la densité), dans l'autre cas chaque réponse impulsionnelle est normalisée par un coefficient constant correspondant à la densité en son centre.

Analogie des deux points de vue avec le concept de champ récepteur

Cette différence de points de vue concernant le processus d'interpolation, correspond exactement aux deux manières théoriques d'obtenir le champ récepteur d'un neurone. Dans un cas, on plante une électrode à la sortie du neurone dont on veut connaître le champ récepteur. On stimule alors successivement différentes parties du champ spatial et on récupére la réponse à chaque stimulus. Ceci permet de déduire l'influence de chaque point de la rétine sur le neurone considéré et donc de tracer le profil du champ récepteur. Une manière alternative de procéder (mais infaisable en pratique) serait de planter une électrode dans chaque cellule du voisinage du neurone considéré (ce neurone inclus), puis de stimuler uniquement le neurone considéré. Cela revient à faire rentrer dans le système une impulsion de Dirac, et donc de mesurer sa réponse impulsionnelle.

Une idée d'amélioration à apporter à ce type d'interpolation serait, plutôt que de normaliser localement les fonctions de base, de les étirer et les contracter localement à l'aide d'une fonction de distorsion $\gamma(x)$, comme c'est le cas pour l'interpolation de Yen utilisant le sinus cardinal (p. 76) :

$$\tilde{f}(x) = \sum_i f(x_i) g(\gamma(x - x_i))$$

En effet, une normalisation ne permet pas de décaler les zéros du sinus cardinal, seules des contractions/dilations le permettent. Cela ne devrait pas modifier grandement les résultats (car une normalisation et une distorsion permettent tous les deux d'amplifier ou de diminuer une valeur), mais cela permettrait d'obtenir de meilleurs résultats dans le cas d'un échantillonnage fortement aléatoire, avec de grandes zones sans échantillon, puisqu'alors la taille de la réponse impulsionnelle serait variable. Cela éviterait les normalisations accidentelles par zéros qui peuvent se produire lorsqu'il n'y a pas d'échantillon dans la zone délimitée par la réponse impulsionnelle.

3.5 Splines

L'interpolation utilisant le sinus cardinal comme fonction interpolatrice ne peut pas être utilisée en pratique du fait de la non compacité de son support et de sa décroissance lente. Au contraire, il est plus intéressant de chercher à reconstruire un signal à partir des translatés d'une fonction de base $\varphi(x)$ qui soit à support borné [Condat 05] :

$$\tilde{f}(x) = \sum_i c(x_i) \varphi(x - x_i) \qquad (\forall x \in \mathbb{R})$$

Les splines d'ordre n, qui sont définies comme les n convolutions successives de la fonction porte avec elle-même, vérifient cette propriété de compacité du support. La spline cubique (ordre 3) est généralement utilisée. Cependant ces fonctions de base ne sont pas des fonctions interpolatrices, elles ne valent pas 1 en 0 et 0 aux positions des autres échantillons. On parle d'interpolation *généralisée* [Condat 06]. Ainsi les $c(x_i)$ seront différents des $f(x_i)$ (les valeurs de f mesurées), et doivent être calculés lors d'une étape de pré-filtrage :

$$c(x) = \sum_i f(x_i) p(x - x_i)$$

où p est un filtre défini comme l'inverse de la version discrétisée de φ. On peut calculer la fonction interpolatrice (appelée aussi fonction cardinale [Unser 99]) qui vaut :

$$\varphi_{int}(x) = \sum_i p(x_i) \phi(x - x_i)$$

Si ϕ est une fonction interpolatrice, alors sa version discrétisée est une impulsion de Dirac, qui est son propre inverse, et il n'est pas besoin d'effectuer d'étape de préfiltrage. L'intérêt de l'interpolation généralisée est que la fonction équivalente φ_{int} peut être à support infini, tandis que la fonction de base (ou fonction *génératrice*) φ sera toujours à support compact.

Cependant, on peut ne pas obliger la fonction reconstruite à avoir la propriété $\tilde{f}(x_i) = f(x_i)$, auquel cas on ne parle plus d'interpolation mais d'approximation. Une approximation

est une condition moins stricte sur la fonction reconstruite qui peut permettre de réduire l'erreur quadratique commise [Condat 05].

Nous avons implémenté la méthode d'approximation présentée dans [Lee 97]. Cet algorithme effectue, de manière itérative, une approximation de la fonction à différentes échelles de grilles régulières, en utilisant des splines multi-échelles (*multilevel B-splines*). Chaque itération permet d'affiner l'approximation.

3.6 Interpolation orientée contour

La méthode décrite dans ce paragraphe est inspirée de l'article [Zhu 98]. Dans cet article Zhu *et al.* transposent un algorithme utilisé pour le demosaïçage dans le cas régulier. L'interpolation prend en compte les contours de l'image afin de ne pas interpoler de part et d'autre d'un contour.

La détection de contour dans le cas d'une image sous-échantillonnée aléatoirement est délicate. En effet dans le cas d'un filtre de sobel de taille 3×3 trop peu d'échantillons sont pris en compte dans la fenêtre de filtrage pour donner un résultat satisfaisant. Afin de prendre en compte un plus grand nombre de pixels, les auteurs choisissent d'utiliser des filtres de taille 3×7 et 7×3. Comme un nombre différent d'échantillons risquent d'être considérés de part et d'autre du filtre, la normalisation du filtre doit être adaptée à chaque pixel considéré. La sortie du filtrage est ensuite convertie en une image binaire par seuillage (Zhu *et al.* utilisent un seuil égal à 20%). On dispose alors de deux images binaires, une pour les contours horizontaux, une autre pour les contours verticaux. Afin de ne prendre en compte que de véritables contours et également afin de ne pas sous-estimer les contours, une ouverture morphologique est appliquée aux deux images binaires. Une ouverture morphologique est une érosion suivie d'une dilatation. L'ouverture permet d'éliminer les objets isolés plus petits que l'élément structurant. L'élément structurant choisi est un segment vertical dans le cas de l'image contenant les contours verticaux et un segment horizontal dans le cas de l'image contenant les contours horizontaux. Les auteurs conseillent une taille d'élément structurant de 10 pixels.

L'interpolation d'un pixel dépend de la présence d'un contour ou non, ainsi que du type de contour. Si un pixel n'est pas sur un contour, alors une interpolation par convolution normalisée prenant en compte la globalité du voisinage est effectuée. En présence d'un contour, l'interpolation ne se fait que dans le sens du contour.

Cette méthode peut être vue comme une amélioration de l'algorithme de convolution normalisée. Toutefois cette amélioration est tributaire de la fiabilité de la détection des contours de l'image.

3.7 Evaluation des algorithmes

Afin de comparer les performances du point de vue qualité des différentes méthodes d'interpolation, nous avons fait la moyenne des PSNR mesurés sur la base d'image de Kodak. Pour chaque image seul le plan vert est considéré, et le pourcentage d'échantillons supprimés est de 50, puis de 80. Le sous-échantillonnage utilisé est l'échantillonnage de Poisson, c'est-à-dire

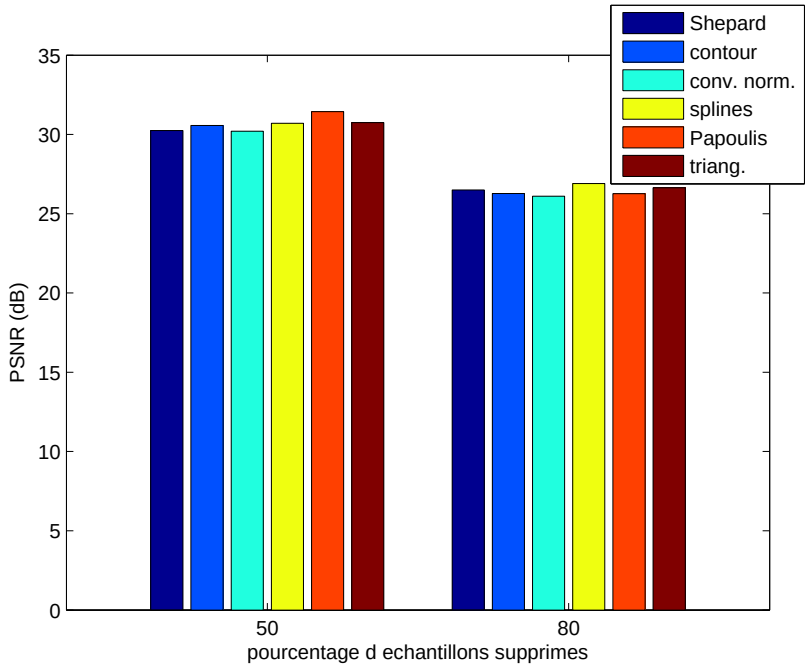


FIG. *4.25: PSNR moyen sur 23 images en niveaux de gris, pour 50 et 80% d'échantillons supprimés, pour les différents algorithmes d'interpolation.*

le cas le plus défavorable. La grille d'échantillonnage est renouvelée pour chaque image. Les résultats sont donnés figure 4.25.

Les écarts de résultats entre les différentes méthodes ne sont pas franchement significatifs. L'algorithme de Papoulis-Gerchberg qui offre les meilleurs performances en terme de PSNR à 50% d'échantillons, celui basé sur les splines est le plus efficace à 20% d'échantillons. Il faut garder à l'esprit que les valeurs concerant les splines est spécifique à l'algorithme implanté. Le choix d'un autre algorithme aurait peut-être donné de meilleurs résultats.

L'interpolation orientée contour n'est pas franchement meilleure que celle de Shepard en terme de PSNR ($0.3dB$ à 50%), les images reconstruites semblent visuellement meilleure, avec des contours nets. On remarque que cette méthode est moins bonne que l'interpolation de Shepard lorsque 80% des échantillons sont supprimés. Ceci semble cohérent puisque la détection de contour n'est pas fiable lorsqu'il y a très peu d'échantillons, et la méthode est justement tributaire du résultat de cette détection.

Compte tenu du faible écart entre les performances des différents algorithmes, on peut d'ores et déjà dire que ce ne sera pas le choix de l'un d'entre eux qui sera crucial pour le demosaïçage. Ce sera plutôt la manière dont on réussira à exploiter la corrélation entre les canaux couleurs, comme c'est le cas pour le demosaïçage de la mosaïque de Bayer. En effet, l'algorithme par sélection de fréquence n'utilise qu'une simple interpolation bilinéaire pour interpoler les données. Son efficacité réside dans l'exploitation judicieuse de la corrélation des plans couleur.

Dans la suite de ce travail de thèse, nous emploierons la convolution normalisée. Cette technique est efficace en temps de calcul - elle ne prend que le temps de la convolution - et surtout elle autorise un moyen d'action sur l'image reconstruite car on peut contrôler les caractéristiques du filtre utilisé.

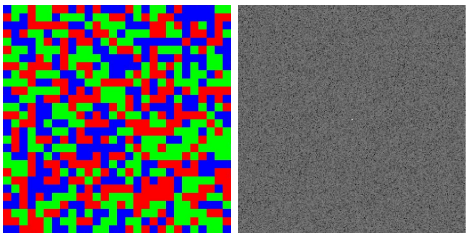
4 Génération de mosaïques aléatoires

Nous avons vu précédemment (p. 86) que pour une mosaïque de filtre couleur dont la grille globale est régulière, les propriétés d'échantillonnage des sous-mosaïques sont reportées sur la chrominance et non sur l'intensité (l'intensité est échantillonnée régulièrement). Comme nous souhaitons étudier l'influence de l'irrégulier sur l'échantillonnage de la couleur, nous avons choisi de faire des mosaïques basées sur une grille globale régulière. Cela revient à considérer que l'optique de la caméra est accordée à la fréquence d'échantillonnage globale de la mosaïque, comme c'est le cas en fovea pour la rétine. Il ne peut donc pas y avoir de repliement spectral de l'intensité avec elle même, mais uniquement du repliement entre intensité et chrominance.

Du fait de cette contrainte de grille régulière, l'échantillonnage à gigue n'a pas été utilisé. Il aurait pu être utile pour des perturbations d'amplitude inférieure au pixel (ce qui sort du cadre d'étude), pour limiter le repliement de l'intensité avec elle-même.

4.1 Mosaïque de type Poisson

Une matrice de filtres couleurs de type Poisson est une matrice dont chaque type de photo-récepteur rouge, vert et bleu suit une distribution aléatoire de type Poisson. Une approximation d'une matrice d'échantillonnage de Poisson sur une grille réguliére peut être obtenue en seuillant les valeurs d'une matrice de bruit blanc. Une matrice de bruit blanc est une matrice dont chaque élément est un tirage aléatoire suivant une loi uniforme entre zéro et un. Le niveau du seuil est choisi en fonction de la densité d'échantillons souhaitée. La TF de cet échantillonnage fait intervenir l'ensemble des fréquences, Figure 4.26(b). Cela correspond au spectre d'un bruit blanc. La présence de basses fréquences se traduit par la présences d'amas, c'est-à-dire des régions où un type de récepteurs est soit le seul présent, soit fait défaut. Cela signifie que la densité locale d'un type de photo-récepteurs est variable. La figure 4.26(a) donne un exemple de matrice de filtres de couleur d'échantillonnage de Poisson.



(a) (b)

FIG. *4.26: (a) Matrice d'échantillonnage de Poisson. On remarque la présence d'amas ; (b) Spectre d'amplitude d'un des canaux de la matrice d'échantillonnage de Poisson. Les porteuses sont diluées sur l'ensemble du spectre.*

4.2 Mosaïque de type Poisson à disque

Une matrice de type Poisson à disque est une mosaïque dont la répartition de chaque sous-mosaïque est de type Poisson à disque. La construction d'une telle matrice est un problème

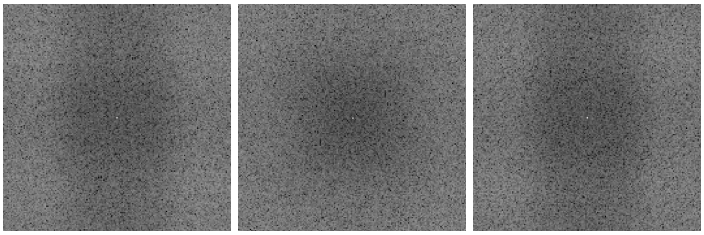


FIG. *4.27: Modules des spectres des matrices des canaux rouge, vert et bleu (de gauche à droite). Les basses fréquences sont atténuées. Deux bandes d'énergie verticales sont présentes dans le cas du rouge et du bleu. Leur disposition n'est donc pas idéale.*

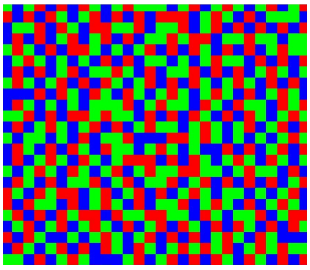


FIG. *4.28: Matrice couleur obtenue. On remarquera l'absence d'amas d'un même type de récepteurs ou encore l'absence de zones dont un type de récepteur fait défaut.*

plus complexe. En effet, si l'on ajoute les échantillons un à un (lancer de fléchettes), alors le temps de convergence est prohibitif. Ce problème a été étudié notamment dans le cadre du *half-toning* [Parker 99]. Dans le cadre des mosaïques de filtres couleur, un problème supplémentaire provient du fait que les trois sous-mosaïques doivent être mutuellement exclusives (un seul filtre par pixel). On ne peut donc pas créer de manière indépendante les trois sous-mosaïques. La méthode retenue ici est le pavage de l'image par un motif de base de taille 100x100 obtenu par "lancer de fléchettes".

La figure 4.27 montre les spectres des trois canaux RGB. La forme du spectre du vert correspond bien à celle escomptée, c'est-à-dire un dirac en zéro et un bruit dont les basses fréquences sont atténuées. Le rouge et le vert possèdent les mêmes caractéristiques, cependant le spectre n'est pas isotrope. On peut observer en effet deux bandes d'énergies verticales le long des bords droit et gauche qui sont dues à la façon dont les échantillons ont été répartis. La matrice de couleurs obtenue est visible sur la figure 4.28.

Une fois ce motif de base obtenu, il est répliqué sur l'ensemble de l'image. Cette périodisation du signal a pour conséquence de "discrétiser" son spectre. Un quadrillage apparaît alors sur le spectre. Ce quadrillage est visible à une grande échelle mais disparait si on regarde un spectre local de dimension inférieure ou égale à celle du motif répliqué.

4.3 Matrice irrégulière basée sur un motif périodique

Nous avons également généré des matrices constituées d'un motif de base de petite taille répliqué sur l'ensemble de l'image (comme c'est le cas pour les mosaïques régulières telles que Bayer). Une telle mosaïque est donc à mi chemin entre une mosaïque aléatoire et une mosaïque irrégulière, la taille du motif de base permettant de contrôler la proportion d'aléatoire introduit. Un motif de petite taille nous permet de tenter d'optimiser son arrangement afin que la disposi-

tion des échantillons soit la plus homogène possible, tout en restant irrégulière. La Figure 4.29 montre un exemple de motif de base irrégulier que nous avons optimisé manuellement. Il a pour propriété de moduler les porteuses dans les coins du spectre, comme nous le verrons dans le chapitre suivant.

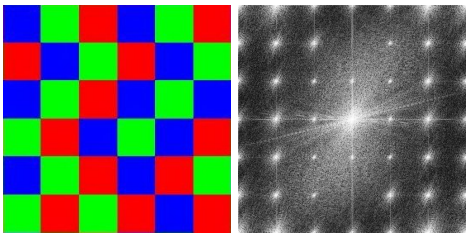


FIG. *4.29: Motif de base "irrégulier" de taille* 6×6 *pixels - la mosaïque complète est obtenue par pavage de ce motif - et spectre d'amplitude d'une image échantillonnée par cette mosaïque.*

5 Conclusion

Dans les deux cas d'échantillonnage (régulier et irrégulier), la même quantité d'énergie spectrale est repliée. C'est la répartition spectrale de ce repliement qui diffère. Les composantes fréquentielles du signal ne respectant pas le théorème d'échantillonnage sont modulées à la fréquences d'échantillonnage, créant ainsi des signaux de plus basses fréquences qu'ils ne le sont en réalité. Dans le cas de l'échantillonnage régulier, le théorème d'échantillonnage qui s'applique est le théorème classique de Shannon. C'est-à-dire que toutes les fréquences critiques sont modulées autour de la même fréquence d'échantillonnage. Par conséquent, un signal "cohérent" (présentant des régularités, des périodicités) qui est insuffisamment échantillonné restera un signal cohérent, de plus basse fréquence. C'est ce qu'on appelle le phénomène de Moiré dans le cas d'une image. Dans le cas d'un échantillonnage irrégulier, le théorème de Shannon s'applique localement, car la fréquence d'échantillonnage est variable. Les fréquences critiques sont alors modulées par une porteuse irrégulière, c'est-à-dire une porteuse large bande. Les répliquats donc répartis sur l'ensemble du spectre de Fourier. Dans le domaine spatial, ils apparaissent sous forme de bruit.

Par conséquent, l'utilisation de l'aléatoire peut être bénéfique dans toutes les applications particulièrement sensibles au phénomène de Moiré. Parmi les procédés utilisant l'aléatoire, citons par exemple le lancer de rayon stochastique dans le domaine de la synthèse graphique, le *halftoning* dans le domaine de l'impression, le *dithering* dans le domaine de la compression, etc.

Cependant, l'échantillonnage irrégulier n'est pas forcément avantageux. Du fait de la variabilité locale de la densité en échantillons, la fréquence de Nyquist locale sera tantôt supérieur et tantôt inférieure à la fréquence de Nyquist globale. Ainsi, si un signal est à la limite de la fréquence de Nyquist, tout en restant inférieur, un échantillonnage régulier pourra restituer le signal sans erreur. Au contraire, un échantillonnage irrégulier commettra des erreurs aux positions spatiales où la fréquence de Nyquist est plus basse, c'est-à-dire les positions où la densité en échantillon est plus faible. On comprend alors qu'il est très important pour un échantillonnage d'être le plus homogène possible dans sa distribution spatiale. Il faut éviter le plus possible

de moduler les réplicats au niveau des basses fréquences. D'autant plus que les images naturelles possèdent beaucoup d'énergie en basse fréquence.

De nombreux algorithmes d'interpolation de signaux échantillonnés irrégulièrement existent. Comme nous l'avons vu, aucun d'eux ne se détache vraiment des autres en terme de qualité des images reconstruites. Par conséquent, l'utilisation de l'un ou de l'autre ne sera pas crucial dans notre stratégie de démosaïçage d'une matrice irrégulière. De plus, le choix de l'un d'eux ne permet pas de résoudre notre problème. En effet, dans tous les cas, l'interpolation revient à sélectionner - de manière plus ou moins judicieuse - les basses fréquences du signal échantillonnés. Or l'intensité peut avoir de l'énergie en haute fréquences, au-delà des porteuses de chrominances. Ces algorithmes ne permettent pas de récupérer ces fréquences. Ainsi ce sera davantage la manière dont on saura exploiter la corrélation spectrale, et éventuellement spatiale, qui sera cruciale. Nous répondons à ce problème dans le chapitre suivant.

Chapitre 5

Démosaïçage généralisé et implications pour le système visuel humain

DANS ce chapitre, nous nous intéressons au démosaïçage d'une mosaïque quelconque, c'est-à-dire que celle-ci peut être régulière ou irrégulière. L'objectif est de concevoir un algorithme qui fonctionne quelque soit la mosaïque utilisée, à partir du moment où les positions des filtres de couleur sont connus. Nous nous restreignons toutefois au cas où la support de la mosaïque globale est une grille cartésienne et où la fréquence de coupure optique est en accord avec le pas d'échantillonnage de cette grille. Cette seconde restriction signifie que l'image en mosaïque peut être simulée par le sous-échantillonnage d'une image couleur dont aucun plan ne souffre de repliement spectral.

Nous effectuons dans un premier temps une comparaison des performances de différentes mosaïques à l'aide d'un algorithme linéaire de minimisation de l'erreur quadratique (*Linear Minimum Mean Square Error* LMMSE). Nous nous restreignons pour cela au cas particulier où les mosaïques sont constituées du pavage d'un motif de base. La mosaïque de Bayer, par exemple, rentre dans cette catégorie puisqu'elle est constituée d'un motif de base de taille deux pixels sur deux pixels. Nous faisons converger nos filtres par une étape d'apprentissage sur une base de d'images.

Nous tentons ensuite de construire les filtres d'estimation de intensité et de chrominance sans passer par une étape d'apprentissage. Une méthode par apprentissage nous donne certes une solution, mais il est difficile d'interpréter cette solution sans une théorie en amont. Il y a deux approches possibles au démosaïçage par passage dans l'espace intensité-chrominance : soit on construit un filtre estimateur de intensité et on en déduit la chrominance, soit on construit un filtre estimateur de chrominance et on en déduit l'intensité. Ces deux approches sont formellement équivalentes, cependant la deuxième solution offre de nombreux avantages que nous décrirons.

Nous discutons enfin des implications du modèle développé dans ce chapitre au niveau de la biologie du système visuel humain, notamment sur le lieu de la séparation de la couleur et de l'information spatiale, sur le type de connexions neuronales spécifiques nécessaires, ou encore sur les classes de cônes impliqués dans la construction des champs récepteurs des cellules ganglionnaires.

1 Modèle intensité-chrominance dans le cas d'une mosaïque quelconque

On rappelle ici le modèle intensité-chrominance, en des termes généraux, pour mettre en évidence le fait qu'il y a forcément des oppositions de couleur en haute fréquence. Il n'est en effet pas évident que lorsqu'une sous-mosaïque module l'information à une certaine fréquence, au moins une des deux autres mosaïques va également moduler à cette même fréquence, en opposition. Ainsi, quelle que soit la fréquence de la porteuse composant la mosaïque, hormis la fréquence nulle (et les multiples de la fréquence d'échantillonnage de la mosaïque globale, comme nous l'avons vu au chapitre sur l'échantillonnage irrégulier), il y aura une opposition de couleurs, car la deuxième mosaïque qui module à la même fréquence aura un effet de phase opposée.

Soit m la grille d'échantillonnage sans discrimination du type de filtre couleur. Nous nous restreignons au cas où m est une grille carrée d'impulsions de Dirac. m peut se décomposer en trois sous-mosaïques m_R, m_V et m_B, car chaque pixel de la grille est couvert d'un filtre de couleur :

$$\left(\forall (x,y) \in \{1..N\}^2\right) \qquad m(x,y) = m_R(x,y) + m_V(x,y) + m_B(x,y) \tag{5.1}$$

où $m_i(x,y)$ vaut 1 ou 0 selon que le filtre de type i ($i \in \{R,V,B\}$) est présent ou non au pixel (x,y). Nous pouvons calculer la transformée de Fourier (numérique) de la grille d'échantillonnage m et de chacune des sous-grilles m_i. m étant un signal numérique constant, sa TF est une impulsion de Dirac, tandis que la TF de la grille m_R (resp. m_V et m_B) est composée d'une composante continue r_0 (resp. v_0 et b_0) et de diverses fréquences n pondérées chacune par le coefficient r_n (resp. v_n et b_n) :

$$\begin{cases} \hat{m} & = \delta_0 \\ \hat{m}_R & = r_0\delta_0 + \sum_{n\neq 0} r_n\delta_n \\ \hat{m}_V & = g_0\delta_0 + \sum_{n\neq 0} v_n\delta_n \\ \hat{m}_B & = b_0\delta_0 + \sum_{n\neq 0} b_n\delta_n \end{cases} \tag{5.2}$$

où δ_n est l'impulsion de Dirac à la fréquence spatiale n. Par exemple pour la mosaïque de Bayer, n décrit l'ensemble de fréquences suivant [Alleysson 05] :

$$n \in \{(0,0.5);(0.5,0);(0.5,0.5)\}$$

r_0, g_0 et b_0 sont les valeurs moyennes de chaque sous-mosaïque, ou en d'autres termes, la probabilité à chaque position spatiale d'avoir un filtre de la couleur correspondante. Appelons les respectivement p_R, p_V et p_B. Par unicité de la transformée de Fourier, nous pouvons conclure (en utilisant les Equations (5.1) et (5.2)) :

$$\begin{cases} p_R + p_V + p_B = 1 \\ r_n + v_n + b_n = 0 \qquad (\forall n \neq 0) \end{cases}$$

Soit $I = \{C_R, C_V, C_B\}$ une image couleur, *i.e.* avec trois plans couleur. L'image en mosaïque résultante I_m est obtenue par :

$$I_m(x,y) = \sum_{i\in\{R,V,B\}} C_i(x,y).m_i(x,y)$$

et sa transformée de Fourier est :

$$\hat{I}_m(\nu) = \underbrace{\sum_i p_i \hat{C}_i(\nu)}_{\hat{\phi}(\nu)} + \sum_{n \neq 0} \underbrace{r_n \hat{C}_R(n-\nu) + g_n \hat{C}_V(n-\nu) + b_n \hat{C}_B(n-\nu)}_{\hat{\psi}_n(n-\nu)} \tag{5.3}$$

$\tilde{\phi}$ est une combinaison linéaire à coefficients positifs de canaux couleur, cette quantité est donc homogène à une intensité. $\tilde{\psi}_n(n-\nu)$ est une combinaison linéaire de canaux couleur dont la somme s'annule, modulée à la fréquence n. C'est donc un signal d'oppositions de couleur, ou encore de chrominance, modulé à la fréquence n. On peut conclure que, quelque soit la fréquence de la porteuse de la sous-mosaïque considérée, au moins une des deux autres sous-mosaïques contiendra une porteuse à la même fréquence et dont le coefficient sera de signe contraire.

2 Comparaison des performances des différents motifs d'échantillonnage pour un capteur couleur

S'il est évident que l'algorithme de reconstruction intervient dans la qualité d'une image reconstruite à partir de sa version en mosaïque, il est important de garder en tête que la qualité est également tributaire du type de mosaïque utilisée. Par exemple, la mosaïque à bande verticale donnera de plus mauvais résultats que la mosaïque de Bayer sur des motifs verticaux. La raison est que pour chaque colonne, la matrice à bande verticale ne possède qu'une sorte de filtre (soit rouge, soit vert, bleu) tandis que la matrice de Bayer en possède toujours deux (soit rouge et vert, soit bleu et vert).

Il semble donc intéressant de faire une étude systématique de la qualité des images couleur reconstruites en fonction de l'arrangement spatial de la mosaïque. Pour cela nous avons développé une méthode permettant de calibrer automatiquement, à partir d'une base d'images, les filtres de démosaïçage linéaire adaptés à la forme de la mosaïque [de Lavarène 07b]. Cela permet d'effectuer une évaluation objective et systématique, basée ici sur le critère de minimisation de l'erreur quadratique.

2.1 Méthode linéaire de démosaïçage par minimisation de l'erreur quadratique

2.1.1 Description de la méthode

La méthode que nous décrivons ici peut s'appliquer, en théorie, à n'importe quelle forme de mosaïque. En pratique, pour des questions de temps de calcul et de ressources, elle s'applique à des mosaïques constituées d'un motif répétitif de petite taille. Ce petit motif peut quant à lui être quelconque. Dans la description de la méthode, nous considérons le motif de Bayer. Mais ce motif peut-être remplacé par n'importe quel autre.

Reconstruction directe dans l'espace RVB

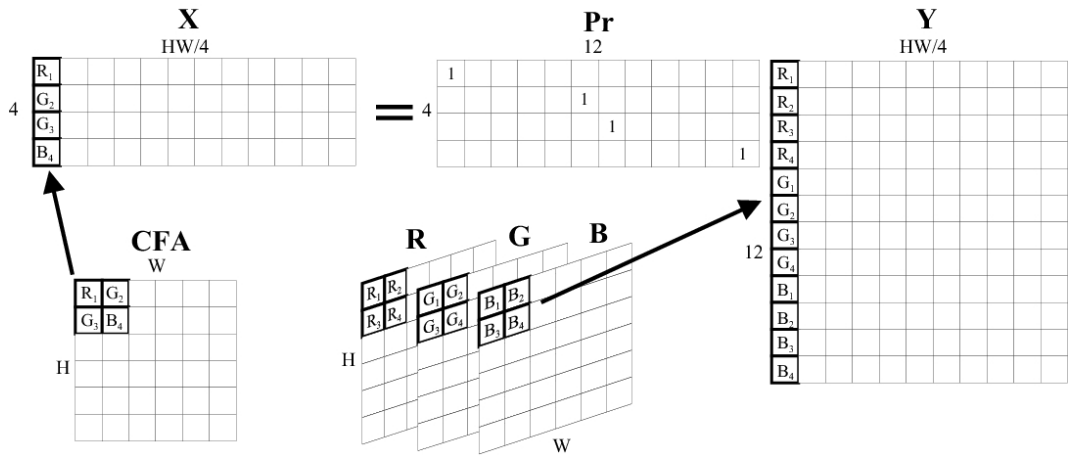


FIG. *5.1: Illustration de la conctruction de l'image en mosaïque* $\mathbf{X}$ *sous forme de produit matriciel entre une matrice d'échantillonnage* $\mathbf{Pr}$ *et l'image couleur* $\mathbf{Y}$.

Dans [Trussell 02, Taubman 00], les auteurs utilisent une notation "empilée" qui déplie l'image hyperspectrale de taille $H \times W \times P$ en un vecteur colonne de taille $HWP \times 1$, où H,W et P sont respectivement la hauteur, la largeur et le nombre de bandes spectrales de l'image. Cette notation permet d'exprimer le modèle de formation d'image sous forme de produit matriciel entre l'image hyperspectrale, une matrice de "floutage" (exprimant les propriétés du système optique), une matrice de sensibilités spectrales et une matrice d'échantillonnage. Nous considèrerons ici uniquement le cas où $\mathbf{Y}$ est une image couleur à trois composantes ($P = 3$) et $\mathbf{X}$ est une image en mosaïque :

$$\mathbf{X} = \mathbf{PrY} \tag{5.4}$$

avec $\mathbf{Pr}$ l'opérateur de projection exprimant le processus d'échantillonnage qui convertit une image à trois couleurs par pixel en une image à une couleur par pixel, c'est-à-dire une image en mosaïque.

Taubman [Taubman 00] a introduit le concept de "superpixel." Un superpixel est en ensemble connexe de pixels correspondant au motif de base dont est constituée une mosaïque (dans le cas où la mosaïque présente une périodicité). Pour la mosaïque de Bayer le motif de base est composé de quatre pixels arrangés en un carré deux par deux : une pixel rouge, deux verts et un bleu (Figure 5.1). A l'échelle du superpixel, la mosaïque est régulière, puisqu'on a un pavage du superpixel. En rajoutant l'hypothèse que le processus d'acquisition est invariant sur l'ensemble de l'image, hypothèse largement utilisée dans le domaine, nous pouvons construire des filtres de reconstruction spatialement invariants à l'échelle du superpixel, ou, autrement dit des filtres *invariants par bloc* [Hel-Or 05]. Dans Portilla *et al.* [Portilla 05], le changement de filtre à l'intérieur du superpixel se fait en estimant 4 filtres.

Cependant, tandis que Taubman et Trussell dépliaient l'image couleur $\mathbf{Y}$ sous forme de vecteur colonne $12HW \times 1$, ce qui imposait une matrice d'échantillonnage de dimension conséquente $3HW \times 12HW$, nous avons choisi de déplier le superpixel en lignes en une matrice $12 \times \frac{HW}{4}$, comme le montre le schéma de la Figure 5.1. Ceci permet l'utilisation d'une matrice d'échantillonnage $\mathbf{Pr}$ de taille 4×12, ce qui est bien plus commode à manipuler.

Dans le cas du démosaïçage "linéaire", une matrice $\mathbf{D}$ reconstruit, par produit matriciel, une

image couleur $\tilde{\mathbf{Y}}$ à partir de l'image en mosaïque $\mathbf{X}$:

$$\tilde{\mathbf{Y}} = \mathbf{D}\mathbf{X} \tag{5.5}$$

Dans un souci de faciliter la compréhension de la reconstruction de l'image, considérons dans un premier temps la reconstruction d'un seul superpixel uniquement à partir des pixels composant ce superpixel dans l'image en mosaïque. $\mathbf{Y}$ est alors un superpixel, une image de taille $2 \times 2 \times 3$, dépliée en un vecteur colonne de taille 12×1. L'image en mosaïque $\mathbf{X}$ est elle aussi un superpixel déplié, mais contient uniquement quatre valeurs. C'est un vecteur colonne de dimension 4×1. Par conséquent, la matrice reconstruction $\mathbf{D}$ sera de taille 12×4, contenant les coefficients de la reconstruction de 4 pixels à trois composantes couleur à partir de 4 pixels à une seule composante :

$$\begin{array}{ccccc} \tilde{\mathbf{Y}} & = & \mathbf{D} & & \mathbf{X} \\ \begin{bmatrix} R_1 \\ \vdots \\ R_4 \\ G_1 \\ \vdots \\ G_4 \\ B_1 \\ \vdots \\ B_4 \end{bmatrix} & = & \begin{bmatrix} D^1_{R_1} & \dots & D^4_{R_1} \\ \vdots & \ddots & \vdots \\ D^1_{R_4} & \dots & D^4_{R_4} \\ D^1_{G_1} & \dots & D^4_{G_1} \\ \vdots & \ddots & \vdots \\ D^1_{G_4} & \dots & D^4_{G_4} \\ D^1_{B_1} & \dots & D^4_{B_1} \\ \vdots & \ddots & \vdots \\ D^1_{B_4} & \dots & D^4_{B_4} \end{bmatrix} & & \begin{bmatrix} R_1 \\ G_2 \\ G_3 \\ B_4 \end{bmatrix} \end{array}$$

L'extension à tous les pixels de l'image couleur se fait naturellement, en considérant que l'image doit être dépliée selon les lignes. Il y a $\frac{H}{2}\frac{W}{2}$ superpixels dans l'image, en supposant que H et W sont pairs. Par conséquent le vecteur $\mathbf{X}$ sera de taille $4 \times \frac{HW}{4}$, et le *vecteur* couleur correspondant $\mathbf{Y}$ sera de taille $12 \times \frac{HW}{4}$, comme montré à la Figure 5.1. La matrice $\mathbf{D}$ reste quant à elle inchangée. Nous insistons sur le fait que ce dernier point est avantageux par rapport aux méthodes décrites par de [Taubman 00] et [Trussell 02] où les images étaient dépliées verticalement et impliquaient des filtres de reconstruction de taille $12HW \times 3HW$.

La prise en compte d'un plus grand noyau pour la reconstruction, un voisinage de superpixels, est obtenu en copiant le voisinage local de chaque pixel dans la direction verticale pour $\mathbf{X}$, et en étendant $\mathbf{D}$ dans la direction horizontale. Par exemple pour un voisinage de superpixel de dimension $n \times n$, $\mathbf{Y}$ sera toujours de taille $12 \times \frac{HW}{4}$ tandis que $\mathbf{X}$ sera de taille $4n^2 \times \frac{HW}{4}$ et $\mathbf{D}$ de taille $12 \times 4n^2$. Dans la suite de ce chapitre et pour simplifier les notations, les images dépliées sous forme de vecteurs seront toujours implicitement étendues verticalement avec leurs voisinages locaux.

Le produit matriciel de l'équation (5.5) effectue une combinaison linéaire des valeurs des voisinages de chaque superpixel pondérée par les coefficients du filtre $\mathbf{D}$. En d'autres termes, le produit matriciel effectue un produit de convolution (invariant par bloc) entre $\mathbf{D}$ et $\mathbf{X}$. La matrice de démosaïçage $\mathbf{D}$ peut être vue comme un ensemble de trois sous-matrices $\mathbf{D}_i$, $i \in \{R, V, B\}$ décrivant les trois plans couleur. Les trois sous-matrices correspondent respectivement aux

filtres de reconstruction de chaque plan couleur :

$$\mathbf{D} = \begin{bmatrix} \mathbf{D}_R \\ \mathbf{D}_V \\ \mathbf{D}_B \end{bmatrix}, \quad \mathbf{D}_i = \begin{bmatrix} D_{i_1}^1 & \dots & D_{i_1}^{4n^2} \\ \vdots & \ddots & \vdots \\ D_{i_4}^1 & \dots & D_{i_4}^{4n^2} \end{bmatrix}$$

Chaque ligne de la sous-matrice $\mathbf{D}_i$ correspond au filtre de reconstruction de la composante i à une position dans le superpixel dans $\mathbf{Y}$ à partir des $4n^2$ éléments du voisinage associé au superpixel dans $\mathbf{X}$.

Nous souhaitons obtenir la matrice de reconstruction minimisant l'erreur quadratique moyenne ξ avec l'image originale :

$$\xi(\mathbf{D}) = E[\|\mathbf{Y} - \tilde{\mathbf{Y}}\|^2] \tag{5.6}$$

Il est à noter que $\mathbf{D}$, (Equation (5.5)), ne peut pas être définie comme l'inverse de $\mathbf{Pr}$ puisque $\mathbf{Pr}^T\mathbf{Pr}$ est singulière. La solution classique à ce problème est la solution de Wiener donnée par :

$$\boxed{\mathbf{D} = (E[\mathbf{Y}\mathbf{X}^T])(E[(\mathbf{X}\mathbf{X}^T)])^{-1}} \tag{5.7}$$

Cette solution s'obtient par annulation de la dérivée par rapport à $\mathbf{D}$ de l'erreur quadratique (Equation (5.6)). Vérifions ce résultats en dérivant l'erreur quadratique :

$$\frac{\partial \xi(\mathbf{D})}{\partial \mathbf{D}} = E\left[2(\mathbf{Y} - \mathbf{D}\mathbf{X})\frac{\partial(\mathbf{Y} - \mathbf{D}\mathbf{X})}{\partial \mathbf{D}}\right]$$

Calculons maintenant la dérivée partielle de DX par rapport à d_{ij}. Il faut faire attention au fait que nos images dépliées sont des matrices rectangulaires et non des vecteurs comme c'est le cas classiquement. Si on fixe une colonne l dans la matrice X, le coefficient (i,l) du produit matriciel DX s'écrit :

$$[\mathbf{D}\mathbf{X}]_{il} = \sum_{k=1}^{4n^2} d_{ik} X_{kl}$$

comme seul le terme pour $k = j$ fait intervenir d_{ij}, la dérivée partielle par rapport à d_{ij} est :

$$\frac{\partial [\mathbf{D}\mathbf{X}]_{il}}{\partial d_{ij}} = X_{lj}$$

ce qui correspond à la transposée de la l^{eme} colonne de $\mathbf{X}$. Finalement, la dérivée de $\mathbf{D}\mathbf{X}$ est $\mathbf{X}^T$ et celle ξ s'écrit par conséquent :

$$\frac{\partial \xi(\mathbf{D})}{\partial \mathbf{D}} = E\left[-2\mathbf{Y}\mathbf{X}^T + 2\mathbf{D}\mathbf{X}\mathbf{X}^T\right]$$

En annulant cette dérivée, on obtient bien la solution donnée par l'Equation (5.7)

On peut calculer la matrice $\mathbf{D}$ par l'equation (5.7) en utilisant une base d'image couleur, c'est-à-dire à trois couleurs par pixel. L'utilisation d'une base de donnée signifie que nous connaissons de manière explicite$\mathbf{Y}$ et que nous pouvons simuler l'image en mosaïque $\mathbf{X}$ par l'équation (5.4). Le calcul ne requiert que l'inversion d'une matrice de taille $4n^2 \times 4n^2$ (n étant la taille du voisinage en superpixels). Ce résultat est similaire à celui de [Trussell 02] mais le

calcul se fait de manière bien plus efficace grâce à la façon dont nous déplions les images. En outre, le fait que les calculs soient effectués dans le domaine spatial plutôt que dans le domaine fréquentiel nous permet de controler la taille de la réponse impulsionnelle des filtres (c'est-à-dire la taille des matrices $\mathbf{D}_i$) et d'éviter ainsi toute troncature arbitraire de la réponse impulsionnelle. Une approche similaire a été récemment utilisée dans [Portilla 05] à travers la définition de matrices de covariance spatio-chromatique définies pour les quatre éléments du superpixel.

Reconstruction par estimation de l'intensité

Le codage en intensité et chrominance est une représentation différente de l'image couleur. $\mathbf{Y}$ peut s'écrire de manière équivalente comme la somme de sa intensité $\mathbf{\Phi}_c$ et de sa chrominance $\mathbf{\Psi}_c$: $\mathbf{Y} = \mathbf{\Phi}_c + \mathbf{\Psi}_c$. Si nous appelons respectivement $\mathbf{P}_c$ et $\mathbf{M}_c$ les estimateurs linéaires de l'intensité et la chrominance, en utilisant les notations de la figure 5.1, nous pouvons écrire :

$$\mathbf{Y} = \mathbf{P}_c\mathbf{Y} + \mathbf{M}_c\mathbf{Y}$$

avec $\mathbf{P}_c + \mathbf{M}_c = \mathbf{I}_{12}$ afin que la décomposition soit conservative ($\mathbf{I}_{12}$ étant la matrice identité de taille 12×12). Définir une décomposition intensité-chrominance revient à définir les matrices $\mathbf{P}_c$ et $\mathbf{M}_c$. l'intensité $\mathbf{\Phi}_c$ est généralement définie comme une pondération positive des valeurs R, V et B à chaque pixel. Nous supposons que les coefficients de pondérations, appelés p_i (avec $i \in \{R, V, B\}$) sont positifs et que leur somme vaut 1, ou encore, que l'intensité est un barycentre des valeurs R, V et B.

De plus, l'intensité est achromatique, elle ne contient pas d'information d'opposition chromatique. Elle a donc une seule valeur par position spatiale. Elle peut être indifféremment représentée par un scalaire en niveaux de gris ou par un vecteur couleur ayant trois valeurs d'intensité identiques. La partie chromatique $\mathbf{\Psi}_c$ de l'image est alors composée d'un vecteur à trois composantes, donnant les différences des plans R, V et B au plan d'intensité. Puisque la somme des p_i est égale à un, la somme des composantes chromatiques s'annule à chaque pixel. Finalement, la définition de l'intensité et de la chrominance se résume à :

$$\begin{aligned} \mathbf{P}_c &= [\ 1 \quad 1 \quad 1\ ]^T \otimes \underbrace{[\ p_R \quad p_G \quad p_B\ ] \otimes \mathbf{I}_4}_{\mathbf{P}} \\ \mathbf{M}_c &= \mathbf{I}_{12} - \mathbf{P}_c = (\mathbf{I}_3 - [\ 1 \quad 1 \quad 1\ ]^T \otimes [\ p_R \quad p_G \quad p_B\ ]) \otimes \mathbf{I}_4 \end{aligned} \tag{5.8}$$

où $\otimes$ est le produit de Kronecker, ou produit tensoriel. Pour deux matrices $A = (a_{ij})_{nm}$ et B, leur produit de kronecker se définit par :

$$A \otimes B = \begin{pmatrix} a_{11}B & \cdots & a_{1m}B \\ \vdots & \ddots & \vdots \\ a_{n1}B & \cdots & a_{nm}B \end{pmatrix}$$

Avec cette notation, $\mathbf{P}_c$ est une matrice de dimension 12×12 composée de la sous-matrice $\mathbf{P}$ répliquée trois fois dans la direction verticale. L'estimateur de chrominance est une matrice de dimension 12×12. Puisque la somme de ses trois composantes pondérées par p_i s'annule

à chaque position spatiale, la chrominance est intrinsèquement un signal à deux dimensions, comme définie dans les standards usuels.

Maintenant que nous avons défini la décomposition en intensité et chrominance d'une image couleur, intéressons nous au cas d'une image en mosaïque. Une image en mosaïque peut être vue comme une image à niveau de gris (comme la matrice $\mathbf{X}$), c'est-à-dire avec une seule composante par position spatiale, bien que cette composante corresponde alternativement à l'une au l'autre composante couleur. Ceci est illustré par le fait que $\mathbf{Pr}$ peut être décomposé en une somme de trois opérateurs de sous-échantillonnage $\mathbf{m}_i$ opérant sur les superpixels de chaque plan couleur comme suit :

$$\mathbf{Pr} = [\ 1 \quad 0 \quad 0\ ] \otimes \mathbf{m}_R + [\ 0 \quad 1 \quad 0\ ] \otimes \mathbf{m}_V + [\ 0 \quad 0 \quad 1\ ] \otimes \mathbf{m}_B \tag{5.9}$$

avec

$$\begin{cases} \mathbf{m}_R = diag([\ 1 \quad 0 \quad 0 \quad 0\ ]) \\ \mathbf{m}_V = diag([\ 0 \quad 1 \quad 1 \quad 0\ ]) \\ \mathbf{m}_B = diag([\ 0 \quad 0 \quad 0 \quad 1\ ]) \end{cases}$$

où $diag$ correspond à la matrice diagonale, ayant son argument comme diagonale. Ici, l'opérateur "somme" agit comme le multiplexeur, les plans couleur étant déjà sous-échantillonnés par les fonctions $\mathbf{m}_i$. Une image en mosaïque peut être représentée, de manière équivalente, sous forme d'une image couleur $\mathbf{X}_c$ dans laquelle la plupart des valeurs sont égales à zéro puisque les deux tiers des composantes chromatiques sont absentes :

$$\mathbf{X}_c = \mathbf{Pr}^T\mathbf{PrY} = \mathbf{Pr}^T\mathbf{X}$$

Le terme $\mathbf{Pr}^T\mathbf{Pr}$ correspond à une matrice de sous-échantillonnage 12×12 ayant seulement quatre valeurs égales à un (dans le cas de Bayer). Elle permet de sélectionner dans l'image couleur $\mathbf{Y}$ les pixels correspondant à l'image en mosaïque $\mathbf{X}$ et remplit le reste par des zéros. Comme montré dans le membre de droite de l'équation, $\mathbf{Pr}^T$ agit comme un démultiplexeur puisqu'il isole chacun des plans couleur de la mosaïque.

Nous pouvons maintenant réécrire $\mathbf{Pr}$ en terme d'estimateurs d'intensité et de chrominance. Chaque matrice d'échantillonnage $\mathbf{m}_i$ peut-être décomposée comme suit : $\mathbf{m}_i = p_i\mathbf{I}_4 + \mathbf{m}_i - p_i\mathbf{I}_4$. Par conséquent, d'après l'équation 5.9, $\mathbf{Pr}$ peut être réécrit :

$$\mathbf{Pr} = [\ p_R \quad p_G \quad p_B\ ] \otimes \mathbf{I}_4 + \mathbf{Pr} - [\ p_R \quad p_G \quad p_B\ ] \otimes \mathbf{I}_4 \tag{5.10}$$

Par identification entre les équations 5.8 et 5.10 nous avons :

$$\begin{aligned} \mathbf{Pr} &= \mathbf{P} + \mathbf{M} \\ &= \mathbf{P} + \mathbf{Pr}(\mathbf{I}_3 \otimes \mathbf{I}_4) - \mathbf{Pr}([\ 1 \quad 1 \quad 1\ ]^T \otimes [\ p_R \quad p_G \quad p_B\ ] \otimes \mathbf{I}_4) \\ &= \mathbf{P} + \mathbf{PrM}_c \end{aligned}$$

Le projecteur $\mathbf{Pr}$ est la somme d'un estimateur d'intensité $\mathbf{P}$ définit comme pour les images couleur (Equation 5.8) et d'un projecteur de chrominance $\mathbf{M}$, qui est la projection de l'estimateur de chrominance $\mathbf{M}_c$ défini pour les images couleur (Equation 5.8). Cela montre que le sous-échantillonnage et la projection dus à la mosaïque de filtres n'affectent pas la composante

d'intensité, mais sont reportés sur la chrominance. Il s'en suit que la chrominance démultiplexée contenue dans l'image en mosaïque est égale à la chrominance sous-échantillonnée d'une image couleur (i.e. $\mathbf{Pr}^T\mathbf{M} = \mathbf{Pr}^T\mathbf{Pr}\mathbf{M}_c$). Une manière de retrouver la chrominance est donc de démultiplexer la composante de chrominance et de l'interpoler.

Nous avons défini jusqu'à présent des estimateurs d'intensité et de chrominance à partir d'une image couleur telle qu'elle serait contenue dans une image en mosaïque. Cependant, nos estimateurs dépendent des coefficients p_i pour lesquels nous n'avons comme contrainte seulement qu'ils soient positifs et que leur somme soit égale à 1. Nous pouvons de plus contraindre les paramètres p_i afin d'assurer que l'estimateur de chrominance s'annule sur un superpixel (donc sur l'ensemble de l'image). Cela garantit que l'intensité et la chrominance auront deux supports fréquentiels disjoints, puisque dans ce cas la chrominance ne sera pas modulée en bande de base. Cela revient à annuler la somme des trois diagonales de la matrice $\mathbf{M}$, soit, pour la mosaïque de Bayer :

$$\begin{cases} 1 - 4p_R = 0 \\ 2 - 4p_G = 0 \\ 1 - 4p_B = 0 \end{cases}$$

ce qui donne

$$\{p_R = \frac{1}{4}, \quad p_G = \frac{1}{2}, \quad p_B = \frac{1}{4}\}$$

Nous utiliserons donc ces coefficients pour définir l'intensité et la chrominance contenue dans une image échantillonnée par la matrice de Bayer.

Plutôt que d'estimer directement l'image couleur à partir de l'image en mosaïque comme précédemment, nous pouvons dans un premier temps estimer l'intensité $\tilde{\mathbf{\Phi}}$:

$$\tilde{\mathbf{\Phi}} = \mathbf{H}_\Phi \mathbf{X} \tag{5.11}$$

$\mathbf{H}_\Phi$ étant le filtre d'intensité. Une fois l'intensité estimée, nous retrouvons la chrominance modulée par la différence entre l'image en mosaïque et l'intensité $\tilde{\mathbf{\Psi}} = (\mathbf{X} - \tilde{\mathbf{\Phi}})$. Comme suggéré, nous démultiplexons la chrominance en la multipliant par $\mathbf{Pr}^T$ avant interpolation pour obtenir la chrominance pleinement échantillonnée $\tilde{\mathbf{\Psi}}_\mathbf{c}$:

$$\tilde{\mathbf{\Psi}}_\mathbf{c} = \mathbf{H}_\Psi \mathbf{Pr}^T \tilde{\mathbf{\Psi}}$$

où $\mathbf{H}_\Psi$ est la matrice contenant les trois filtres interpolateurs de chrominance. Finalement l'image couleur reconstruite $\tilde{\mathbf{Y}}$ sera la somme des deux composantes :

$$\tilde{\mathbf{Y}} = \tilde{\mathbf{\Phi}}_c + \tilde{\mathbf{\Psi}}_c \tag{5.12}$$

où $\tilde{\mathbf{\Phi}}_c = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}^T \otimes \tilde{\mathbf{\Phi}}$.

Contrairement à la reconstruction directe dans l'espace RVB où une seule phase d'apprentissage suffisait, nous avons ici deux étapes d'apprentissage afin d'obtenir respectivement :

– l'estimateur d'intensité, calculé à partir de l'image en mosaïque $\mathbf{X}$ (laquelle est simulée à partir de la base d'images en mettant à zéro les composantes chromatiques appropriées) et à partir de l'intensité $\mathbf{\Phi}$ (laquelle est également calculée à partir de la base d'images) :

$$\mathbf{H}_\Phi = (E[\mathbf{\Phi}\mathbf{X}^T])(E[(\mathbf{X}\mathbf{X}^T)])^{-1}$$

– l'interpolateur de chrominance, calculé à partir de la chrominance $\mathbf{\Psi_c}$ et de la chrominance sous-échantillonnée $\mathbf{\Psi}$ (calculées à partir de la base d'image) :

$$\mathbf{H}_\Psi = (E[\mathbf{\Psi_c}(\mathbf{Pr}^T\mathbf{\Psi})^T])(E[(\mathbf{Pr}^T\mathbf{\Psi})(\mathbf{Pr}^T\mathbf{\Psi})^T)])^{-1}$$

avec $\mathbf{\Phi} = \mathbf{PY}$, $\mathbf{\Psi}_c = \mathbf{M}_c\mathbf{Y}$ and $\mathbf{\Psi} = \mathbf{X} - \mathbf{\Phi}$.

Le grand avantage de cette décomposition est que la chrominance a une largeur de bande très étroite en comparaison des fréquences de Nyquist de chaque sous-mosaïque. Des filtres d'ordres très faibles sont donc suffisants pour l'interpolation. Au contraire, l'estimateur d'intensité doit avoir un gabarit très contraint aux fréquences situées à la frontière entre l'information d'intensité et l'information de chrominance modulée. Un filtre d'ordre élevé est donc requis pour l'estimation de l'intensité (typiquement 7×7 ou 9×9). Cependant, un seul filtrage d'ordre élevé est nécessaire, contrairement à la méthode par reconstruction directe qui nécessite trois filtrages d'ordre élevé, ce qui rend l'algorithme bien plus efficace d'un point de vue computationnel.

Lien entre la méthode RVB et la méthode par intensité-chrominance

Dans la méthode par estimation directe des valeurs RVB, les filtres effectuent deux combinaisons linéaires optimales (au sens des moindres carrés) du voisinage, à chaque pixel, afin de retrouver les deux valeurs chromatiques manquantes. Dans la méthode utilisant l'information d'intensité, notre filtre estime dans un premier temps la combinaison linéaire du voisinage permettant d'estimer la meilleure intensité possible, puis les filtres de chrominance effectuent l'interpolation des valeurs manquante de chrominance. Nous pouvons montrer que ces deux méthodes sont formellement équivalentes.

Supposons que l'on connaisse le filtre RVB $\mathbf{D} = \begin{bmatrix} \mathbf{D}_R^T & \mathbf{D}_G^T & \mathbf{D}_B^T \end{bmatrix}^T$, $\mathbf{D}_i$ étant le filtre spatialement variant permettant de calculer le plan couleur i. l'intensité étant définie comme la pondération des trois plans couleur, ces filtres peuvent être également utilisés pour estimer l'intensité :

$$\tilde{\mathbf{\Phi}} = \sum_i p_i(\mathbf{D}_i\mathbf{X}) = \left\{\sum_i p_i\mathbf{D}_i\right\}\mathbf{X}$$

Par conséquent le filtre d'intensité s'obtient par :

$$\mathbf{H}_\phi = \sum_i p_i\mathbf{D}_i$$

Nous pouvons de plus en déduire la chrominance :

$$\tilde{\mathbf{\Psi}} = \mathbf{H}_\Psi\mathbf{Pr}^T(\mathbf{X} - \tilde{\mathbf{\Phi}}) = \tilde{\mathbf{Y}} - \tilde{\mathbf{\Phi}}_\mathbf{c}$$

En remplaçant $\tilde{\mathbf{\Phi}}$ et $\tilde{\mathbf{Y}}$ par leurs expressions avec $\mathbf{D}$, nous obtenons les filtres interpolateurs de la chrominance :

$$\mathbf{H}_\Psi = \{\mathbf{D_i} - \mathbf{H}_\phi\}\{\mathbf{I}_4 - \mathbf{H}_\phi\}^{-1}$$

Supposons maintenant que nous connaissons le filtre d'estimation de l'intensité $\mathbf{H}_\phi$ ainsi que les filtres interpolateurs de la chrominance $\mathbf{H}_\mathbf{\Psi}$. Est-il possible d'obtenir les filtres d'inter-

polation directe RVB ? L'image couleur $\mathbf{Y}$ est obtenue par :

$$\begin{aligned}\tilde{\mathbf{Y}}_i &= \tilde{\mathbf{\Phi}}_\mathbf{c} + \mathbf{H}_\Psi \mathbf{Pr}^T(\mathbf{X} - \tilde{\mathbf{\Phi}}) \\ &= \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}^T \otimes \mathbf{H}_\Phi X + \mathbf{H}_\Psi \mathbf{Pr}^T(\mathbf{X} - \mathbf{H}_\Phi \mathbf{X}) \\ &= \left\{ \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}^T \otimes \mathbf{H}_\Phi + \mathbf{H}_\Psi \mathbf{Pr}^T(\mathbf{I_4} - \mathbf{H}_\Phi) \right\} \mathbf{X}\end{aligned} \tag{5.13}$$

Nous concluons :

$$\mathbf{D} = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}^T \otimes \mathbf{H}_\Phi + \mathbf{H}_\Psi \mathbf{Pr}^T(\mathbf{I_4} - \mathbf{H}_\Phi) \tag{5.14}$$

ce qui prouve que nous pouvons obtenir les filtres RVB étant donnés les filtres d'intensité-chrominance. Finalement, nous avons démontré que les deux méthodes sont équivalentes.

2.1.2 Application au cas de la mosaïque de Bayer et validation du modèle intensité-chrominance

Avant d'appliquer la méthode à diverses mosaïques, intéressons nous au cas de la mosaïque de Bayer afin de voir si nous retrouvons les résultats d'Alleysson *et al.* [Alleysson 05], à savoir, un filtre d'intensité qui atténue les hautes fréquences dans les directions horizontale, verticale et diagonale. Le résultat de l'apprentissage de la matrice $\mathbf{D}$ pour la méthode RVB est donné à la Figure 5.2, et le résultat pour la méthode par intensité-chrominance est donné Figure 5.3 (seul le filtre d'intensité est représenté, car les filtres de chrominance sont très proches des filtres bilinéaires). Sur ces figures sont présentés les spectres d'amplitude des filtres obtenus pour chaque position dans le superpixel. L'apprentissage a été effectué en utilisant la méthode "*leave-one-out*", c'est-à-dire que l'image testée ne fait jamais partie de la base d'apprentissage. On observe que le filtre d'intensité obtenu est spatialement variant (Figure 5.3), contrairement à la solution proposée Par Alleysson *et al.*. En effet, si l'on retrouve la même forme de filtre pour les pixels R et B avec une atténuation des porteuses de chrominance, le filtre sur le pixel G quant à lui diffère. Ce dernier n'atténue pas les fréquences horizontales et verticales. Le filtre d'intensité proposé p. 60 est donc sous-optimal . Ainsi cette simulation nous permet de constater que le modèle est valide, mais que la solution qui en est dérivée n'est pas optimale. Cette variation spatiale du filtre d'intensité avait été proposée par Lian *et al.* [Lian 05], cependant leur justification est peu convaincante. Il est difficile d'expliquer à ce stade du discours la raison de cette forme de filtre, mais une explication sera donnée dans la prochaine section.

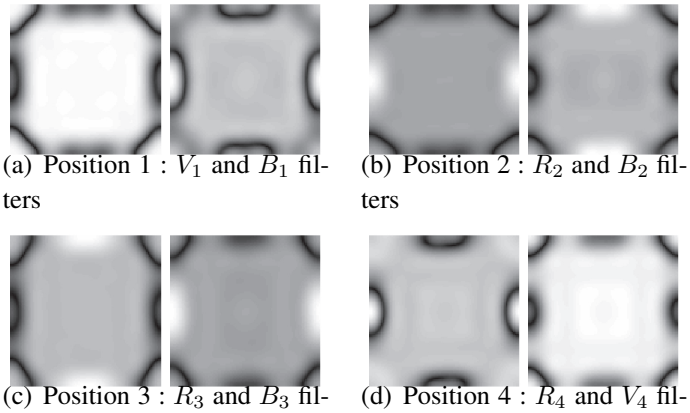
Par ailleurs, le filtre correspondant aux pixels G étant moins contraint (il atténue uniquement les coins du spectres), il est possible de réduire son ordre (et donc son efficacité algorithmique) sans perte de qualité. Typiquement, une filtre 9×9 peut être réduit à un ordre 5×5.

Nous constatons également que les deux méthodes, par estimation directe dans l'espace RVB ou par passage par l'intensité, donnent les mêmes performances (Table 5.1), ce qui va dans le sens de la démonstration du paragraphe précédent montrant l'équivalence mathématique entre les deux méthodes. Par ailleurs, la forme fréquentielle des filtres RVB (Figure 5.2) est difficilement interprétable sans le modèle d'échantillonnage intensité-chrominance. Grâce à ce modèle nous pouvons interpréter la large composante passe-bas de ces filtres comme étant la sélection de la composante d'intensité. Une composante haute-fréquence est présente ou non dans les coins et le milieu des côtés du spectre, selon la position dans le superpixel (le filtre est donc spatialement variant) afin de démoduler la composante de chrominance.

	R	G	B	ΔE	Complexité
RVB 7x7	38.15 ($\pm$2.55)	40.88 ($\pm$2.41)	36.86 ($\pm$2.55)	1.15 ($\pm$0.34)	98HW
RVB 9x9	38.87 ($\pm$2.56)	41.43 ($\pm$2.42)	37.29 ($\pm$2.58)	1.00 ($\pm$0.34)	162HW
Lum 7x7	38.40 ($\pm$2.49)	40.88 ($\pm$2.37)	36.74 ($\pm$2.47)	1.18 ($\pm$0.34)	47HW
Lum 9x9	38.92 ($\pm$2.52)	41.34 ($\pm$2.36)	37.21 ($\pm$2.51)	1.01 ($\pm$0.33)	63HW

TAB. *5.1: PSNR et ΔE moyens, et écart-type sur l'ensemble de la base d'images pour les deux méthodes (par estimation directe et par estimation de l'intensité) appliquées à la mosaïque de Bayer, pour des tailles de noyau 7×7 et 9×9.*

Au niveau de la complexité algorithmique, la méthode par estimation de l'intensité est bien plus efficace, puisqu'elle ne nécessite qu'un filtre d'ordre élevé (par exemple un noyau 9×9, soit 81 opérations par pixel) plutôt que deux dans le cas de la méthode RVB (deux filtres car deux valeurs sont calculées à chaque position spatiale, soit 162 opérations par pixel pour un filtre 9×9). *A fortiori* dans le cas de la mosaïque de Bayer, puisque sur les pixels G, l'ordre peut être réduit sans perte de qualité. La complexité indiquée Table 5.1 tient compte de cette optimisation. Au coût de l'étape d'estimation de la luminance, il faut ajouter le coût de l'interpolation bilinéaire de la chrominance (6 op/pixel) et celui de la soustraction et de l'addition de la luminance (4 op/pixel).



(a) Position 1 : V_1 and B_1 filters (b) Position 2 : R_2 and B_2 filters

(c) Position 3 : R_3 and B_3 filters (d) Position 4 : R_4 and V_4 filters

FIG. *5.2: Spectres d'amplitude des filtres estimés pour la reconstruction directe à chaque position dans le superpixel de Bayer (les filtres sont normalisés pour l'affichage). Les filtres diffèrent de l'un à l'autre par leurs fréquences de coupure. Les caractéristiques fréquentielles de ces filtres sont expliquables par la décomposition en intensité-chrominance.*

FIG. *5.3: Spectre d'amplitude du filtre d'intensité pour chaque position (1, 2, 3 et 4) dans le superpixel de Bayer. Aux positions 2 et 3 (pixels verts), l'intensité peut être estimée avec une acuité horizontale et verticale maximale.*

	R	G	B	Moyenne
Bayer	38.53 (±2.67)	41.22 (±2.47)	37.25 (±2.59)	39.00 (±2.58)
Bandes verticales	34.50 (±2.81)	34.61 (±2.82)	34.50 (±2.69)	34.54 (±2.77)
Bandes horizontales	32.95 (±2.46)	33.09 (±2.48)	33.16 (±2.26)	33.07 (±2.40)
Bandes diagonales	38.17 (±2.48)	38.84 (±2.42)	38.20 (±2.59)	38.40 (±2.50)
Lukac	38.69 (±2.45)	41.24 (±2.37)	38.25 (±2.51)	39.39 (±2.44)
Lukac (tournée)	38.50 (±2.42)	40.96 (±2.34)	38.07 (±2.53)	39.18 (±2.43)
Pseudo-random	38.90 (±2.50)	40.12 (±2.41)	39.44 (±2.67)	**39.49** (±2.53)
Pseudo-random (tournée)	38.87 (±2.49)	40.16 (±2.40)	39.51 (±2.64)	**39.51** (±2.51)

TAB. *5.2: Valeurs de PSNR (dB) et écarts-types entre image originale et image reconstruite, moyennées sur les 24 images de la base Kodak, pour différentes configurations de mosaïques.*

	PSNR		ΔE S-CIELab	
Optique circulaire	sans	avec	sans	avec
Bayer	39.00	39.66	1.06	1.01
Proposée	**39.49**	**40.90**	**1.01**	**0.89**

TAB. *5.3: Valeurs de PSNR et de ΔE (dans l'espace S-CIELab) moyennées sur les 24 images de la base Kodak, pour la mosaïque de Bayer et la mosaïque pseudo-irrégulière proposée. Les chifres montrent que le gain en qualité apporté par la mosaïque proposée est plus important lorsqu'une optique circulaire est simulée.*

2.2 Evaluation des performances des différentes mosaïques

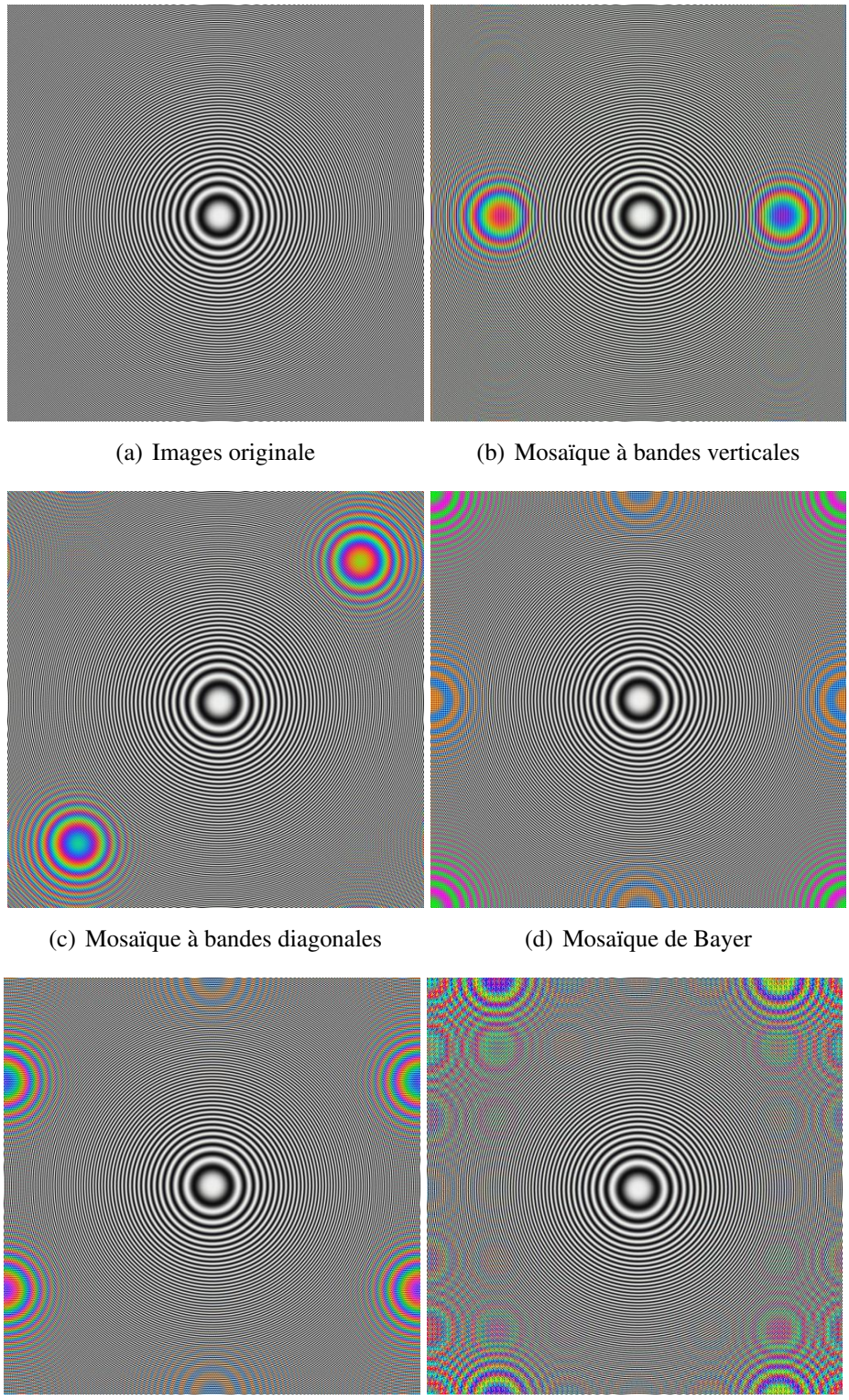
La méthode LMMSE donne les filtres de démosaïçage linéaires optimaux au sens des moindres carrés, et ce quelque soit le motif du capteur. Il nous permet donc de comparer les performances des différents capteurs. Lukac a proposé une méthode adaptative s'appliquant à n'importe quelle mosaïque [Lukac 05a] et effectue un comparatif des performances de différentes mosaïques [Lukac 05a]. Le fait que notre algorithme soit linéaire le rend plus "objectif", car il ne sera pas biaisé par certains motifs dans l'image.

Les résultats des simulations sur la base d'images sont donnés dans la Table 5.2 en terme de PSNR. On constate que toutes les mosaïques ne présentent pas la même qualité d'image. Comme nous pouvions nous y attendre, la mosaïque de Bayer donne de meilleurs résultats que les mosaïques basées sur des bandes. Cependant, la mosaïque de Lukac ainsi que la mosaïque pseudo-irrégulière (proposée à la fin du chapitre précédent) donnent toutes les deux des résultats supérieurs à ceux de la mosaïque de Bayer. La mosaïque irrégulière a un gain de 0.5dB de moyenne par rapport à Bayer, et la mosaïque de Lukac un gain de 0.2dB ou 0.4dB selon son orientation.

Nous avons également fait des simulations sur l'image "Circular Zone Plate" (CZP), dont les résultats sont montrés à la Figure 5.4. Cette image achromatique permet de bien faire apparaître les défauts de fausses couleurs. Elle est générée par une sinusoïde de révolution, modulée en fréquence par la position spatiale dans l'image, avec l'origine des axes pris au centre de l'image :

$$(x, y \in [-256, 256]^2) \qquad CZP(x, y) = 128 + 128 * sin(\frac{\pi}{512}(x^2 + y^2)) \tag{5.15}$$

La particularité de cette image est que son contenu fréquentiel au voisinage d'un point (x,y) dans



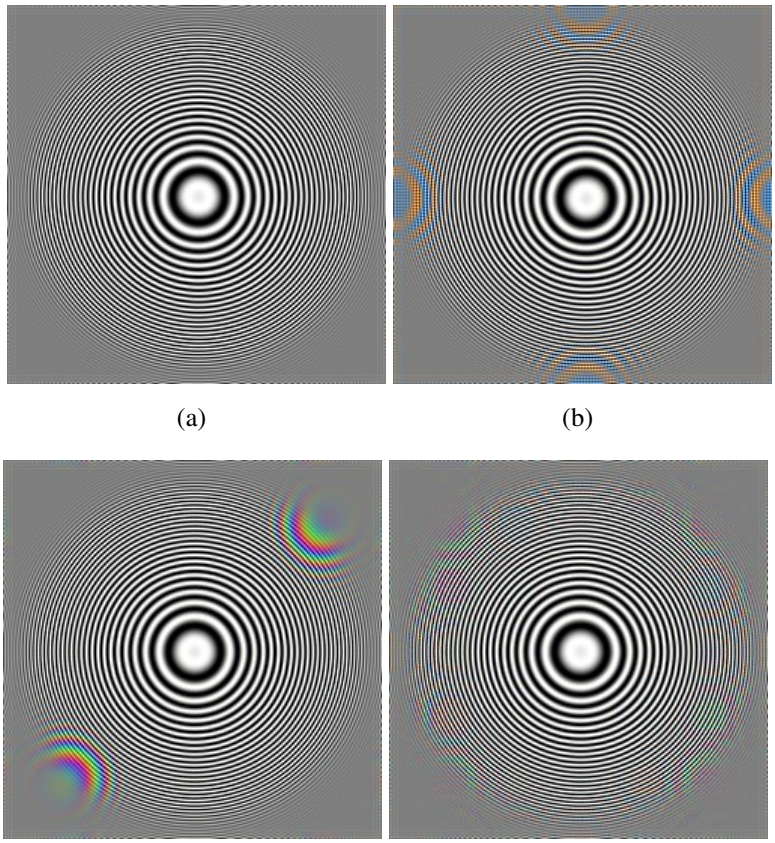
(a) Images originale (b) Mosaïque à bandes verticales

(c) Mosaïque à bandes diagonales (d) Mosaïque de Bayer

(e) Mosaïque de Lukac (f) Mosaïque pseudo-aléatoire

FIG. *5.4: Démosaïçage de l'image CZP. (a) image CZP originale ; (b-f) images reconstruites après sous-échantillonnage à travers différentes mosaïques.*

le domaine spatial correspond à la fréquence (x,y) dans le domaine de Fourier. On constate que l'on obtient bien des fausses couleurs aux positions des porteuses de chrominance telles qu'on



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

FIG. *5.5: Démosaïçage de l'image CZP (256x256 pixels) avec simulation d'une optique circulaire passe-bas. (a) Image originale (b-d) Reconstruction à partir de la mosaïque de Bayer, diagonale à bandes et de la mosaïque irrégulière proposée.*

peut les observer sur les spectres d'amplitude des mosaïques, excepté pour la mosaïque pseudo-irrégulière. En effet, pour cette dernière, seul le contenu fréquentiel local importe, alors que la transformée de Fourier nous donne l'information sur le contenu fréquentiel de l'image prise dans sa globalité. Ainsi des porteuses localisées aux très basses fréquences apparaissent sur le spectre (en fait tous les multiples de $1/n$, n étant la taille du superpixel), mais elles n'interviennent pas dans le problème du démosaïçage. Ainsi, cette image nous donne des informations sur les fréquences spatiales qui causent effectivement du recouvrement spectral entre l'intensité et la chrominance.

Tentons d'expliquer pourquoi les deux mosaïques (irrégulière et Lukac) donnent de meilleurs résultats que la mosaïque de Bayer. Deux propriétés majeures interviennent dans la quantité de recouvrement entre intensité et chrominance, et donc dans la qualité de l'image reconstruite : 1) les fréquences des porteuses de chrominances, 2) leurs orientations. La chrominance doit être, naturellement, modulée à des fréquences les plus éloignées possible du centre du spectre. Cependant les directions horizontales et verticales sont les orientations les plus sensibles aux fausses couleurs dans les images naturelles [Dubois 05]. Ainsi, si l'on considère la distance au centre, les mosaïques à bandes présentent les plus mauvaise caractéristiques, puisqu'elles modulent au tiers de la fréquence de Nyquist. Au contraire, les mosaïques de Bayer et de Ny-

quist modulent la chrominance aux fréquences $\frac{1}{2}$, optimisant ainsi ce critère. Cependant, la mosaïque de Lukac préserve totalement une des deux directions horizontale ou verticale (selon l'orientation du motif), tandis que la mosaïque de Bayer module dans les deux directions. La mosaïque de Lukac est donc meilleure que celle de Bayer si l'on considère ce second critère, et c'est sans doute pourquoi il y a une différence de 1dB de moyenne pour le plan B. Le motif pseudo-aléatoire préserve quant à lui les deux directions horizontales et verticales. Toutes les chrominances ne sont pas modulées à $\frac{1}{2}$ de la fréquence de Nyquist, mais celles qui ne le sont pas possèdent moins d'énergie. Par conséquent le motif pseudo-aléatoire effectue un bon compromis entre le critère sur les fréquences et celui sur les orientations.

Les considérations que nous venons de faire sont renforcées si l'on considère que l'optique du système est circulaire. Si cette dernière est réglée de manière à avoir un maximum de résolution achromatique dans les directions verticales et horizontales sans pour autant, avoir de repliement spectral, alors les fréquences admissibles pour la chrominance se trouvent dans les coins du spectre. La Figure 5.5(a) montre le résultats de la reconstruction de l'image CZP avec une telle optique circulaire. Les fausses couleurs sont substantiellement réduites dans le cas de la mosaïque irrégulière proposée. Ceci est confirmée par les mesures de qualité objective moyennées sur la base d'images Kodak consignées Table 5.3 dans laquelle on constate, par exemple, un gain de plus de 1 dB par rapport à la mosaïque de Bayer en présence de l'optique passe-bas circulaire.

D'autre part, de l'effet de grille, tel que décrit par Alleysson *et al.*, est présent dans les images reconstruites. Il est cependant faible, du fait du bon calibrage des filtres.

3 Nouveau modèle de séparation des canaux spatiaux et chromatiques

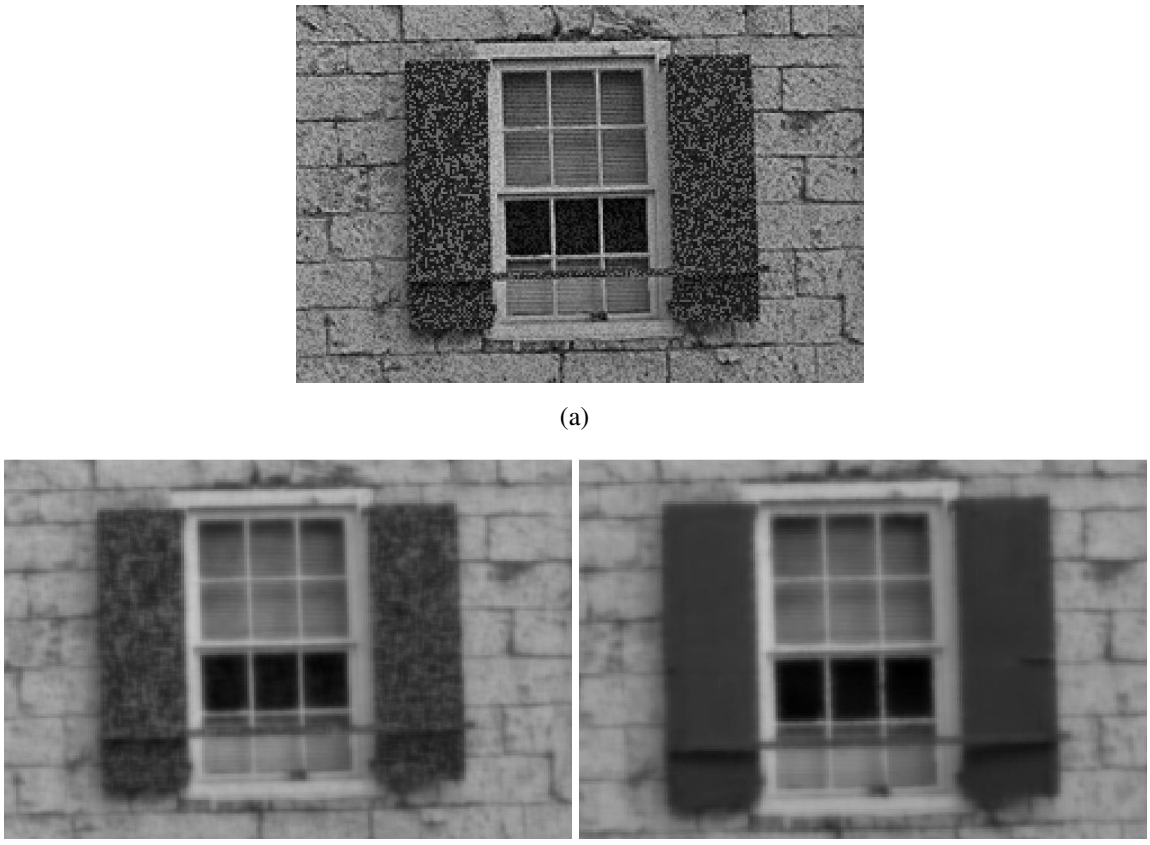
Dans cette section, nous tentons de reconstruire une image couleur à partir d'une mosaïque quelconque par un algorithme similaire à celui par sélection de fréquence (cf. p.60), c'est-à-dire par une estimation de l'intensité puis par démodulation des porteuses de chrominance. Contrairement à la méthode LMMSE que nous venons de décrire, dans laquelle les filtres sont construits par apprentissage sur une base d'images, nous essayons ici de les modéliser, c'est-à-dire que nous essayons de les déduire d'un modèle théorique.

3.1 Approche par estimation de l'intensité

3.1.1 Une première approche, par convolution normalisée

Le modèle d'échantillonnage que nous avons rappelé au début de ce chapitre nous montre que, dans une image en mosaïque, l'intensité est en bande de base, contrairement à la chrominance qui se retrouve modulée à des fréquences non nulles. Il est donc naturel de chercher à estimer l'intensité par filtrage passe-bas de l'image en mosaïque I_m :

$$\Phi(x, y) = (f * I_m)(x, y)$$



(a)

(b) (c)

FIG. *5.6: (a) Image en mosaïque représentée en niveaux de gris. (b)Résultat d'un filtrage passe-bas spatialement invariant. L'image est bruitée, on récupère un bruit de chrominance. (c) Résultat d'un filtrage passe-bas par convolution normalisée sur chacun des plans couleur. Le résultat est dépourvu du bruit associé à la grille d'échantillonnage.*

où $*$ représente le produit de convolution et f un filtre passe-bas spatialement invariant. Cependant, dans le cas d'un échantillonnage irrégulier, la mosaïque reste perceptible après un tel filtrage. Le résultat est une intensité bruitée, telle qu'on peut le voir sur la Figure 5.6(b). Ce problème rejoint en fait celui de l'interpolation d'un signal échantillonné irrégulièrement et dont une des solutions étaient la convolution normalisée (cf. p.93), qui permettait de prendre en compte la densité locale des échantillons.

Nous avons vu (p.60) que l'intensité dans une mosaïque dont les proportions globales en photorécepteurs seraient p_R, p_V et p_B était définie par :

$$\Phi(x, y) = \sum_i p_i C_i(x, y)$$

Cependant, si l'on mesure localement ces densités, par exemple en faisant glisser une fenêtre et en comptant le nombre d'échantillons présents dans la fenêtre - autrement dit en faisant une convolution -, alors leurs valeurs peuvent varier avec la position de la fenêtre dans l'image. Ces variations seront d'autant plus grande que la mosaïque sera irrégulière. Ainsi les trois composantes chromatiques n'apporteront pas la même contribution au filtrage passe-bas selon la position dans l'image. Il s'agit donc de régulariser ces contributions afin qu'elles soient les mêmes

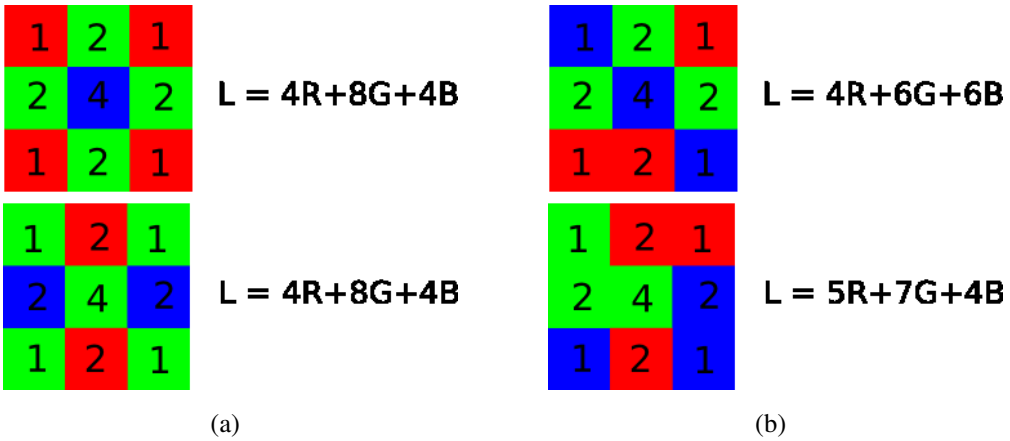


FIG. *5.7: Illustration de la variation de la définition de l'intensité dans le cas irrégulier si le filtre passe-bas est uniforme. Les coefficients du filtre sont représentés pour chaque voisinage. (a) Deux positions dans la mosaïque de Bayer. Les contributions des canaux à l'intensité sont constantes quelque soit la position. (b) Deux positions dans une mosaïque irrégulière. Les contributions à l'intensité varient avec la position considérée.*

quelque soit la position spatiale, exactement comme c'est le cas lorsque l'on effectue une interpolation par convolution normalisée (paragraphe 3.4 du chapitre 4). La Figure 5.7 illustre la composition du signal d'intensité estimé par un filtre passe-bas 3×3 à deux positions différentes dans une mosaïque, avec dans un cas la mosaïque régulière de Bayer, et dans un autre une mosaïque aléatoire.

L'image en mosaïque peut s'écrire comme la somme de ses trois plans couleur démultiplexés (c'est-à-dire les plans couleur dont les informations chromatiques manquantes sont mises à zéro) :

$$I_m = I_m.m_R + I_m.m_V + I_m.m_B$$

Par conséquent, un filtre f s'appliquant sur l'image en mosaïque I_m peut être considéré, par linéarité de la convolution, comme s'appliquant sur chacun des plans couleur démultiplexés :

$$\begin{aligned} \Phi &= f * I_m \\ &= f * (I_m.m_R + I_m.m_V + I_m.m_B) \\ &= f * (I_m.m_R) + f * (I_m.m_V) + f * (I_m.m_B) \end{aligned}$$

On peut alors normaliser la convolution sur chacun des plans couleur et multiplier par les densités respectives :

$$\Phi(x, y) = \sum_{i \in \{R,V,B\}} p_i \frac{(f * (I_m.m_i))(x, y)}{(f * m_i)(x, y)}$$

Le résultat de cette convolution normalisée donne effectivement une intensité dépourvue du "bruit" de chrominance qui était présent de la cas du filtrage uniforme (Figure 5.6(c)). Une manière de comprendre ce qui se passe est de considérer que l'intensité est un axe dans l'espace RVB, défini par les coefficients p_R, p_V et p_B. Si ces coefficients varient d'un point à l'autre de l'espace, alors la droite d'intensité pointe dans différentes directions d'un point à l'autre. La convolution normalisée permet de définir une unique droite d'intensité dans l'espace RVB pour toute l'image.

Par ailleurs, dans le cas de l'utilisation d'un filtre à Réponse Impulsionnelle Finie (RIF), on peut ne faire qu'une seule convolution (spatialement variante) au lieu d'une convolution par plan couleur. Une convolution spatialement variante ne pouvant être décrite globalement, plaçons nous au pixel x et considérons que le noyau du filtre est analogique (d'où l'emploi de l'intégrale dans le produit de convolution). Le calcul de la sortie du filtrage s'écrit :

$$\Phi(x) = \sum_{i \in \{R,V,B\}} p_i \frac{\int I_m(u).m_i(u) f(u-x) du}{(f * m_i)(x)}$$

Faisons passer les dénominateurs dans les intégrales et factorisons par l'image en mosaïque I_m :

$$\Phi(x) = \int I_m(u) \underbrace{\sum_{i \in \{R,V,B\}} p_i \frac{m_i(u) f(u-x)}{(f * m_i)(x)}}_{f_{lum}(u-x)} du$$

Ainsi notre filtre d'intensité centré sur la position x s'écrit sous la forme :

$$f_{lum}(u-x) = p_R f^R(u-x) + p_V f^V(u-x) + p_B f^B(u-x)$$

où les f^i sont les restrictions normalisées du filtre f aux grilles m_i. C'est le fait que les grilles m_i soient à supports spatiaux disjoints qui permet d'effectuer une seule convolution, directement sur l'image en mosaïque (au lieu d'une convolution sur chaque plan couleur). De plus le filtre f_{lum} est effectivement spatialement variant, car chaque restriction f^i suit la forme de la grille m_i.

Cependant, dans le cas d'une implémentation récursive (cas d'un filtre à réponse impulsionnelle infinie), on est obligé d'exécuter trois convolutions distinctes pour la normalisation. Cela vient du fait que si l'on effectue un filtrage récursif directement sur l'image en mosaïque, les informations provenant des différents pixels du voisinage se retrouvent mélangées. Lors du calcul de la sortie du filtre, nous n'avons pas accès aux différents canaux couleurs, uniquement à la sortie aux pixels voisins. On ne peut donc pas normaliser la contribution de chacun des canaux.

Les limites de la méthode

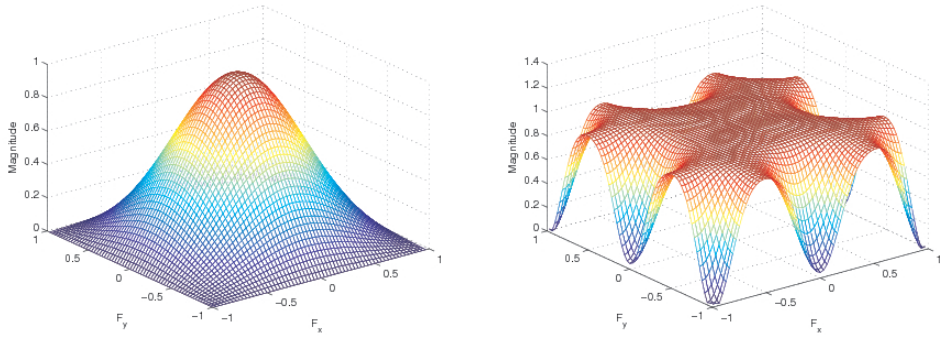
Plaçons nous dans le cas particulier où les trois photorécepteurs seraient localement en même proportions que les densités globales (c'est le cas de la mosaïque de Bayer) :

$$(f * m_i)(x) = p_i$$

Cela signifie que les restrictions du filtre f aux différentes sous-mosaïques sont normalisées par 1, et par conséquent le filtre d'intensité n'est autre que le filtre f :

$$f_{lum}(u-x) = \sum_i p_i f^i(u-x) = f(u-x)$$

Ainsi, dans ce cas particulier, le filtre d'intensité est un simple filtre passe-bas *isotrope* (Figure 5.8(a)). Il ne comporte donc aucune information sur la localisation de la chrominance dans le plan de Fourier, si ce n'est qu'elle n'est pas en bande de base. Dans le cas particulier de la



(a) (b)

FIG. *5.8: Filtres d'intensité pour la mosaïque de Bayer : (a) d'après la présente méthode ; (b) d'après l'algorithme par sélection de fréquences (2.2.3). On voit la sous-optimalité du filtre proposé en (a).*

mosaïque de Bayer, on est alors bien loin du filtre proposé par Alleysson *et al.* qui atténue uniquement les zones du spectre comportant des porteuses de chrominance (Figure 5.8(b)). Cette méthode est donc sous-optimale.

Le fait que filtre d'intensité soit trop passe-bas entraîne un lissage excessif de l'image. De plus, le fait que la chrominance soit prise comme le complémentaire d'un signal d'intensité mal estimé (*i.e.* $f_{chr} = \delta - f_{lum}$), entraîne une erreur d'estimation de la chrominance. Cela se traduit par des fausses couleurs dans les zones de hautes fréquences d'intensité (Figure 5.9), mais qui ont néanmoins le mérite d'hériter des bonnes propriétés de l'échantillonnage irrégulier, à savoir d'être des erreurs de reconstructions non cohérentes, au contraire des erreurs extrêmement structurées dues au repliement spectral dans le cas de l'échantillonnage régulier.

3.1.2 Estimation d'une intensité de meilleure acuité spatiale

Le filtre f_i, composant le filtre d'intensité proposé dans le paragraphe précédent, est la restriction du filtre passe-bas isotrope f à la grille m_i :

$$f^i(u-x) = m_i(u) f(u-x)$$

Or cette grille m_i peut se décomposer, nous avons utilisé cette propriété à plusieurs reprises, en une partie constante p_i et une partie variable $\tilde{m}_i$. La partie variable module le signal en hautes fréquences, c'est-à-dire là où est modulée la chrominance, tandis que la partie constante laisse le filtre en bande de base, là où se trouve l'intensité :

$$f^i(u-x) = p_i f(u-x) + \underbrace{\tilde{m}_i(u) f(u-x)}_{modulation}$$

Par conséquent le filtre f^i, par l'intermédiaire de la position des échantillons de type i dans le voisinage, nous renseigne sur les fréquences auxquelles est modulée la chrominance par la sous-mosaïque m_i. La position dans le domaine fréquentiel de l'intensité est quant à elle indépendante du voisinage. Ainsi, le voisinage local sur chaque plan i va nous permettre d'estimer

FIG. *5.9: Démosaïçage par estimation de l'intensité par convolution normalisée. La mauvaise estimation des hautes fréquences d'intensité se traduit par des contours irréguliers et par des fausses couleurs.*

la chrominance. Un candidat pour être un filtre linéaire estimateur de chrominance serait donc de la forme :

$$f_{chr} = \alpha_R f^R + \alpha_V f^V + \alpha_B f^B$$

où les α_i sont des coefficients à déterminer. l'intensité étant le signal complémentaire de la chrominance dans l'image en mosaïque, le filtre d'intensité associé serait de la forme :

$$f_{lum} = \delta - (\alpha_R f^R + \alpha_V f^V + \alpha_B f^B)$$

où δ est l'impulsion de Dirac. Nous pouvons appliquer la condition sur la contribution à l'intensité de chaque canal à hauteur de p_i. Ainsi la somme des coefficients associés à un certain type de filtre i doit être égal à p_i :

$$\forall i \in \{R, V, B\} \qquad \int f_{lum}(u - x).m_i(u)du = p_i$$

Or ces sommes de coefficients vont dépendre de la couleur du pixel sur lequel est centré le filtre. Supposons que l'on soit sur un pixel rouge (le raisonnement est identique pour les pixels vert et bleu), alors la restriction du filtre d'intensité à la grille rouge a la particularité de

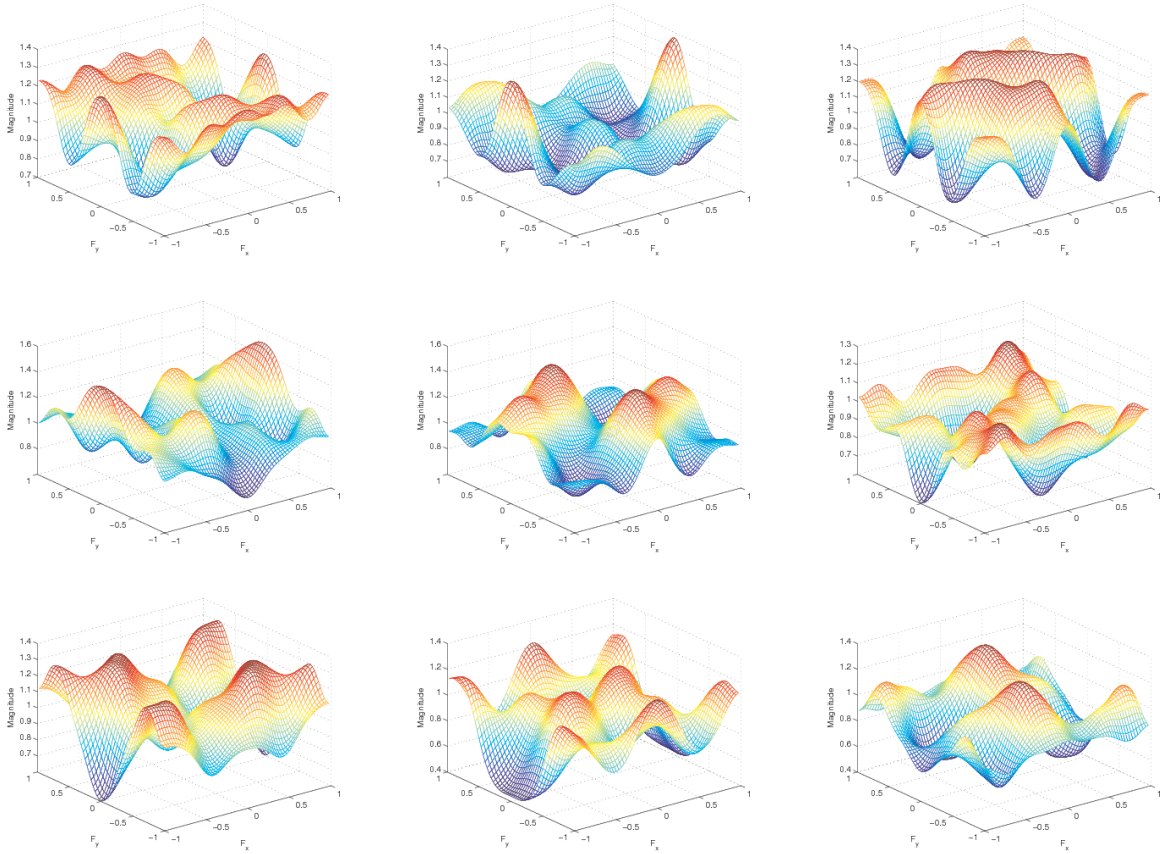


FIG. *5.10: Réponses fréquentielles du filtre d'intensité modélisé, pour différentes positions spatiales dans une mosaïque irrégulière. Comme on peut le voir, le filtre est spatialement variant : les fréquences atténuées par le filtre dépendent de la position spatiale dans la mosaïque.*

contenir une impulsion de dirac. Ainsi, comme les f_i sont normalisés, nous avons les contraintes suivantes sur les α_i :

$$\begin{cases} 1 - \alpha_R = p_R \\ -\alpha_V = p_V \\ -\alpha_B = p_B \end{cases}$$

soit $\alpha_R = 1 - p_R$, $\alpha_V = -p_V$ et $\alpha_B = -p_R$. Et donc notre filtre, à une position rouge s'écrit :

$$f_{lum,R} = \delta - ((1 - p_R)f^R - p_V f^V - p_B f^B) \tag{5.16}$$

On a donc construit un filtre estimateur d'intensité dans le cas général. Si on observe sa réponse fréquentielle en différents points de l'image (Figure 5.10), on constate qu'il est spatialement variant. Ces variations avec la position spatiale ont deux causes. La première vient du caractère intrinsèquement irrégulier de la mosaïque. En effet, si on considère une seule sous-mosaïque m_i, alors les porteuses se déplacent de manière continue avec le déplacement dans l'image. La deuxième cause est d'une autre nature et peut facilement être expliquée si l'on examine le cas particulier de la mosaïque de Bayer, ce que nous faisons au prochain paragraphe.

La Figure 5.11 montre, pour différentes positions spatiales dans le motif pseudo-aléatoire présenté p.102, les filtres modélisés construits à partir d'un filtre f gaussien ainsi que les filtres

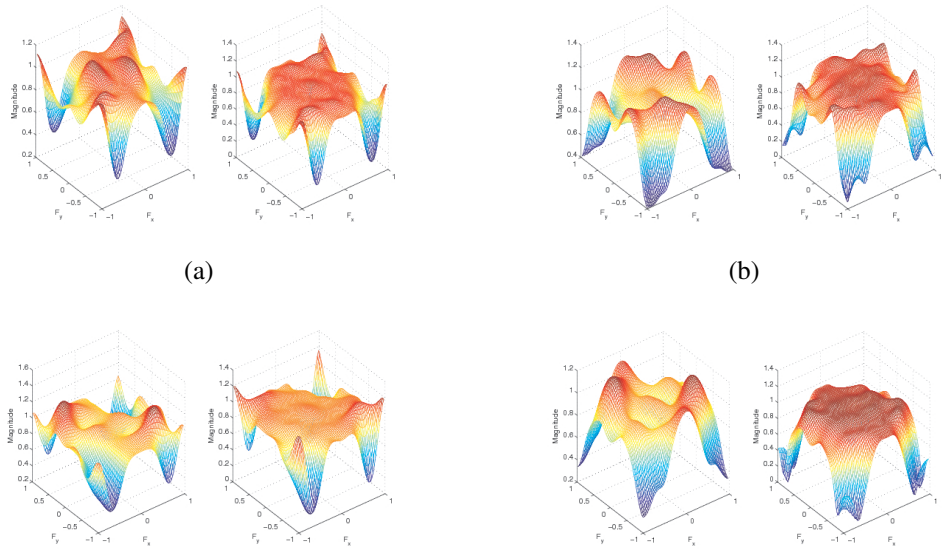
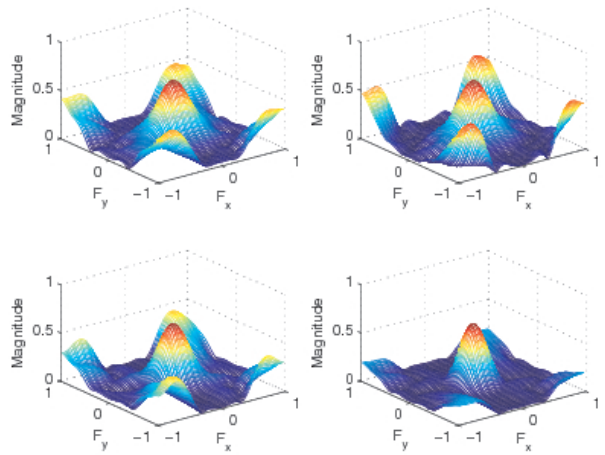


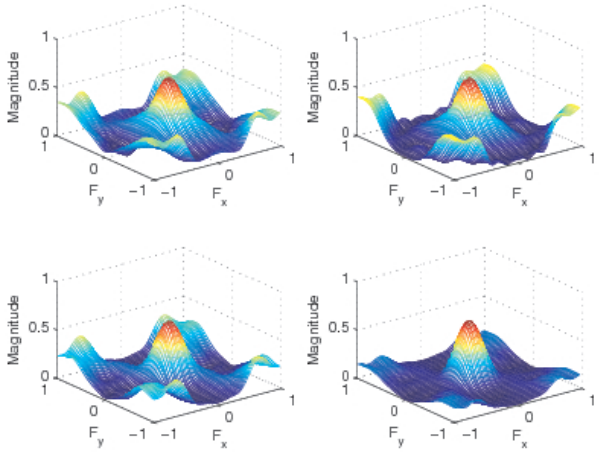
FIG. *5.11: Les figures (a) à (d) montrent, pour 4 configurations de voisinages, la réponse fréquentielle du filtre construit à partir d'un filtre f gaussien (sous-figures gauches) et celle du filtre LMMSE obtenu par apprentissage (sous-figure droites) . Les filtres modélisés et les filtres LMMSE ont des gabarits similaires. On note toutefois quelques différences, les filtres modélisés semblent présenter davantage d'oscillations.*

LMMSE obtenus par apprentissage. On constate que les deux types de filtres ont des gabarits proches, ils atténuent sensiblement aux mêmes fréquences. Cependant ces filtres ne sont pas identiques, les filtres théoriques semblent présenter plus d'oscillations. Afin de mettre en évidence ces différences, la Figure 5.12 montre une autre représentation de ces filtres, pour deux positions spatiales. Sur cette figure, les filtres sont décomposés en deux parties. Supposons que le pixel courant soit R. Notre filtre théorique est donc défini par l'Equation (5.16). A gauche, est représentée la composante positive $(1 - p_R)f^R$, et à droite la composante négative $-p_V f^V - p_B f^B$. On constate que la composante positive du filtre théorique est très proche de celle du filtre LMMSE. Bien que l'apprentissage ait pu optimiser plus finement la forme des filtres, les mêmes fréquences sont laissées passées dans les deux cas, et avec les mêmes facteurs d'amplification. Au contraire, la composante négative du filtre présente de grandes différences au niveau de l'amplitude de certaines fréquences. L'amplitude du filtre en basse fréquence est correcte, tandis qu'en hautes fréquences, le filtre n'amplifie pas suffisamment par rapport au filtre LMMSE. Cela se traduit sur le filtre d'intensité par le fait que les porteuses de chrominance sont moins atténuées. Nous ne sommes pas en mesure d'expliquer cette différence. Il peut s'agir soit d'un problème de modélisation de notre part, soit simplement d'une amélioration faite par les filtres LMMSE en tirant partie des statistiques des images.

Au niveau de la qualité des images reconstruites, les filtres théoriques d'intensité donnent un résultat en deçà de 3dB des performances des filtres LMMSE sur la mosaïque pseudo-aléatoire (Table 5.4), avec davantage d'effet de grille et de fausses couleurs. La Figure 5.13 montre un exemple d'image pour lequel des fausses couleurs apparaissent avec les filtres modélisés, qui n'apparaissent pas avec les filtres LMMSE.



(a)



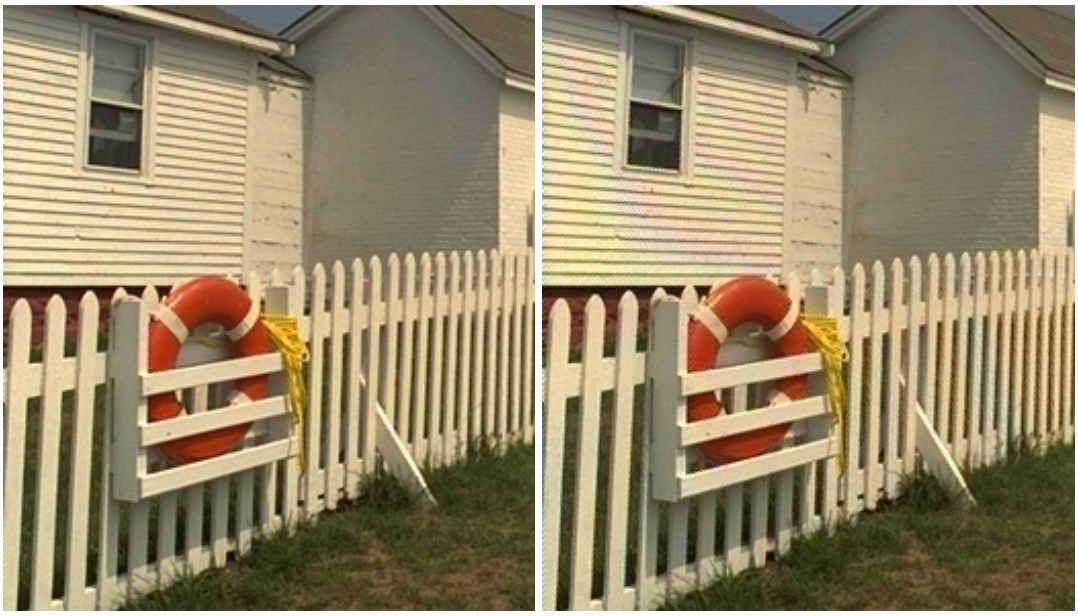
(b)

FIG. *5.12: Comparaison des filtres construits et des filtres LMMSE décomposés en deux partie, pour deux positions dans la mosaïque pseudo-aléatoire ((a) et (b)). Pour chaque sous-figure, la ligne du haut correspond à la méthode LMMSE, et la ligne du bas correspond à la méthode théorique. Pour chaque méthode, le filtre de gauche correspond à la partie positive du pourtour du filtre, et le filtre de droite correspond à la partie négative (cf. explications dans le texte). On constate que l'amplitude de la partie négative du filtre ne remonte pas autant dans les hautes fréquences que ne le fait la composante négative du filtre LMMSE.*

3.1.3 Retour au cas particulier de la mosaïque de Bayer

Afin d'avoir une idée de la validité de notre nouveau filtre, observons sa forme dans le cas particulier de la mosaïque de Bayer. Les pixels rouge et bleu ayant des rôles symétriques, considérons un pixel courant rouge ou bleu. Le filtre d'intensité s'écrit alors :

$$f_{lum,RB} = \delta - (0.75f^R - 0.5f^V - 0.25f^B) \tag{5.17}$$



(a) (b)

FIG. *5.13: Reconstruction d'une partie de l'image lighthouse à partir des filtres LMMSE (a) et des filtres théoriques (b). Pour certains motifs des fausses couleurs apparaissent avec les filtres théoriques qui n'apparaissent pas avec les filtres LMMSE.*

Type de filtre	R	G	B
Modélisé	36.56 ($\pm$2.38)	37.81 ($\pm$2.31)	36.87 ($\pm$2.46)
LMMSE	38.74 ($\pm$2.54)	39.78 ($\pm$2.43)	39.07 ($\pm$2.64)
Modélisé	35.84 ($\pm$2.36)	37.27 ($\pm$2.35)	36.34 ($\pm$2.46)
LMMSE	38.90 ($\pm$2.50)	40.12 ($\pm$2.41)	39.44 ($\pm$2.67)

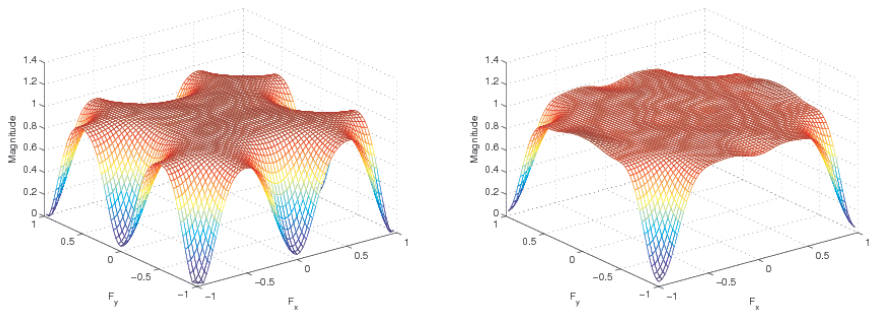
TAB. *5.4: PSNR moyen obtenu sur deux mosaïques pseudo-aléatoires en utilisant les filtres modélisés (à partir de gaussiennes d'écarts-type $\sigma_R = \sigma_G = 2.9$ et $\sigma_B = 2.6$ et en utilisant les filtres LMMSE. Les filtres LMMSE donnent des résultats nettement supérieurs sur les deux mosaïques.*

Pour un filtre de "base" f gaussien, on observe bien une réponse fréquentielle (Figure 5.14(a)) en accord avec le filtre de l'algorithme dit par sélection fréquentiel et avec le filtre obtenu par méthode LMMSE pour les pixels rouge et bleu. Le filtre correspondant au pixel vert est :

$$f_{lum,V} = \delta - (-0.25f^R + 0.5f^V - 0.25f^B) \tag{5.18}$$

Il correspond également au filtre obtenu par la méthode LMMSE (Figure 5.14(b)), corroborant le fait que le filtre optimal pour la mosaïque de Bayer doivent être spatialement variant, bien que cette dernière soit une grille régulière.

Pourquoi le filtre d'intensité, aux positions correspondant aux filtres verts, ne doit-il pas filtrer les porteuses horizontales et verticales associées aux grilles R et B ? La réponse n'a *a priori* rien d'évident, puisque le concept de fréquence est difficilement associable à des positions spatiales aussi précises que l'échelle d'un pixel (le filtre d'intensité varie considérablement d'un pixel à l'autre). Lian *et al.* [Lian 05] avaient émis eux aussi l'hypothèse que le filtre d'intensité pouvait ne pas couper les fréquences horizontales et verticales sur les positions vertes. Ils jus-



(a) (b)

FIG. *5.14: Modules des réponses fréquentielles du filtre d'intensité pour la mosaïque de Bayer, pour les pixels rouge et bleu (à gauche) et le pixel vert (à droite). On retrouve les mêmes filtres que par apprentissage lors de la méthode LMMSE.*

tifiaient ce choix par le fait que, dans la décomposition des trois sous-mosaïques de Bayer en somme de cosinus (cf. p.61), les cosinus des grilles R et B associées aux fréquences horizontales et verticales s'annulaient sur les positions V. Il me semble bien hasardeux de conclure à l'annulation de fréquences par des considérations d'annulation de sinusoïdes en certains points du domaine spatial. En revanche, la manière dont est construit notre filtre d'intensité nous permet de conclure sans ambiguité. En effet, notre filtre a la particularité de mettre en opposition la grille verte d'une part, et les grilles rouge et bleue d'autre part (Equation (5.18)). De plus, les filtres f^R et f^B (qui donc s'additionnent si le pixel central est V) n'ont pas besoin d'être normalisés, car dans la mosaïque de Bayer, les densités locales sont les mêmes que les densités globales. Ainsi, le filtre $0.25f^R + 0.25f^B$ ne peut pas faire apparaître dans sa réponse fréquentielle des modulations associées aux grilles R et B car celles-ci sont transparentes. En revanche, sa réponse fréquentielle portera l'information sur la modulation de la grille $m_R + m_B$, c'est-à-dire la grille $1 - m_V$. Le filtre d'intensité peut se réécrire finalement :

$$f_{lum,V} = \delta - 0.5(f^V - f^{\hat{V}})$$

ce qui montre clairement que sur un pixel V, les porteuses de chrominance ne sont modulées que par la grille m_V. Il est important de comprendre que ceci vienne du fait que f^R et f^B sont normalisés par le même coefficient. Si deux canaux s'additionnent dans la construction du filtre et que l'un doit être amplifié par rapport à l'autre, alors la réponse impulsionnelle du filtre contiendra des hautes fréquences (Figure 5.15).

Séparabilité du filtre

Une question importante par rapport à l'efficacité de l'exécution d'un algorithme en traitement d'image, est la question de la séparabilité d'un filtre. Le filtre d'intensité que nous avons construit est-il séparable ? Autrement dit, allons-nous pouvoir l'implémenter par filtrage successif des lignes puis des colonnes ?

Intéressons-nous aux conditions nécessaires pour qu'un filtre soit séparable. Soit h un noyau de taille $n \times n$ ($n \in \mathbb{N}$) séparable en lignes et en colonnes. Supposons que les noyaux séparés

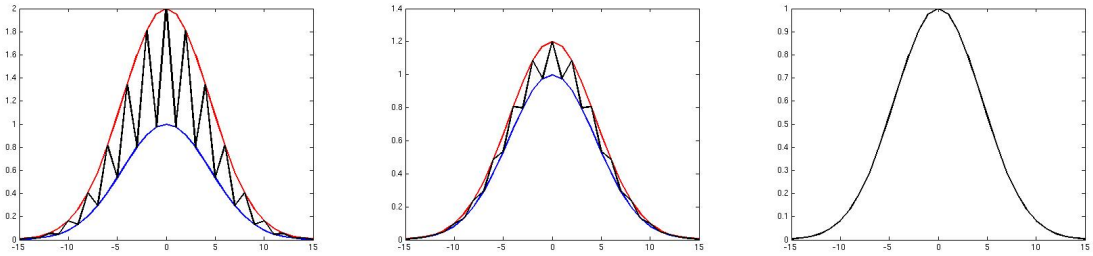


FIG. *5.15: Illustration du fait que lorsque deux canaux ont le même facteur de normalisation dans la construction du filtre d'intensité, leur multiplexage spatial ne fait alors pas intervenir de hautes fréquences (cas de f^R et f^B sur les pixels verts dans la mosaïque de Bayer). (a) les deux réponses impulsionnelles n'ont pas le même facteur de normalisation. Leur multiplexage introduit des hautes fréquences ; (b) L'amplitude de ces hautes fréquences diminue si les normalisations sont proches. (C) les deux réponses impulsionnelles ont le même facteur de normalisation. Le signal résultant est lisse.*

h_l et h_c soient transposés l'un de l'autre et qu'ils soient symétriques :

$$h_l = [h_1 \cdots h_{max} \cdots h_1]$$
$$h_c = [h_1 \cdots h_{max} \cdots h_1]^T$$

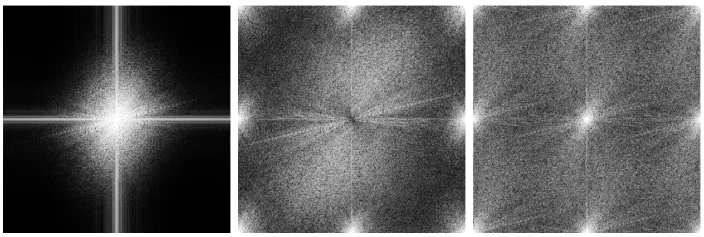
Ces conditions de symétrie sont fortes, elles ne sont vraies que lorsque la mosaïque possède elle-même des symétries, c'est-à-dire lorsque la mosaïques est régulière. Cependant, on peut intuiter qu'une mosaïque régulière a plus de chance d'être associée à un filtre d'intensité séparable qu'une mosaïque irrégulière. Ainsi, si on montre que pour une mosaïque régulière le filtre n'est pas séparable, la probabilité est grande que ce ne soit pas le cas non plus pour une mosaïque irrégulière. Le noyau h est le résultat du produit matricielle entre h_l et h_c :

$$h = h_l h_c = \begin{pmatrix} h_1^2 & & h_1 h_{max} & & h_1^2 \\ & \ddots & & & \\ h_{max} h_1 & & h_{max}^2 & & h_{max} h_1 \\ & & & \ddots & \\ h_1^2 & & h_1 h_{max} & & h_1^2 \end{pmatrix}$$

On remarque que tous les termes diagonaux (des deux diagonales) du filtre de convolution h doivent être positifs. Or dans notre cas, les coefficients peuvent être négatifs (pondérés par $-(1-p_i)$) d'après la construction du filtre d'intensité. Ainsi, le filtre d'intensité n'est pas séparable dans le cas général. Pour que ce soit le cas, il faudrait le démultiplexer sur chaque plan couleur. On rejoint alors la solution proposée dans le paragraphe suivant.

3.2 Approche par estimation de la chrominance

Nous venons de décrire une manière de construire des filtres d'estimation de l'intensité dans une mosaïque quelconque. Intéressons nous maintenant à l'approche duale, qui consiste à estimer le signal de chrominance plutôt que le signal d'intensité. Comme nous le verrons, cette approche, qui ne consiste pas à prendre le filtre complémentaire au filtre d'intensité, possède des propriétés intéressantes quant à l'implémentation des filtres et quant aux conséquences pour un modèle de vision des couleurs chez l'humain.



(a) (b) (c)

FIG. *5.16: Spectres d'amplitude dans le cas de la mosaïque de Bayer. (a) Basses fréquences d'intensité ϕ^{LF}. (b) Hautes fréquences de l'image en mosaïque I_m^{HF}. (c) Canal rouge $\{I_m^{HF}\}_R$ de (b).*

3.2.1 Description de l'approche

Soit f un filtre passe-bas dont les caractéristiques seront précisées plus tard. Soit I_m^{BF} (pour **B**asses **F**réquences) l'image en mosaïque filtrée par f :

$$I_m^{BF}(x,y) = (f * I_m)(x,y)$$

avec '*' le symbole du produit de convolution. En reprenant les notations de l'Equation (5.3) nous avons :

$$(f * I_m)(x,y) = (f * \phi)(x,y) + \sum_{n \neq 0} (f * \psi_n)(x,y)$$

Mais comme les chrominances ψ_n sont modulées en hautes fréquences, on en déduit :

$$f * \sum_{n \neq 0} \psi_n \approx 0$$

et par conséquent :

$$I_m^{BF} \approx f * \phi$$

Ces deux dernières équations sont vérifiées à condition que la fréquence de coupure f_c du filtre f soit inférieure à la fréquence $(n - f_{max})$, f_{max} étant la fréquence maximale des signaux de chrominance. I_m^{BF} contient alors une version grossière (c'est-à-dire floue) - mais néanmoins dépourvue de repliement spectral - de l'intensité ϕ (Fig. 5.16(a)). Nous appellerons désormais cette image ϕ^{BF}.

Soit I_m^{HF} le signal complémentaire de ϕ^{BF} par rapport à l'image en mosaïque (Fig. 5.16(b)) :

$$I_m^{HF} = I_m - \phi^{BF} \tag{5.19}$$

D'après la représentation fréquentielle d'une image en mosaïque (Equation 5.3), I_m^{HF} contient les oppositions de couleurs modulées ainsi que les détails d'intensité (les hautes fréquences d'intensité ϕ^{HF}). Dans le domaine fréquentiel nous avons :

$$\hat{I}_m^{HF}(\nu) = \hat{\phi}^{HF}(\nu) + \sum_{n \neq 0} \hat{\psi}_n(n - \nu) \tag{5.20}$$

avec $\phi = \phi^{BF} + \phi^{HF}$.

Examinons maintenant le démultiplexage de I_m^{HF}. Le démultiplexage d'une image en mosaïque consiste en l'isolation de chaque canal couleur de l'image en mosaïque en trois composantes distinctes. Les composantes qui n'ont pas été échantillonnées sont mises à zéro, suivant les fonctions m_i. En pratique, cela signifie copier les pixels d'une même classe chromatique et les copier dans une matrice préalablement initialisée à zéro. Ceci est formellement équivalent à multiplier l'image en mosaïque I_m^{HF} par les sous-mosaïques m_i :

$$\{I_m^{HF}\}_i(x,y) = I_m^{HF}(x,y)m_i(x,y)$$

Du fait de l'équation 5.19, il vient :

$$\{I_m^{HF}\}_i = I_m m_i - \phi^{LF} m_i$$

où en fait $I_m m_i = C_i m_i$, et $C_i = \phi + \{\psi\}_i$. Par conséquent :

$$\begin{aligned}\{I_m^{HF}\}_i &= (\phi + \psi_i)m_i - \phi^{LF} m_i \\ &= (\phi^{HF} + \psi_i)m_i\end{aligned}$$

La sous-mosaïque m_i peut se décomposer en la somme d'un terme constant p_i et d'une partie variable $\tilde{m}_i$, $m_i = p_i + \tilde{m}_i$ (comme dans l'équation 5.2), ce qui mène à :

$$\{I_m^{HF}\}_i = \underbrace{p_i(\phi^{HF} + \psi_i)}_{\text{bande de base}} + \underbrace{(\phi^{HF} + \psi_i)\tilde{m}_i}_{\text{modulé}}$$

Nous avons donc deux termes dans l'expression de $\{I_m^{HF}\}_i$: un terme est en bande de base, il s'agit de la somme de la chrominance i et des détails d'intensité, l'autre module le même signal en haute fréquence (figure 5.16(b)). De surcroît, si la fréquence de coupure f_c du filtre passe-bas f est choisie de manière adéquate, alors le signal de chrominance est à support fréquentiel disjoint des autres signaux. La condition sur la fréquence de coupure est qu'elle soit au moins égale à la largeur de bande de la chrominance. Ainsi, un simple filtrage passe-bas sur chacun des plans de l'image démultiplexée permet de récupérer la chrominance. En pratique, nous utilisons le même filtre f que pour l'estimation de ϕ^{LF} :

$$\{\psi\}_i = \frac{1}{p_i} f * \{I_m^{HF}\}_i \tag{5.21}$$

Connaissant la mosaïque et les composantes de chrominance, nous pouvons en déduire l'intensité par soustraction entre l'image en mosaïque et la chromiance remodulée :

$$\phi = I_m - \sum_i \psi_i m_i$$

Une reconstruction sans erreur est bien sûr illusoire. En pratique, il y a du repliement spectral entre la chrominance et l'intensité modulée, se traduisant par de fausses couleurs ou du bruit le long des contours. Pour réduire ces défauts, une mise en oeuvre adaptative de l'algorithme est présentée au paragraphe 3.2.3.

Pour résumer, l'approche par estimation de la chrominance peut se diviser en cinq étapes :

1. séparation de l'image mosaïque I_m en une composante basse fréquence ($I_m^{LF} = \phi^{LF}$) et une composante haute fréquence (I_m^{HF}) en utilisant le filtre passe-bas f,
2. demultiplexage de la composante haute fréquence,
3. filtrage par f de chacun des plans de l'image démultiplexée afin d'obtenir les composantes de chrominance ψ_R, ψ_V and ψ_B,
4. récupération de la composante d'intensité ϕ par soustraction entre l'image en mosaïque et la chrominance remodulée[1],
5. addition de l'intensité à la chrominance pour obtenir C_R, C_V and C_B.

3.2.2 Choix de la fonction d'interpolation

L'efficacité la méthode linéaire (en terme de qualité et de temps de reconstruction) repose sur le choix de f. Un net avantage vis-à-vis de l'approche par estimation de l'intensité est que les filtres passe-bas n'ont pas de contrainte particulière de forme. Ils n'ont qu'une contrainte au niveau du choix de la fréquence de coupure. Par souci d'efficacité algorithmique, nous choisissons d'implémenter f sous forme d'un filtre récursif séparable de premier ordre dont la transformée en z est [de Lavarène 07a] :

$$F(z_1, z_2) = F_1(z_1)F_2(z_2) \tag{5.22}$$

où z_1 et z_2 sont les variables horizontales et verticales, et avec

$$F_k(z_k) = (1-a)^2 \frac{1}{1 - az_k^{-1}} \frac{1}{1 - az_k} \tag{5.23}$$

Ce filtre est de premier ordre. Il ne fait donc intervenir que les quatre voisins les plus proches, d'où une grande efficacité algorithmique. Sa réponse fréquentielle est représentée figure 5.17. Comme le lecteur peut le constater sur cette figure, F atténue davantage dans les directions diagonales que dans les directions horizontale et verticale. Par conséquent, une numérateur d'ordre 1 est nécessaire pour les mosaïques dont les m_i modulent dans les directions verticale ou horizontale (à l'instar de la mosaïque de Bayer pour les canaux rouge et bleu) afin d'annuler le filtre aux fréquences $(0, \pm 0.5)$ et $(\pm 0.5, 0)$. Cependant ce numérateur est superflu pour les mosaïques ne modulant pas la chrominance aux fréquences évoquées. Par exemple, le filtre suivant possède ces propriétés :

$$\frac{1}{8}\begin{bmatrix} 1 & 0 & 1 \\ 0 & 4 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

Le paramètre a du filtre récursif contrôle la fréquence de coupure du filtre et, par voie de conséquence, le compromis entre fausses couleurs et effet de grille. Pour la base Kodak, $a = 0.5$ est un bon choix empirique.

Le filtrage récursif comme modélisation du filtrage rétinien

[1] la remodulation de la chrominance est purement virtuelle, puisqu'elle ne consiste qu'en la non-prise en compte de certains échantillons, suivant m_i.

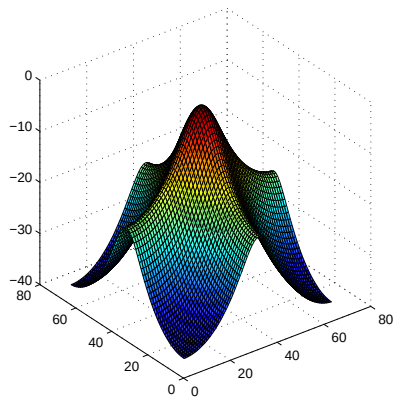


FIG. *5.17: Réponse fréquentielle du filtre à réponse impulsionnelle infinie (dB).*

Ce filtre récursif du premier ordre peut-être vu comme une modélisation électrique du filtrage rétinien [Beaudot 94]. Les jonctions gap liant des cellules entre elles sur une même couche sont modélisés par des résistances R. Dans notre cadre d'étude, nous ne considérons que la dimension spatiale, autrement dit nous ignorons les effets capacitifs. La loi des noeuds appliquée au schéma de la Figure 5.18 nous donne l'équation aux différences suivante :

$$\frac{x(k)-y(k)}{r}+\frac{y(k-1)-y(k)}{R}+\frac{y(k+1)-y(k)}{R}=0$$

En posant $\alpha = \frac{r}{R}$, la transformée en z de cette équation nous donne la fonction de transfert suivante :

$$G(z)=\frac{Y(z)}{X(z)}=\frac{1}{1+2\alpha-\alpha(z^{-1}+z)}$$

Cette fonction peut s'écrire sous la forme d'une composante causale et d'une composante anticausale :

$$G(z)=(1-a)^2\frac{1}{1-az^{-1}}\frac{1}{1-az} \qquad \text{avec} \qquad a=1+\frac{1}{2\alpha}-\sqrt{(1+\frac{1}{2\alpha})^2-1}$$

Cependant, un tel filtrage n'est pas symétrique, car on applique le filtrage anticausal sur le résultat du filtrage causal, créant ainsi un déphasage. On lui préfère l'implémentation suivante :

$$G(z)=\frac{1-a}{1+a}\left(\frac{1}{1-az^{-1}}+\frac{1}{1-az}-1\right) \tag{5.24}$$

qui a l'avantage d'être symétrique, puisque les filtrages causaux et anticausaux sont séparés. Cette dernière équation provient de l'égalité suivante :

$$\frac{1}{1-az^{-1}}+\frac{1}{1-az}-1=\frac{1-a^2}{(1-az^{-1})(1-az)}$$

Le comportement du filtre est comparable à un processus de diffusion : le potentiel d'une cellule donnée à un instant donné se propage vers les cellules voisines tout en diminuant peu à peu. Le fait que ce filtre soit implémentable par un réseau de résistance fait que notre algorithme peut être implémenté sur une rétine analogique.

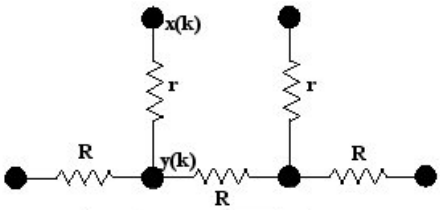
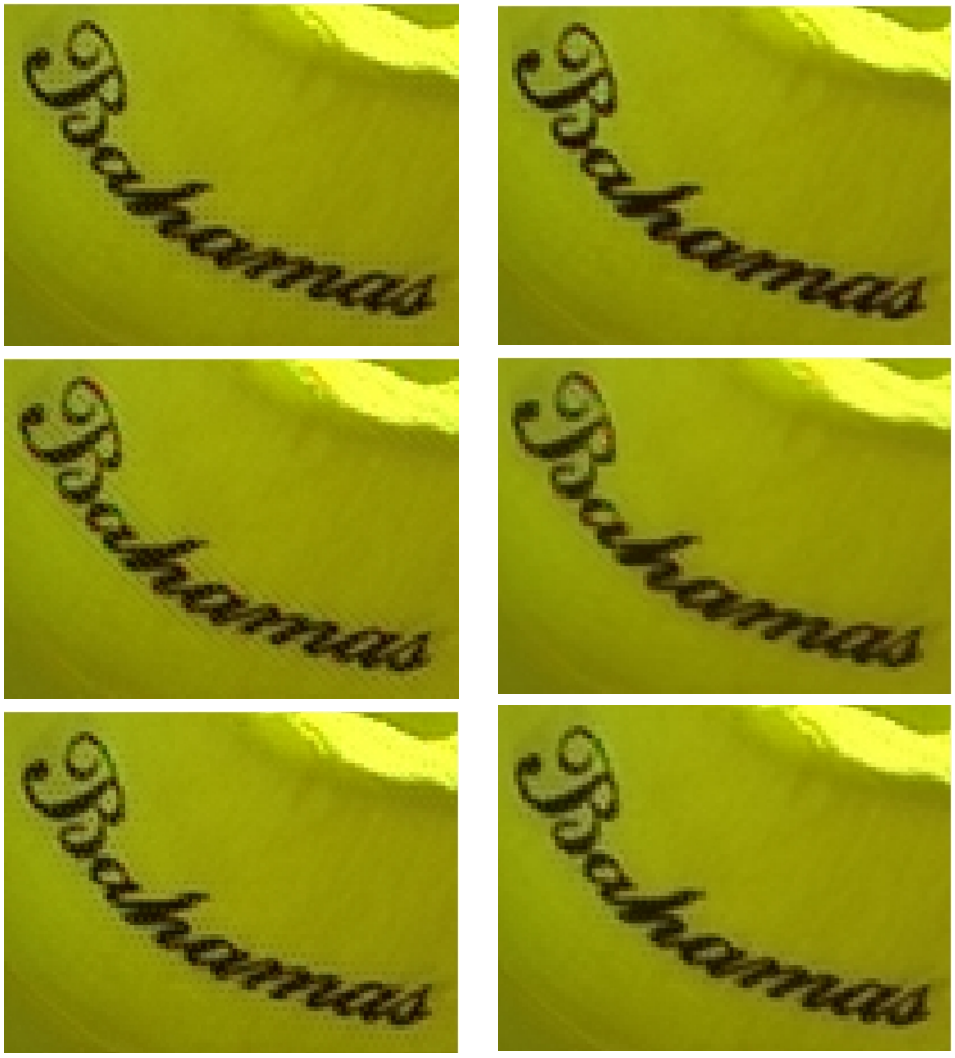


FIG. *5.18: Réseau de résistance modélisant les connexions rétiniennes.*



(a) (b)

FIG. *5.19: Extrait d'une image reconstruite en utilisant la méthode linéaire (a) et la méthode adaptative (b). De haut en bas, les mosaïques utilisées sont les mosaïques de Bayer, à bandes diagonales et de Lukac*

3.2.3 Mise en oeuvre adaptative

Nous proposons deux mises en oeuvre adaptatives possibles. L'une est implémentée, l'autre ne l'est pas encore.

Filtrage adaptatif puis soustraction des résidus d'intensité

l'intensité grossière ϕ^{BF} (mais dénuée, dans une certaine mesure, de repliement spectrale)

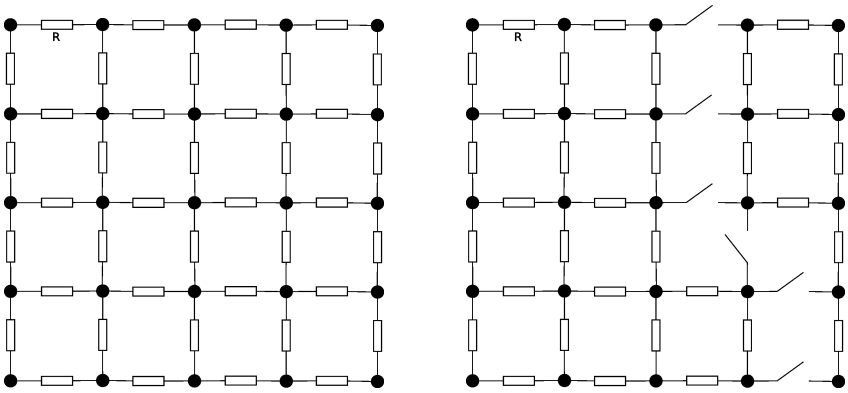


FIG. *5.20: A gauche : réseau de résistance ; à droite : réseau de résistance scindé par un contour. La diffusion électrique ne se propage pas en travers du contour. En conséquence les contours de chrominance sont plus nets.*

peut-être mise à profit d'une manière adaptative. L'utilisation d'une intensité basse-fréquence peut paraître sous-optimale si l'on considère que les algorithmes de démosaïçage s'appliquant sur la mosaïque de bayer utilisent directement les valeurs de l'image en mosaïque I_m plutôt qu'une version basse-fréquence. Cependant, il faut garder à l'esprit que ces méthodes calculent des gradients sur l'image en mosaïque uniquement entre pixels de même couleur. Or ces pixels ne sont pas adjacents sur la mosaïque. L'image en mosaïque contiendra autant d'information de contour que sa version filtrée passe-bas. De plus, ϕ^{BF} a le grand avantage d'être totalement indépendant de la mosaïque, puisque cette image ne contient que de l'information spatiale.

Nous calculons les gradients horizontaux w_1 et verticaux w_2 (w_2 est défini de la même manière que w_1, mais dans la direction verticale) :

$$w_1(x, y) = abs(\phi^{BF}(x, y + 1) - \phi^{BF}(x, y - 1)) \tag{5.25}$$

Ces gradients nous permettent de choisir la direction du filtrage de chrominance à chaque pixel, c'est-à-dire qu'ils nous permettent de choisir entre F_1 et F_2. Il faut toutefois, si l'on veut faire en sorte que l'algorithme fonctionne pour n'importe quel arrangement des filtres couleurs, calculer dynamiquement le gain du filtre. En effet, l'interpolation de la chrominance se faisant par convolution normalisée, la normalisation doit rester cohérente avec le filtrage utilisé : il faut filtrer adaptativement les grilles d'échantillonnage.

Ce filtrage adaptatif permet d'obtenir une chrominance à transitions plus franches le long des contours. Cependant des résidus d'information d'intensité (notés res) peuvent être mêlés à la chrominance, si bien que la chrominance ne vérifie plus l'égalité $\sum_i p_i\psi_i = 0$ à chaque position spatiale :

$$\sum_i p_i\psi_i = res \qquad \text{avec} \qquad |res| > 0$$

Il faut donc retrancher ce résidu d'intensité signal de chrominance.

Comme pour les autres méthodes de demosaïçage adaptatif, une étape de post-traitement est nécessaire afin d'actualiser les valeurs de chrominance à partir de l'intensité estimée. La Figure 5.19 montre sur un exemple les reconstructions linéaire et adaptative sur les mosaïques de Bayer, diagonale et de Lukac. Cet exemple illustre le fait que la méthode s'applique sur différents motifs et montre que l'adaptatif permet de réduire l'effet de grille.

Cette méthode est plus rapide que celle proposée par Lukac, qui est également applicable à toute mosaïque [Lukac 05a]. En effet, Lukac est obligé de faire plusieurs passes pour calculer des gradients dans toutes les directions dans une mosaïque irrégulière, comme nous l'avons décrit p.68. La solution proposée est plus directe, puisqu'elle ne nécessite qu'une seule passe pour le calcul des gradients, quelle que soit la mosaïque.

Si on considère que le filtrage récursif correspond à un processus de diffusion au sein d'un réseau de résistance, alors l'algorithme adaptatif revient à considérer que le réseau est rompu au niveau des contours (Figure 5.20). Ainsi la réponse impulsionnelle du filtre s'adapte localement aux contour, elle ne se propage pas en travers de ceux-ci.

Filtrage linéaire puis rehaussement des contours de chrominance

Cette seconde solution que nous proposons n'a pas été testée en pratique, mais elle nous semble plus élégante car elle permettrait de contourner certains inconvénient de la méthode adaptative que nous venons de décrire. La précédente méthode présente, en effet, deux inconvénients :

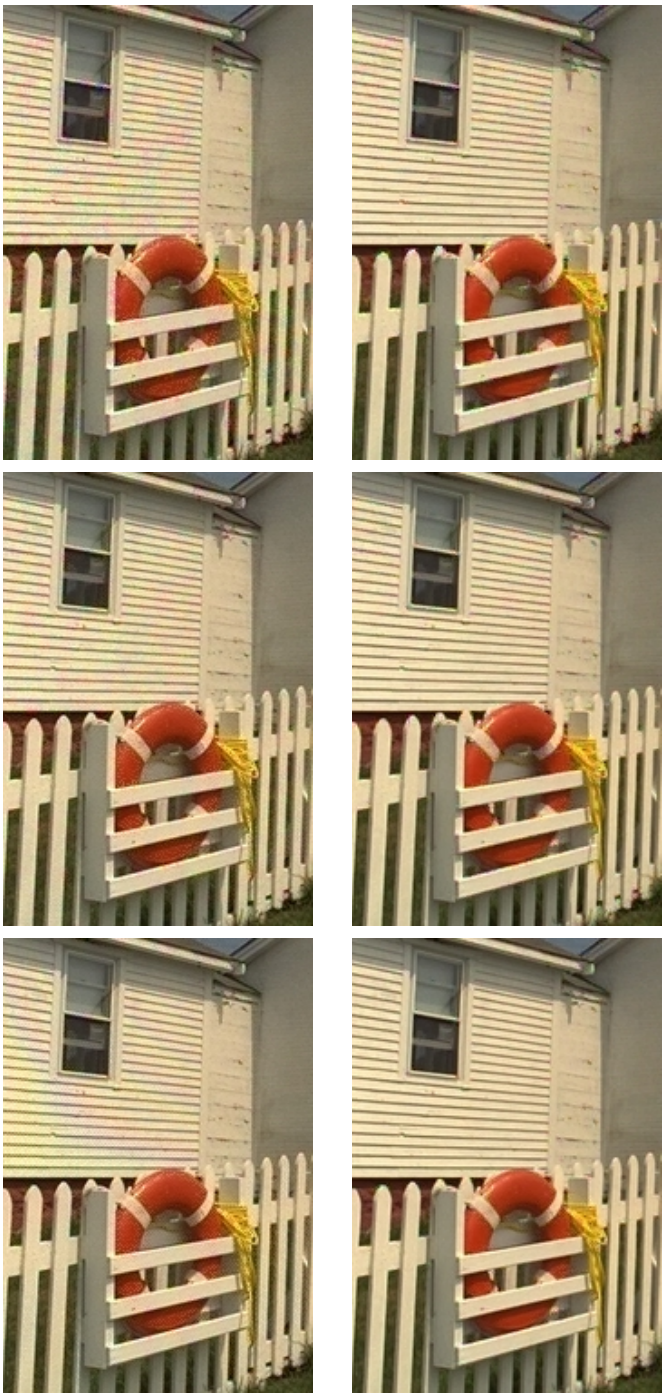
- la nécessité de filtrer adaptativement les grilles d'échantillonnage, afin de faire une convolution normalisée cohérente avec le numérateur. Cela demande des calculs supplémentaires par rapport à un filtrage linéaire.
- il faut retrancher des résidus d'intensité, ce qui, là aussi, rajoute une étape de calcul.

L'effet de grille provient du fait que la chrominance estimée est trop lissée, trop basse-fréquence, en particulier en présence d'un contour. Or nous avons accès à une information sur les contours par le biais de l'intensité basse-fréquence. Ainsi, l'idée serait de faire un réhaussement des contours en fonction de la présence d'un contour d'intensité. Cela permettrait d'avoir une chrominance à variations franches lorsque l'on est en présence d'un contour d'intensité. Afin de ne pas réhausser dans le cas d'une ombre projetée, par exemple, le réhaussement sera plus robuste si on ajoute comme condition la présence d'un contour de chrominance. Ce contour de chrominance serait estimé sur la chrominance obtenu par filtrage linéaire.

Par ailleurs, cette approche semble davantage biologiquement plausible par rapport à la méthode adaptative précédente. Si on considère le canal d'opposition bleu-jaune. Ce canal est calculé au sein même de la rétine et est véhiculé par la voie koniocellulaire. Il semble donc logique de considérer que le calcul de ces oppositions ne soit pas adaptif, puisqu'il est effectué dès les premières synapses du système visuel. De plus, la fonction de sensibilité spectrale des cônes S étant considérablement décorrélée de celles des cônes L et M, l'opposition bleu-jaune devrait être plus large bande que l'opposition rouge-vert. On peut donc penser que la voie d'opposition bleu-jaune subit, à plus forte raison, un traitement cortical afin de retrouver des composantes de hautes fréquences.

3.2.4 Une mosaïque totalement aléatoire ?

La méthode par apprentissage LMMSE présentée au début de ce chapitre permettant de comparer uniquement des mosaïques composée d'un motif de base périodisée. La limite était le temps de calcul pour les inversions matricielles au cours de l'apprentissage. Nous avons maintenant un algorithme fonctionnant quelque soit l'arrangement de la mosaïque, sans restriction.



(a) Linéaire (b) Adaptatif

FIG. *5.21: Résultats des reconstructions linéaire (a) et adaptative (b). Avec, de haut en bas : la mosaïque de Poisson, la mosaïque de Poisson à disque, la mosaïque basée sur un motif* $6x6$.

La Figure 5.21 montre une exemple de reconstruction avec la version linéaire et la version adaptative de l'algorithme sur les trois types de mosaïque que nous avions définies au Chapitre 4, Section 4. Ces mosaïques étaient les suivantes :

- la mosaïque de "Poisson". Les filtres sont obtenus par tirage uniforme,
- la mosaïque de "Poisson à disque", une approximation de la contrainte de distance de

garde entre deux filtres de même couleur,
– un motif périodique 6×6, optimisé manuellement afin que les pixels soient répartie le plus homogènement possible.

La mosaïque de Poisson à disque donne des images d'une meilleure qualité, à la fois subjective (moins de fausses couleurs) et objective (2dB de PSNR d'écart en moyenne sur la base d'image Kodak). Cela vient de la présence de porteuses en basses fréquences dans la cas de la mosaïque de Poisson, ou autrement dit, de la présence de regroupements de pixels de même couleur dans le domaine spatial. Comme nous l'avons vu dans le chapitre sur le système visuel humain (Chapitre 2), des études montrent que l'arrangement des cônes L et M à la surface de la rétine est complètement aléatoire. D'après nos simulations, cet échantillonnage ne serait pas optimal. Ainsi il semblerait que la rétine aurait préféré compenser cet échantillonnage non optimal par d'autres mécanismes (microsaccades, traitements adaptatifs...) plutôt que de développer un mécanisme optimisant l'arrangement des cônes.

La mosaïque de Poisson à disque donne des résultats équivalent au motif 6×6, cependant la périodicité des motifs de Moiré est cassée dans le cas de la matrice de Poisson à disque.

3.2.5 Avantage du modèle pour une chaîne d'acquisition d'images numériques

En passant par l'intensité basse-fréquence et la chrominance, l'algorithme devient indépendant de la mosaïque, c'est-à-dire qu'il fonctionnera si on change la mosaïque (à condition toutefois que la mosaïque soit connue). Ainsi le *pipeline* de traitements en amont et en aval du demosaïçage n'est pas figé pour un type de mosaïque particulière. Il sera plus facilement adaptable si l'on veut changer de mosaïque. Cette dépendance vis-à-vis des mosaïques, et en particulier celle de Bayer, est un frein au passage à un autre arrangement de mosaïque chez les industriels, car cela les oblige à modifier toute leur chaîne de traitement.

Un exemple d'utilisation de cette intensité basse-fréquence a été donné par Meylan *et al.* [Meylan 07] où le modèle d'adaptation de la rétine pour un rendu de grande dynamique est appliqué au niveau de l'image en mosaïque. Les algorithmes HDR sont généralement appliqués après l'étape de demosaïçage, augmentant potentiellement des artefacts de reconstruction. On voit combien le système visuel semble avoir optimisé les traitements non pas en se structurant en module successifs indépendants, mais au contraire en imbriquant chaque traitement afin d'aboutir à une efficacité maximale.

4 Conséquences du modèle pour le système visuel humain

Nous voyons dans cette section la cohérence du modèle que nous venons de présenter avec le système visuel humain et les implications qui en découlent. En particulier, la structure et la construction des champs récepteurs des cellules ganglionnaires nous intéresse. Nous verrons aussi qu'une partie du traitement de la couleur doit se faire à un niveau post-rétinien. Enfin, le modèle que nous proposons implique de connaître au niveau du cortex la classe de cône à laquelle appartient chaque signal. Nous proposons un algorithme simple montrant qu'on peut retrouver de manière non supervisée ces classes.

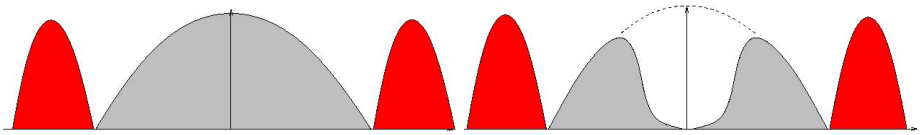


FIG. *5.22: Représentation fréquentielle à 1 dimension du signal au niveau des cônes et à la sortie de la rétine. Le spectre gris représente l'information d'intensité, située en bande de base. Le spectre rouge représente l'information de chrominance modulée en hautes fréquences. A la sortie de la rétine, le spectre est creusé en basses fréquences. Ceci permettra à la chrominance d'être démodulée sans chevaucher l'intensité.*

4.1 Structure des champs récepteurs des cellules ganglionnaires

4.1.1 Cellules ganglionnaires *midget*

Une inhibition latérale pour des supports fréquentiels disjoints

D'après notre modèle par estimation de la chrominance, les basses fréquences d'intensité doivent être retranchées de l'image afin de laisser de la place en bande de base pour la chrominance avant que l'image ne soit transmise au cortex (Figure 5.22). Si l'on appelle Im l'image à la sortie des photorécepteurs, I_{LGN} l'image arrivant au corps genouillé latéral et f un filtre passe-bas, cette opération cela s'écrit :

$$I_{LGN} = I_m - I_m * f$$

De manière équivalente, en factorisant :

$$I_{LGN} = Im * (\delta - f) \tag{5.26}$$

Ainsi, à la sortie de la rétine, le champ récepteur des cellules ganglionnaires *midget* doit être de la forme $\delta - f$. Or, comme f est un filtre passe-bas à gain unitaire, sa réponse impulsionnelle aura une amplitude faible. Ainsi le photorécepteur central présentera une contribution positive prépondérante, tandis que les photorécepteurs voisins présenteront une inhibition moins forte (Figure 5.23). On a donc bien une structure de la forme centre/pourtour, cohérente avec les observations électrophysiologiques. De plus, il est intéressant de remarquer que dans notre modèle le champ récepteur périphérique ne présente pas de sélectivité vis-à-vis des classes de cônes. Cette sélectivité des connexions n'est pas nécessaire au niveau de la rétine, mais elle doit apparaître au niveau du cortex.

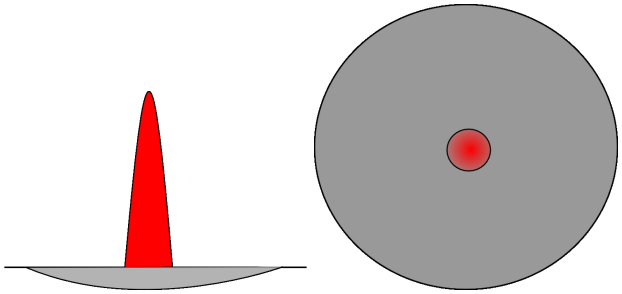


FIG. *5.23: Champ récepteur des cellules ganglionnaires de type P. Le champ associe une pondération élevée positive sur le photorécepteur central, tandis que les photorécepteurs voisins présentent une inhibition diffuse, à caractère passe-bas.*

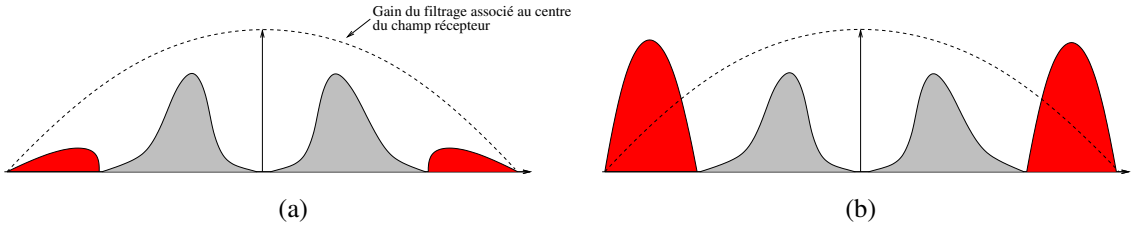


FIG. *5.24: Effet de filtrage passe-bas dû à la taille du centre du champ récepteur des cellules ganglionnaires midget. (a) Cas où le centre est issu de connexions non spécifiques. L'ensemble du spectre est filtré passe-bas, donc les porteuses de chrominance sont atténuées. (b) Cas où le centre n'est connecté qu'à un type de cône. Seules les hautes fréquences d'intensité sont atténuées.*

Pour l'opposition B/Y, il y a des connexions non sélectives à L et M pour former la composante inhibitrice. Les voies parvo et koniocellulaires pourraient donc toutes les deux avoir une contribution négative provenant des signaux L et M de manière non sélective.

Oppositions de couleur en périphérie de la rétine

En périphérie de la rétine, les cellules ganglionnaires sont connectées à un grand nombre de cônes, aussi bien dans le centre de leur champ récepteur que dans le pourtour de leur champ. On peut montrer que si l'on souhaite préserver l'information chromatique, le centre doit nécessairement être spécifique. Le pourtour peut, quant à lui, être non spécifique (à l'image de ce que nous proposons en fovea). En fovea, où les cellules ganglionnaires *midget* ont une ligne directe vers l'information des cônes, on pouvait modéliser (Equation (5.26)) le centre du champ récepteur par une impulsion de Dirac (neutre en terme de convolution). En périphérie, le centre intègre l'information issue de plusieurs cônes. On peut donc le modéliser par un filtrage passe-bas.

Examinons quelles sont les conséquences de ce filtrage sur le spectre du signal véhiculé par les cellules ganglionnaires *midget* selon que les connexions sont spécifiques ou non. Si les connexions sont non spécifiques, alors on effectue un filtrage passe-bas sur le signal en mosaïque. Or, nous avons vu que les hautes fréquences du signal en mosaïque sont composées des porteuses de chrominance et des détails d'intensité. Ce sont donc ces deux informations qui sont atténuées par le filtrage passe-bas associé au centre du champ récepteur *midget* (Figure 5.24(a)). Par conséquent, comme le nombre de cônes présents dans le centre du champ récepteur *midget* augmente avec l'eccentricité, la perception des couleurs devrait diminuer elle aussi avec l'eccentricité. C'est ce qui est mesuré par expériences psychophysiques [Mullen 96].

Supposons maintenant que le centre du champ récepteur *midget* reçoive une information provenant exclusivement d'un type de cône. Alors le filtrage passe-bas s'applique sur les composantes spatialement démultiplexées du signal en mosaïque. Comme nous l'avons vu, le démultiplexage a pour conséquence de démoduler l'information de chrominance en bande de base. Ainsi, le filtrage passe-bas n'atténuera pas l'information d'oppositions de couleur, mais uniquement l'information de détail d'intensité (Figure 5.24(b)). Les oppositions de couleurs sont donc préservées, elles peuvent avoir la même intensité qu'en fovea. Cela correspond à ce qui aurait été mesuré au niveau physiologique [Martin 01].

Un ajustement des poids selon les densités locales ?

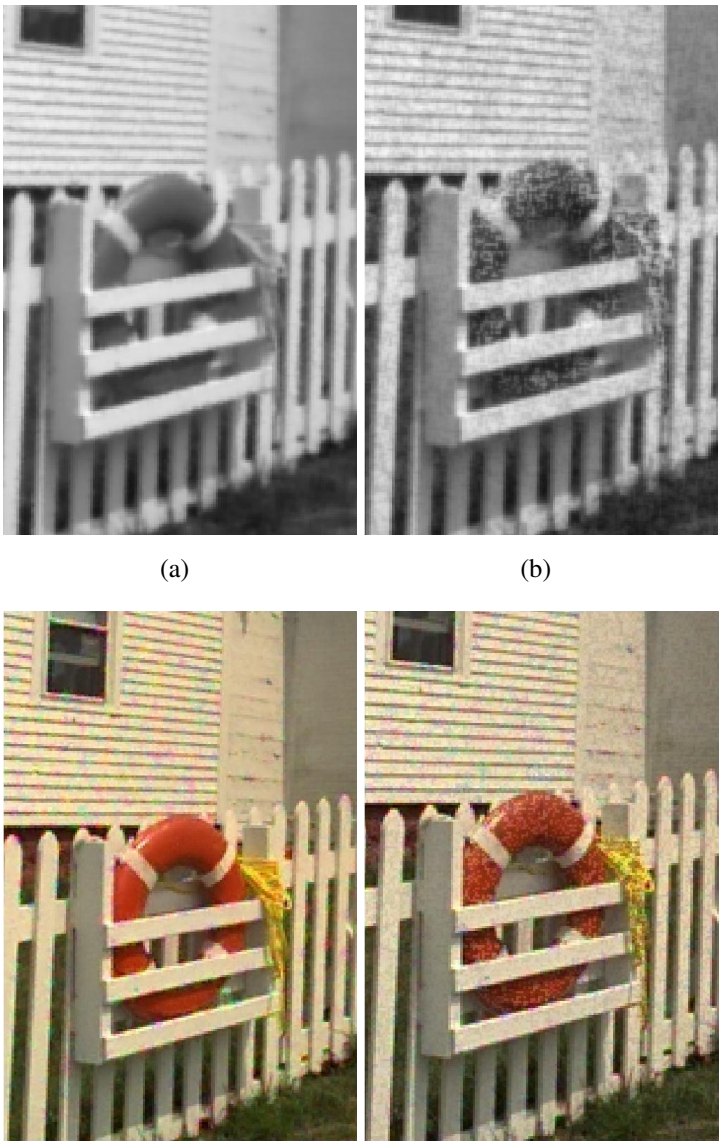


FIG. *5.25: (a) l'intensité basse fréquence est estimée par convolution normalisée, (b) par filtrage uniforme. L'absence de normalisation locale du filtre se traduit par l'apparition de bruit. Cela se ressent au niveau du PSNR des images reconstruites par la méthode linéaire (a-b) : 31.77dB et 30.71dB respectivement. Constante d'espace du filtre :* $k_0 = 1$ *pixel.*

D'après ce que nous avons vu sur l'interpolation par convolution normalisée, il semble nécessaire que les poids provenant des cônes L et des cônes M soient bien adaptés (leur somme doit être égale à p_L et p_M respectivement). Cela impliquerait un apprentissage de la rétine afin d'ajuster correctement ces poids, ce qui est parfaitement envisageable. Cependant il est intéressant de remarquer que le niveau de bruit qui apparaît si on ne normalise pas ces poids dépend de la fréquence de coupure du filtre que l'on choisit. Si le filtre utilisé est très passe-pas, alors la normalisation ou non des poids a une très faible influence sur le résultat (Figure 5.26). Au contraire, si on souhaite extraire des fréquences d'intensité plus élevée, la normalisation aura une plus grande importance (Figure 5.25). Cela se comprend facilement en considérant la taille

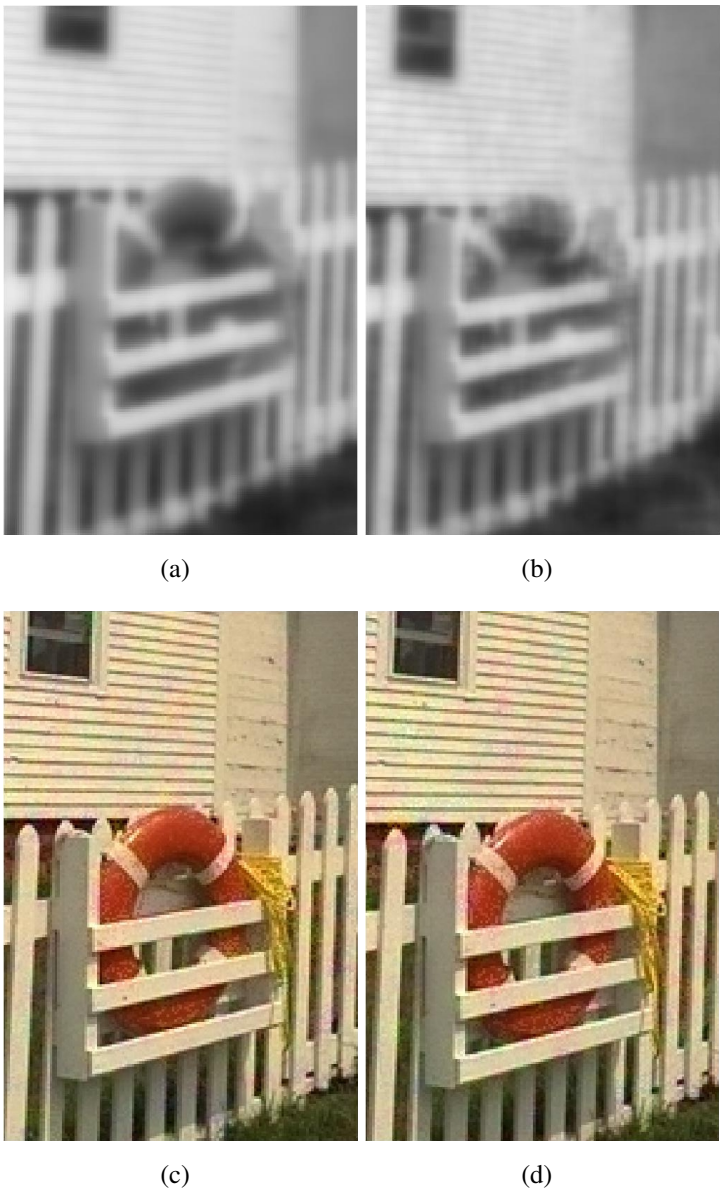


FIG. *5.26: (a) Intensité basse fréquence estimée par convolution normalisée, (b) par filtre uniforme. Lorsque le filtre est très passe-bas, la normalisation locale du filtre n'as plus grande importance. (a-b) Les images reconstruites par la méthode linéaire ont des PSNR identiques : 35.09dB et 35.05dB respectivement. La constante d'espace du filtre passe-bas vaut* $k_0 = 3$ *pixels.*

de la réponse impulsionnelle d'un filtre passe-bas. Si la fréquence de coupure est très basse, alors sa réponse impulsionnelle est grande, effectuant une moyenne sur un plus grand nombre d'échantillons. La densité d'échantillons est alors à une échelle plus globale. Dans le cas limite, si notre filtre passe-bas calcule l'intensité moyenne de toute l'image, alors les densités d'échantillons sont strictement égales à p_L et p_M, et il n'est pas besoin de normaliser. Si la fréquence de coupure est plus élevée, alors le voisinage considéré est plus restreint, et les variations de densités locales seront plus élevées d'un pixel à l'autre, nécessitant une normalisation afin de réguler ces variations. Ainsi, il est tout à fait envisageable, selon les fréquences de coupures des filtres rétiniens, que les connexions n'aient pas à être ajuster en fonction du voisinage proche.

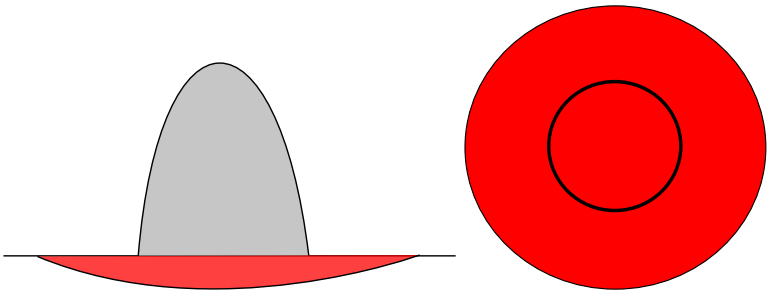


FIG. *5.27: Schéma du champ récepteur des cellules ganglionnaires parasol (coupe et vue de dessus). Le centre reçoit une information des trois types de cônes. Le pourtour retranche une information chromatique, c'est-à-dire qu'il connecte spécifiquement une classe de cône avec une pondération positive et en connecte une autre avec une pondération négative.*

Ainsi il est possible que la rétine n'ait pas à ajuster le poids de ses connexions en fonction des densités locales si le filtrage est suffisamment passe-bas. En contrepartie, cela génère beaucoup de bruit chromatique sur l'intensité (ou effet de grille), et la chrominance est très passe-bas. Il faudrait alors que ce système soit allié à un mécanisme post-rétinien adaptatif.

Il est à noter que l'utilisation d'un filtre très passe-bas permet à la fois de réduire les fausses couleurs (transformées en un bruit achromatique), de récupérer de plus hautes fréquences d'intensité, et de s'affranchir de la nécessité de pondérer les connexions rétiniennes en fonction des densités de photorécepteurs. Par ailleurs, cela rendrait possible l'utilisation d'un filtrage récursif au niveau de la rétine, car le filtrage récursif est incompatible avec une normalisation spécifique à chaque canal, comme nous l'avons vu au chapitre précédent.

4.1.2 Cellules ganglionnaires parasol : une intensité à laquelle on retranche un résidu de chrominance

Nous avons vu dans le paragraphe précédent que dans le cas d'un arrangement irrégulier des photorécepteurs, un filtrage suffisamment passe-bas permet de récupérer un signal d'intensité dépourvu de bruit de chrominance, sans avoir recourir à la convolution normalisée. Nous avons supposé que c'était en accord avec la faible largeur de bande de la chrominance.

Il est rapporté cependant, dans la littérature, que le signal achromatique véhiculé par la voie magnocellulaire a une meilleure résolution spatiale que les signaux d'oppositions de couleurs [Livingstone 02]. Cela signifie que le signal magnocellulaire doit faire face au problème de bruit de chrominance. Nous voyons deux solutions à ce problème. L'une serait d'effectuer une convolution normalisée. L'autre solution serait d'utiliser le filtre décrit dans le paragraphe 3.1 par l'Equation (5.16), en filtrant l'expression obtenu par un passe-bas et que nous appelons ici g. Cette méthode permet de monter plus haut en fréquence que la méthode par convolution normalisée, sans toutefois créer d'effet de grille, grâce au filtre passe-bas g. Le champ récepteur centré sur un cône L - mais dont le centre embrasse plusieurs cônes - serait donc défini de la manière suivante :

$$f_{lum,L} = \underbrace{g}_{centre} - \underbrace{g * ((1-p_L)f^L - p_M f^M - p_S f^S)}_{\text{pourtour inhibiteur}} \tag{5.27}$$

Le champ récepteur serait alors complexe, composé d'un centre achromatique (correspondant

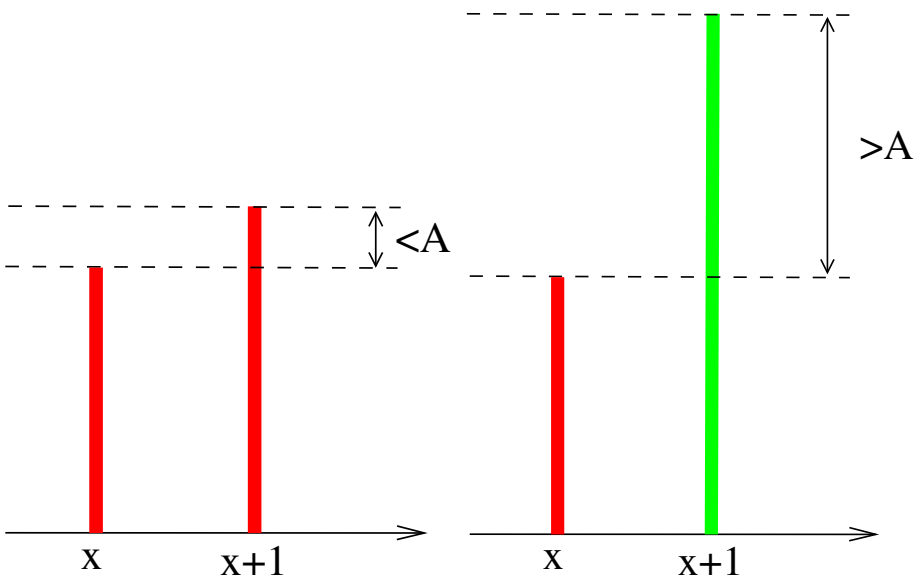


FIG. *5.28: (a) Lorsque deux cônes voisins de même classe échantillonnent un stimulus, qu'il soit chromatique ou achromatique, la différence d'amplitude de leur réponse est contrôlée par la fréquence de coupure de l'optique. (b) Lorsque deux cônes voisins sont de classes différentes, un stimulus chromatique induit systématiquement une différence de gradient entre leurs réponses.*

au filtre passe-bas g) et d'un pourtour inhibiteur impliquant à la fois des oppositions de couleurs et un filtrage passe-bas achromatique (Figure 5.27). Cette inhibition permet de compenser le bruit chromatique qu'il y aurait si le champ récepteur avait été composé uniquement de son centre. Un tel champ récepteur pourrait correspondre aux mesures électrophysiologiques relevées par Smith *et al.* pour lesquelles ni l'hypothèse de sélectivité ni l'hypothèse de non discrimination ne sont entièrement statisfaisantes [Smith 92].

4.2 L'implication du cortex visuel primaire

Dans le modèle que nous proposons, l'information spatiale véhiculée par la voie magnocellulaire pourrait être mise à profit pour obtenir des informations d'oppositions de couleurs dont la localité spatiale serait mieux définie. Comme les voies magno et parvocellulaires sont clairement distinctes jusqu'au cortex visuel primaire, c'est dans ce dernier que devrait avoir lieu ce mécanisme (ou peut-être au niveau du corps genouillé latéral, puisque celui-ci reçoit de nombreuses rétroactions provenant du cortex).

Johnson *et al.* obtiennent un grand nombre de cellules sensibles aussi bien à l'intensité qu'à la couleur, et ce pour des fréquences intermédiaires, en utilisant des stimuli dont le contraste en couleur est aussi fort que le contraste en intensité [Johnson 01]. Cela pourrait correspondre à l'effet de grille que nous obtenons, c'est-à-dire un débordement des hautes fréquences des oppositions de couleur sur l'information intensité.

4.3 Apprentissage non supervisé de la mosaïque des cones

Le modèle décrit dans la section précédente nécessite un filtrage chromatiquement sélectif (en d'autres termes, il nécessite un démultiplexage des sous-mosaïques suivi d'un filtrage) au niveau du cortex visuel. Les neurones ont donc besoin de discriminer les différentes classes d'information chromatique. En d'autres termes, il est nécessaire de savoir si, à telle position spatiale, le signal provient d'un cône de type L, M ou S. D'un autre côté, il semble que la différentiation des cônes en L et M soit un processus aléatoire et indépendant du voisinage [Wernet 06].

Nous montrons ici qu'il est possible de retrouver cette mosaïque chromatique par apprentissage non supervisé en présentant des images couleur devant la rétine. Ahumada et Mulligan ont travaillé sur cette problématique et ont proposé une solution [Ahumada 90] dans le cas d'une mosaïque composée de cônes L et M, où les poids d'un réseau de type Hebbien sont ajustés en utilisant des translations de l'image d'entrée. L'algorithme que nous proposons ne nécessite pas de translation d'une image, mais juste la présentation devant la rétine d'une série d'images suffisamment représentative de la statistique des images naturelles.

Afin de comprendre le mécanisme de cet apprentissage, supposons dans un premier temps que la fréquence de coupure f_c de l'optique de l'oeil soit exagérément basse - nous considérerons le cas général dans un second temps. Dans ce cas, l'image échantillonnée par une même classe de cône ne peut pas contenir d'énergie à des fréquences supérieures à f_c. Par ailleurs, on peut décomposer le signal en série de Taylor au point $x + k$,

$$C_i(x + k) = C_i(x) + kC_i'(x) + O(k^2)$$

Cela signifie que la variation maximale du signal entre les points x et $x + k$ est contrôlée par l'amplitude de la dérivée du signal au point x. On a donc, quelle que soit la position x :

$$|C_i(x + k) - C_i(x)| \leq kA \tag{5.28}$$

A étant l'amplitude maximale de la dérivée. L'amplitude maximale, qui est reliée au contenu fréquentiel du signal (car $(cos(\omega x))' = -\omega sin(\omega x)$). Plus un signal contient de hautes fréquences, plus sa dérivée est énergétique.

Or, et c'est la base de notre modèle d'échantillonnage, la mosaïque globale des cônes va avoir pour effet de moduler l'information relative à la couleur vers les hautes fréquences du spectre, autour des fréquences d'échantillonnage des sous-mosaïques. Il n'y a alors aucune ambiguïté entre une fréquence associée à l'intensité et une fréquence associée à la chrominance[2]. C'est cette propriété que nous utilisons pour distinguer de cônes voisins appartenant à deux classes différentes. Dans le domaine spatial, cette propriété sur les fréquences se traduit par une propriété sur l'amplitude des variations (Equation (5.28)). Ainsi, c'est par le biais de la valeur du gradient entre deux photorécepteurs $C(x)$ et $C(x + k)$ que nous pouvons savoir s'ils appartiennent à la même classe de cônes (Figure 5.28) :

$$\begin{cases} |C(x + k) - C(x)| > kA \Longrightarrow \text{classes distinctes} \\ |C(x + k) - C(x)| \leq kA \Longrightarrow \text{même classe} \end{cases} \tag{5.29}$$

[2]la chrominance ne peut pas dépasser f_c non plus, donc la chrominance modulée ne peut pas descendre en-dessous de $f_e - f_c$ (avec f_e la fréquence d'échantillonnage).

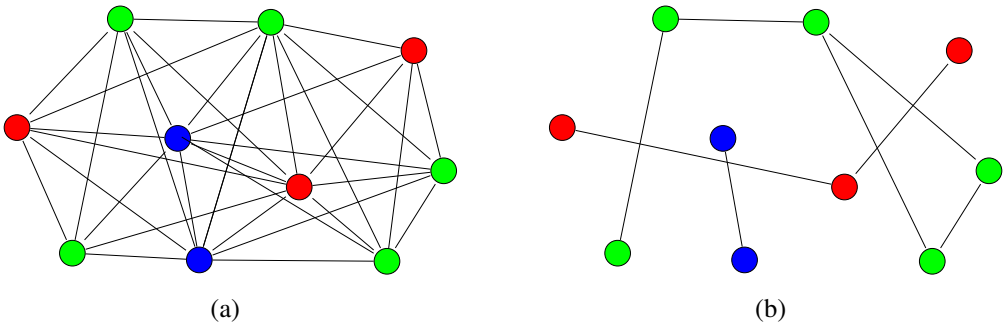


FIG. *5.29: Figure illustrant le principe de l'algorithme d'apprentissage. Chaque photorécepteur est initialement connecté à chacun de ses voisins (a). Chaque liaison se comporte comme un fusible qui cède lorsque le gradient entre deux photorécepteurs est trop grand. Lorsque l'algorithme à convergé, 3 sous-graphes apparaissent, correspondant aux 3 sous-mosaïques (b).*

Le principe de la classification basé sur ce critère de gradient est le suivant (Figure 5.29). Chaque photorécepteur est initialement connecté à tous ses voisins. Les connexions se comportent comme des fusibles vis-à-vis du gradient. Tant que le gradient est inférieur à un seuil, la connexion est maintenue. Si le gradient est trop fort, la connexion cède. Après convergence, trois sous-graphes apparaissent, ils correspondent aux trois sous-mosaïques.

Bien sûr, l'hypothèse sur le comportement de l'optique est trop restrictive par rapport à la réalité. Cependant, on peut supprimer cette hypothèse en considérant les statistiques des images naturelles. Si l'optique ne joue plus le rôle de filtre passe-bas, alors une forte variation dans l'image peut être attribuée aussi bien à une variation en intensité entre deux photorécepteurs quelconques (un contour par exemple) qu'à une variation chromatique due à la transition d'un type de photorécepteur à un autre (une zone homogène colorée par exemple). Cependant, la probabilité d'être sur un contour d'intensité est bien inférieure à celle d'être sur dans zone colorée. En effet, le nombre de pixels associés à des contours est plus faible que le nombre de pixels associés à des surfaces. Cette différence statistique entre contours et couleur va nous permettent d'utiliser notre critère basé sur le gradient. Les connexions qui étaient binaires dans le cas précédent (il y a connexion ou il n'y a pas connexion entre deux photorécepteurs) deviennent maintenant continues : on décrémente la "vraisemblance" qu'il y ait une connexion entre deux photorécepteurs selon l'amplitude du gradient entre les deux réponses.

$$\begin{cases} |C(i) - C(j)| \geq seuil_t \Longrightarrow P_{t+1}(i,j) = P_t(i,j) - \alpha \\ |C(i) - C(j)| \leq seuil_t \Longrightarrow \text{on ne fait rien} \end{cases} \tag{5.30}$$

où chaque t correspond à un stimulus différent.

Cet algorithme permet de retrouver les trois classes de photorécepteurs, sans déterminer toutefois laquelle correspond à L, à M ou à S. Pour pouvoir nommer explicitement ces classes, il faudrait, par exemple, tenir compte des statistiques de chaque classe.

5 Conclusion

Nous avons vu dans ce chapitre que l'on pouvait, quelle que soit l'arrangement de la mosaïque, effectuer un démosaïçage linéaire par estimation de l'intensité et des oppositions de

couleurs dans le plan de Fourier, comme c'était déjà le cas avec la mosaïque de Bayer. Nous avons estimé les filtres d'intensité optimaux au sens des moindres carrés par apprentissage sur une base d'image, pour différents arrangement de la mosaïque. Cette méthode permet de comparer objectivement les performances des différents motifs. Il en ressort que la mosaïque de Bayer, pourtant la plus utilisée dans le domaine de l'imagerie mono-capteur, n'est pas en tête. Elle est devancée par la mosaïque proposée par Lukac et par une mosaïque que nous proposons. L'inconvénient majeur du motif de Bayer est qu'il ne permet pas d'avoir une résolution maximale dans les directions horizontales et verticales, du moins si on emploie une méthode de démosaïçage linéaire. La méthode d'apprentissage peut être également utilisée pour calibrer les filtres d'intensité d'une caméra, si on dispose également d'un capteur achromatique. La base d'apprentissage est alors constituée d'images prises par la caméra avec la matrice de filtres couleur (ces images constituent les images en mosaïque) et des mêmes images prises par la caméra avec le capteur achromatique (ces images constituent les images d'intensité).

La mosaïque doit cependant être constituée du pavage d'un motif de base de quelques pixels de côté pour que l'apprentissage puisse être calculé en pratique. Nous avons donc également étendu le modèle d'Alleysson *et al.* à tout arrangement de mosaïque, sans contrainte sur l'arrangement. Lorsque l'on applique le modèle à la mosaïque de Bayer, on retrouve bien le filtre d'intensité vers lequel a convergé l'apprentissage aux moindres carrés. Cependant il existe une légère différence, pour une mosaïque à motif pseudo-irrégulier, entre la réponse fréquentielle des filtres issus du modèle et celle des filtres par apprentissage. Cela se ressent au niveau des PSNR des images reconstruites. En particulier, les filtres obtenus par apprentissage produisent un très faible niveau d'effet de grille.

Nous proposons deux approches possibles de démosaïçage, en passant soit par une estimation de l'intensité complète, soit par une estimation de l'intensité basse fréquence. La stratégie passant par l'intensité basse fréquence a l'avantage d'être spatialement invariante et d'utiliser des filtres isotropes. Nous en avons fait une implémentation à l'aide de filtres récursifs séparables. Ces filtres récursifs utilisent l'information des quatre plus proches voisins, à comparer avec le filtre d'intensité à réponse impulsionnelle finie obtenu par apprentissage qui doit être de taille 7×7 voire 9×9. Nous proposons, de surcroît, une version adaptative qui permet de réduire l'effet de grille.

La nature aléatoire de l'arrangement des cônes à la surface de la rétine peut conférer de bonnes propriétés en terme de repliement spectral. Cependant, la présence de certains amas d'un seul type de photorécepteur ne nous apparaît pas optimal dans nos simulations. Cela provient du fait que la chrominance est alors modulé à de très basses fréquences et interfère avec l'intensité. Un arrangement optimal serait, comme nous l'avons vu, une disposition aléatoire mais avec une certaine alternance entre les photorécepteurs afin d'éviter ce genre d'amas.

Nous avons également les implications de notre modèle au niveau du système visuel. Il en ressort qu'il n'est pas besoin d'établir des connexions chromatiques spécifiques au niveau de la rétine pour la construction de la voie parvocellulaire. Un traitement achromatique au niveau de la rétine permet d'obtenir les antagonismes spectraux mesurés au niveau des champs récepteur des cellules ganglionnaires. D'autre part, l'information achromatique véhiculée par la voie magnocellulaire permet d'améliorer la séparation entre l'information d'intensité et celle d'opposition de couleurs rouge-vert de la voie parvocellulaire. Cela va dans le sens des études

rapportant une facilitation de la perception d e la couleur par un stimulus achromatique coïncident.

Ingling et Martinez avaient bien compris, en se basant sur la forme des champs récepteur, que la chrominance était localisée en basses fréquences et que l'intensité était passe-bande. C'est explicite chez eux, avec la transformée de Fourier du centre et du pourtour du champ récepteur qui montre des supports fréquentiels disjoints. Néanmoins, ce modèle est incomplet car il ne prend pas en compte le multiplexage spatial. Ainsi, si l'on considère la représentation fréquentielle du signal au niveau des cellules ganglionnaires, comme le font Ingling et Martinez, il est erroné de représenter uniquement des oppositions de couleur en basses fréquence et une intensité passe-bande. Il faut également rajouter tous les réplicats de ces signaux aux fréquences d'échantillonnage. On comprend alors la difficulté à séparer les deux canaux spatial et chromatique, et comment peuvent apparaître des fausses couleurs. De plus, le modèle que nous proposons explicite les composantes d'intensité et d'oppositions de couleurs en fonction des densités p_i des photorécepteurs.

Au niveau de la voie magnocellulaire, nous proposons un champ récepteur plus complexe composé de deux étage. Un étage effectuerait un traitement achromatique - donc sans connexion spécifique - et embrasserait à la fois le centre et le pourtour du champ récepteur, effectuant un simple filtrage passe-bas. Un autre étage retrancherait l'information de couleur par des connexions spécifiques dans le pourtour du champ récepteur. Un tel champ récepteur pourrait être en accord avec les mesures relevées par Smith *et al*, dont ni l'hypothèse de spécificité, ni celle de connexions aléatoires ne sont satisfaisantes.

Chapitre 6

Conclusion et perspectives

LES travaux scientifiques, depuis ceux de Newton, ont permis de mettre en évidence que la vision des couleurs a pour origine les cônes, neurones répartis en trois classes à la surface de la rétine, chaque classe étant sensible à une gamme de longueurs d'onde. Le signal issu des trois classes de cônes est ensuite transformé par le système visuel en un canal de luminance et deux canaux d'opposition de couleurs. Le sous-échantillonnage effectué par chaque sous-mosaïque de cônes - il n'y a qu'un cône par position spatiale - pose un problème complexe d'interpolation. Alleysson *et al.* ont montré que le choix des axes achromatique et d'oppositions de couleurs était judicieux du fait de la représentation de ces signaux dans le plan de Fourier. La faible largeur de bande des oppositions de couleurs permet une grande acuité en luminance.

Cependant, les mécanismes de construction de ces voies perceptuelles sont encore mal compris. En particulier, l'intéraction entre les différentes classes de neurones rétiniens n'est pas clairement établie. L'arrangement irrégulier des cônes augmente encore la difficulté du problème. Le modèle d'Alleysson offre une solution élégante dans le cas d'une mosaïque régulière, mais échoue lorsque celle-ci est irrégulière.

Des traitements rétiniens...

Nous avons présenté dans cette thèse une solution indépendante de l'arrangement de la mosaïque. Il faut toutefois connaître la mosaïque, car un démultiplexage de l'information chromatique est nécessaire à un moment ou à un autre. A partir de la solution proposée, nous émettons l'hypothèse qu'il n'y a pas de traitement chromatique dans l'élaboration de la voie parvocellulaire. Les traitements rétiniens intervenant dans la construction de cette voie seraient achromatiques, mais prépareraient l'information pour l'estimation ultérieure (*i.e.* post-rétinienne) de l'information d'opposition rouge-vert. Ainsi, aucune connexion chromatique spécifique n'est nécessaire pour cette voie.

Concernant la voie achromatique magnocellulaire, il y a deux approches possibles. L'estimation d'une information d'intensité peut être effectuée par un filtrage passe-bas uniforme du signal issu de la mosaïque des cônes. Le caractère irrégulier de l'arrangement spatial des cônes introduit néanmoins du bruit de grille. Une solution pour éliminer ce problème consiste à pondérer les coefficients du filtre par les densité locales en photorécepteurs de chaque classe. Cela signifie que le champ récepteur des cellules ganglionnaires parasol établierait des connexions

non spécifiques, mais que le poids des connexions serait localement adapté pour réguler la contribution de chaque canal au signal d'intensité. La seconde solution, qui a ma préférence, est de retrancher des oppositions de couleurs estimée de manière *uniforme* puis de filtrer le résultat passe-bas de manière *uniforme*. Cette solution permet d'obtenir une meilleure résolution d'intensité, tout en s'affranchissant du bruit de grille. Pour résumer, nous proposons donc, de manière quelque peu contre-intuitive de prime abord, un traitement achromatique pour obtenir un canal véhiculant des oppositions de couleurs, et un traitement chromatique pour obtenir un canal achromatique.

Quel serait le "coût écologique" des types de connexions que nous proposons pour les voies parvo et magnocellulaire ? Nous faisons ici des considérations sur le coût écologique de manière intuitive, en considérant que chaque type de cellule, chaque type de connexion etc... est affecté d'un coût, et que la nature, par le biais de l'évolution, réduit au minimum ces coûts tout en assurant la survie de l'espèce. Par exemple, les cellules de la voie magnocellulaire sont peut-être coûteuses (elles transmettent l'information plus rapidement que les cellules de la voie parvocellulaire), mais ce coût est contrebalancé par leur faible proportion. L'établissement de connexions chromatiquement spécifiques nous apparaît également plus coûteux que des connexions non discriminantes. En effet, des connexions spécifiques signifient soit que la position de chaque cône est explicitement connue et codée dans l'ADN, soit qu'un mécanisme d'apprentissage intervient, ou encoreque chaque cône libère une messager chimique afin d'être identifié.

Les cellules de la voie parvocellulaire sont présentes en large majorité. Il est donc très économique, d'après notre modèle, qu'elles n'aient pas besoin d'établir de connexions spécifiques. Concernant les cellules de la voie magnocellulaire, la solution d'un filtre passe-bas non discriminant est moins coûteuse, cependant la résolution spatiale s'en trouve affectée. La solution basée sur des connexions spécifiques (plus précisément, le signe du poids de la connexion dépend de quel cône provient le signal) est plus coûteux, mais offre une meilleure résolution. Le surcoût est peut-être contrebalancé par la faible proportion de ces cellules par rapport à celle des cellules de la voie parvocellulaire.

... aux traitements post-rétiniens.

Par ailleurs, il n'est pas possible, d'après nos simulations, d'estimer correctement l'intensité et les oppositions de couleurs par filtrage linéaire. La raison est le recouvrement spectral entre entre les signaux achromatique et chromatiques, qui a lieu par exemple le long de contours francs. Nous en concluons qu'un traitement adaptatif doit avoir lieu à un niveau post-rétinien. Ce traitement adaptatif pourrait mettre à profit l'information d'intensité de la voie magnocellulaire pour estimer les hautes fréquences manquantes de la chrominance. Cette hypothèse est en accord avec les travaux rapportant la facilitation de la perception de stimuli chromatiques par l'adjonction de stimuli achromatiques.

Un tel traitement adaptatif, impliquant à la fois la voie magnocellulaire et la voie parvocellulaire pourrait être à l'origine de l'illusion "aquarelle" (*watercolor illusion*). Ce genre d'illusion serait alors la manifestation d'un mécanisme anti-repliement spectral. L'illusion aquarelle apparaît lorsqu'une ligne fermée achromatique est flanquée du côté intérieur d'un contour chro-

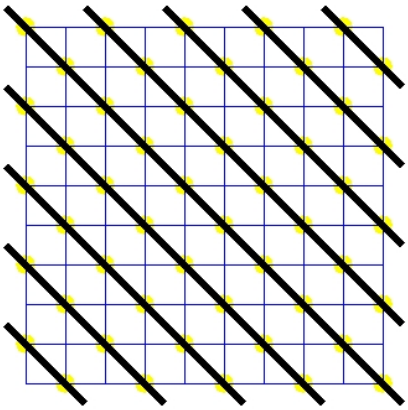


FIG. *6.1: Lignes illusoires de Kitaoka [Kitaoka 01]. Des lignes jaunes illusoires apparaissent, orthogonalement aux lignes noires. De plus, lorsque le regard est bien stabilisé, la couleur jaune a tendance à s'atténuer.*

matique. Alors, une couleur est induite du côté à l'intérieur de la forme. Le traitement adaptatif se baserait, dans cet exemple précis, sur le contour achromatique, qui est le seul contour visible par la voie magnocellulaire, pour déterminer les bords de l'objet. Un mécanisme de remplissage orthogonal au contour achromatique interviendrait alors, afin que la chrominance soit basse fréquence tout en ayant des transitions franches aux bords de l'objet. Il semble important que ce mécanisme soit orthogonal aux contours, afin de pouvoir rendre compte de l'illusion observée Figure 6.1. Dans cette illusion des lignes de couleur apparaissent orthogonalement aux lignes achromatiques. Il serait intéressant donc d'implémenter la méthode de démosaïçage adaptative proposée dans cette thèse de telle manière que ces illusions visuelles soient retrouvées.

Le disque de Benham - disque achromatique induisant des couleurs lorsqu'il est en mouvement - a peut-être lui aussi une origine provenant de l'intéraction entre les voies magno et parvocellulaire. Dans cette illusion, la dimension temporelle entre en jeu. Il serait donc intéressant d'étudier les conséquences des retards temporels entre les voies cellulaires sur la constructions des oppositions de couleur et de l'intensité. Il est possible que les couleurs induites soient la conséquence d'une incohérence entre l'information magnocellulaire et la voie parvocellulaire.

Nous avons également vu que l'échantillonnage irrégulier possédait de bonnes propriétés de réduction de la visibilité du phénomène de Moiré. Cependant, la présence de certains amas d'un seul type de photorécepteur - si les travaux sur le sujet le confirment - ne nous apparaît pas optimal dans nos simulations. Il est possible qu'un tel arrangement soit trop coûteux à obtenir pour la rétine. Il faudrait utiliser un mécanisme permettant l'optimisation de l'arrangement des photorécepteurs lors de leur migration au cours du développement de la rétine. Le fait qu'un tel mécanisme soit très complexe et le fait que la différenciation en cônes L et M soit de nature aléatoire nous amènent à penser que le système visuel a préféré, d'un point de vue du coût écologique, gérer les problèmes induits par les amas de photorécepteurs plutôt que d'éviter ces amas. Une des manières par laquelle l'évolution pourrait avoir résolu ce problème serait, par exemple, l'utilisation des microsaccades oculaires.

Parmi les problèmes que la communauté scientifique de vision des couleurs doit résoudre, on peut citer également le problème de la teinte jaune unique. Des travaux récents ont mon-

tré que la position spectrale du jaune unique était la même chez tous les individus. D'après le modèle que nous avons développé, l'intensité et les oppositions de couleur dépendent de la proportion de chaque type de cône. Ces proportions varient d'un individu à l'autre. Il est donc surprenant que ce ne soit pas le cas pour les teintes uniques. Le mécanisme régulant leur position spectrale reste à découvrir.

Aspects technologiques

Les modèles de vision des couleurs trouvent "naturellement" des applications au niveau des caméras numériques. Il semble évident que les progrès sur la compréhension des traitements effectués par le système visuel humain seront transférés au domaine des capteurs d'image grand public. Le grand public est en effet le plus exigeant vis-à-vis de la qualité des photographies qu'il prend. Il souhaite en effet que son appareil reproduise la scène exactement tel qu'il la voit, c'est-à-dire, bien souvent, dans des conditions qu'un photographe expérimenté qualifierait de mauvaises. En ce sens l'étude du système visuel est une ressource pour le marché des caméras numériques.

Nous avons également développé, au cours de cette thèse, un algorithme linéaire basé sur la minimisation de l'erreur quadratique. Cette méthode nous a permis de faire un comparatif des performances de différentes mosaïques. Il en ressort que la mosaïque de Bayer n'est pas la mosaïque optimale. Cependant, l'inertie de cette dernière dans l'industrie ne devrait pas remettre sa suprématie en cause. La télévision est également un domaine d'application pour une mosaïque améliorée. Un écran de télévision est en effet composé de la juxtaposition de phosphores rouges, verts et bleus. C'est donc une matrice de filtres de couleur. La différence avec les caméras numériques réside dans le fait que le mélange des couleurs s'effectue avant la projection de l'image sur la rétine. Ainsi, dans ce domaine, l'arrangement spatial des phosphores devrait avoir une forte influence sur la qualité de l'image affichée.

La liste des différences entre une caméra artificielle et le système visuel humain est longue. Le potentiel d'inspiration est donc vaste. Néanmoins, l'une d'entre elles est majeure : la camera est un dispositif passif de vision, tandis le système visuel est un dispositif actif. L'oeil est toujours en mouvement, en scrutation. L'oeil humain, avec les mouvements du capteur, couplé à un échantillonnage spatialement variant, est une source d'inspiration inouïe pour les systèmes artificiels. Des projets allant dans le sens de la vision active pour les caméras existent. Ces projets, mêlant sciences techniques et sciences cognitives, sont un bon exemple de transfert de technologie de l'humain vers les systèmes artificiels.

Science et art

Margaret Livingstone a publié de nombreux ouvrages sur l'art et la vision. Elle montre comment la biologie du système visuel permet d'interpréter les oeuvres d'art. Elle explique notamment l'expression mystérieuse de la joconde par le fait que son sourire n'a pas la même expression en basses fréquences qu'en hautes fréquences. Ainsi, selon la partie du champ visuel dans laquelle se trouve le sourire de la Joconde, le sourire est perçu différemment, lui conférant une expression dynamique [Livingstone 02].

Georges Seurat, pionnier du pointillisme, juxtaposait dans ses tableaux des touches de cou-

leur, créant ainsi une mosaïque (Figure 6.2). Lorsque l'on regarde ses oeuvres avec du recul, les touches chromatiques se mélangent, et on perçoit une véritable scène en couleur. Georges Seurat, qui s'intéressait de près aux théories de vision des couleurs, jouait sur les oppositions de couleurs, en juxtaposant par exemple des touches bleues à des touches jaunes. Il appelait lui-même sa démarche de *chromo-luminariste*. Il serait intéressant d'appliquer notre modèle aux oeuvres de Seurat et de comparer les résultats à ce que l'oeil humain en perçoit.

FIG. *6.2: Détail du tableau "La parade", illustrant la technique en mosaïque de George Seurat.*

Annexe A

Base d'images utilisée

La base d'images couleur utilisée dans cette thèse est la base Kodak, largement utilisée au sein de la communauté du démosaïçage :

Annexe B

Quelques exemples de résultats

1 Recontruction minimisant l'erreur quadratique moyenne par apprentissage sur la base d'images Kodak

Les figures B.1, B.2, B.3 et B.4 montre des exemples de reconstruction utilisant l'algorithme par apprentissage supervisé sur la base d'images Kodak, décrit au Chapitre 5, Section 2.

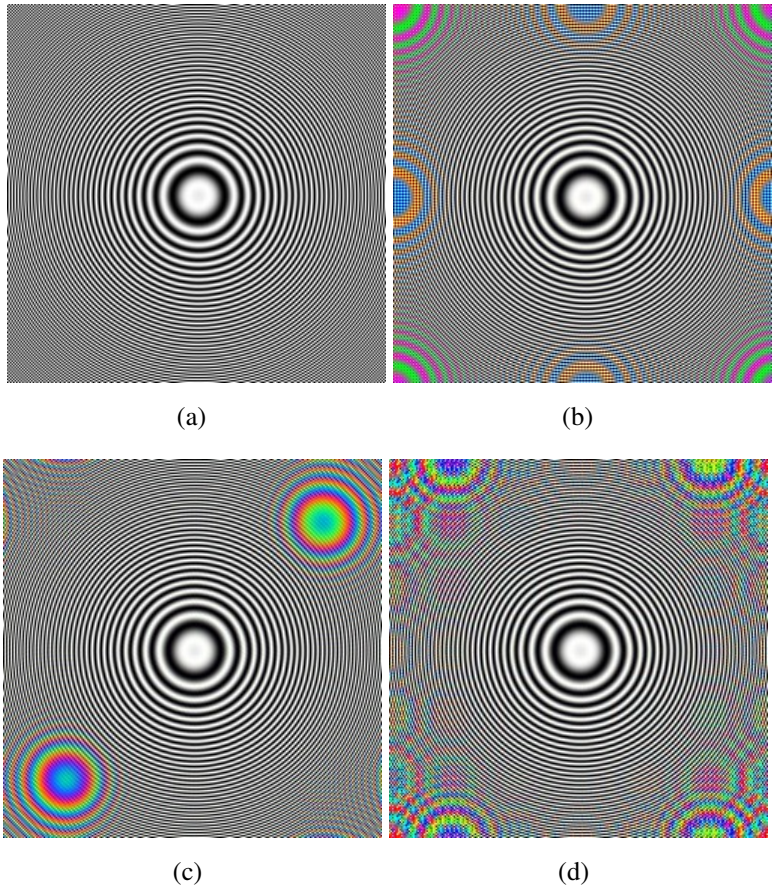
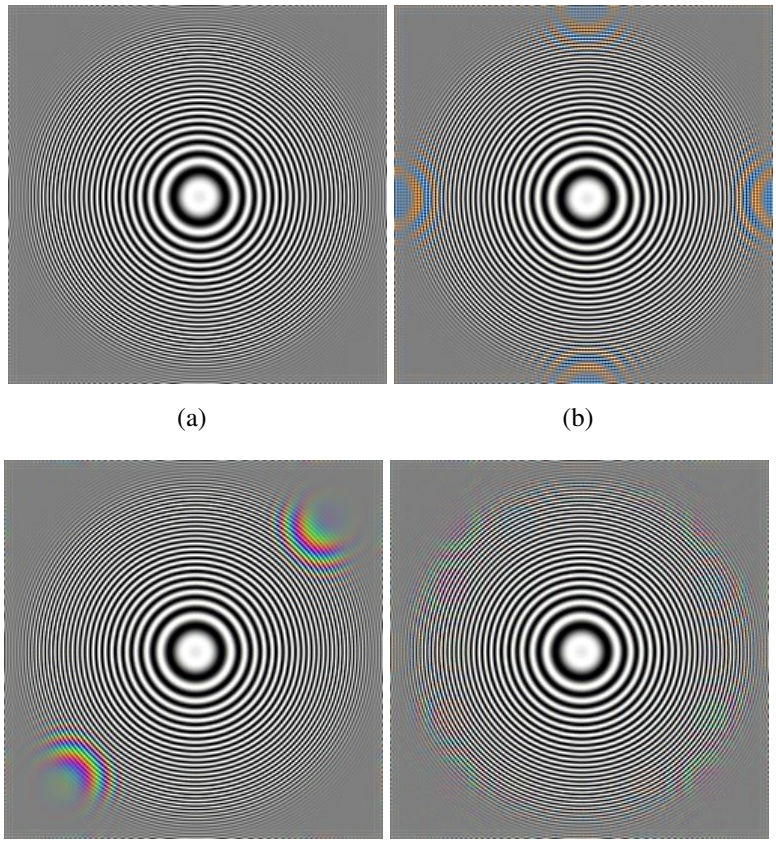


FIG. *B.1: Démosaïçage de l'image CZP (256x256 pixels). (a) Image originale (b-d) Reconstruction à partir de la mosaïque de Bayer, diagonale à bandes et de la mosaïque irrégulière proposée.*



(a) (b)

(c) (d)

FIG. *B.2: Démosaïçage de l'image CZP (256x256 pixels) avec simulation d'une optique circulaire passe-bas. (a) Image originale (b-d) Reconstruction à partir de la mosaïque de Bayer, diagonale à bandes et de la mosaïque irrégulière proposée.*

2 Reconstruction utilisant le modèle proposé

Les Figures B.5 et B.6 illustrent la reconstruction utilisant l'algorithme proposé au Chapitre 5, Section 3 (version linéaire et version adaptative).

(a) Originale

(b) Bandes verticales

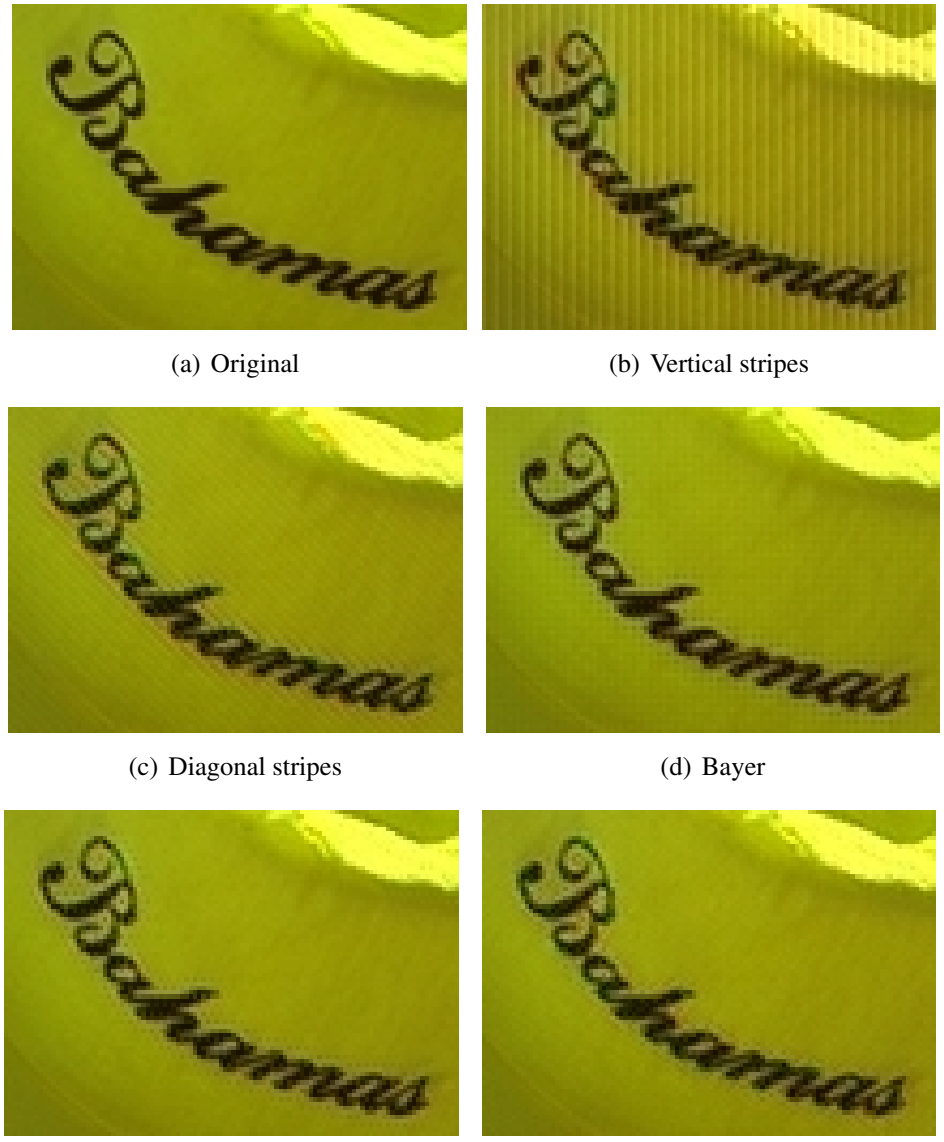
(c) Bandes diagonales

(d) Bayer

(e) Lukac

(f) Proposée (pseudo-irrégulière)

FIG. *B.3: (a) Extraits de l'image "phare" ; (b-f) Extraits reconstruits en utilisant diverses mosaïques.*

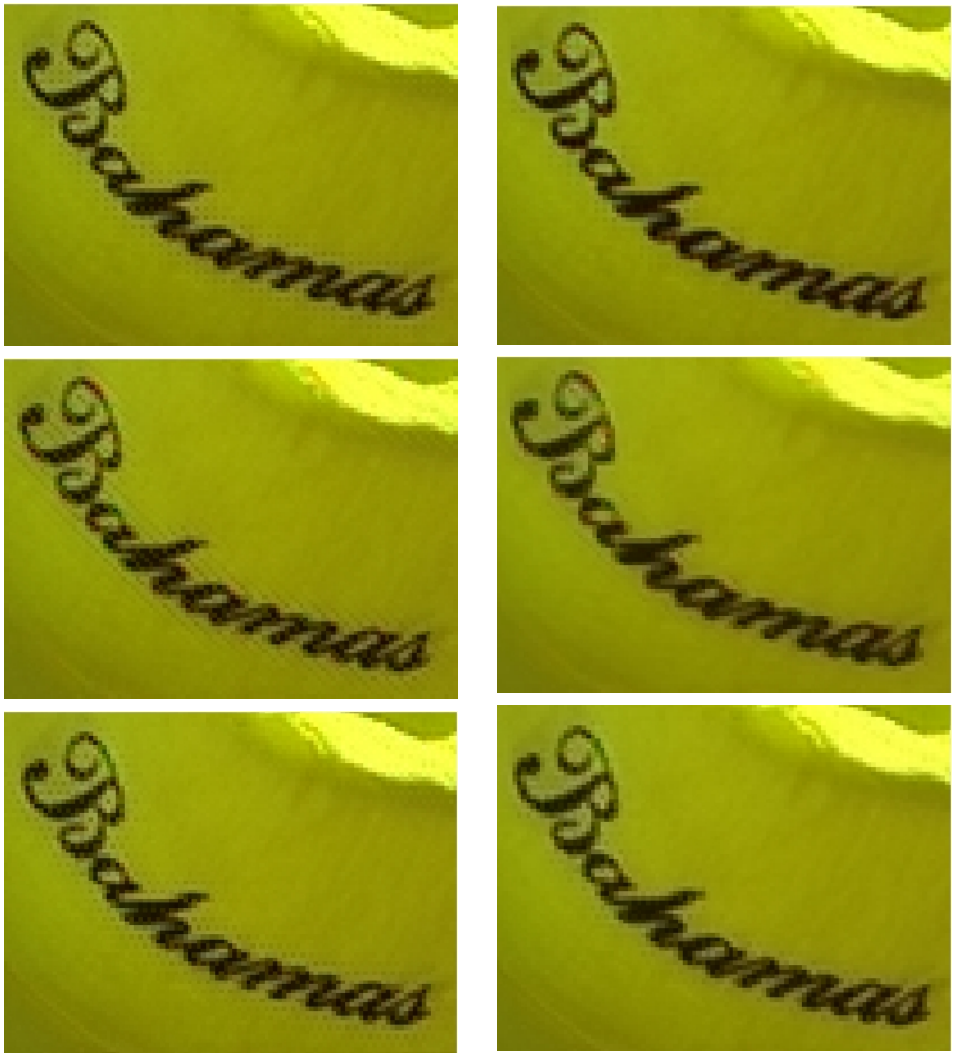


(a) Original (b) Vertical stripes

(c) Diagonal stripes (d) Bayer

(e) Lukac (f) Pseudo-random

FIG. *B.4: (a) Image Bahamas ; (b-f) Image reconstruite en utilisant diverses mosaïques.*



(a) (b)

FIG. *B.5: Image Bahamas reconstruite en utilisant la méthode linéaire (a) et la méthode adaptative (b). De haut en bas, les mosaïques utilisées sont les mosaïques de Bayer, à bandes diagonales et de Lukac*

(a)

(b)

(c)

FIG. *B.6: (a) Image originale ; (b) Reconstruction linéaire ; (c) Reconstruction adaptative.*

Bibliographie

[Adams 96] J.E. Adams & J.F. Hamilton. *Adaptive color plan interpolation in single sensor color electronic camera*. US patent 5,506,619, to Eastman Kodak Company, 1996. Cité page 67.

[Adams 97] J. E. Adams. *Design of practical color filter array interpolation algorithms for digital cameras*. In SPIE, volume 3028, pages 117–125, February 1997. Cité page 67.

[Ahumada 89] A. J. Ahumada & J. I. Yellott. *Reconstructing irregularly sampled images by neural networks*. In B. E. Rogowitz, editeur, Proc. SPIE Vol. 1077, p. 228, Human Vision, Visual Processing, and Digital Display, Bernice E. Rogowitz ; Ed., volume 1077 of *Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference*, pages 228–+, August 1989. Cité page 79.

[Ahumada 90] Albert J. Ahumada & Jeffrey B. Mulligan. *Learning Networks for the Red-Green Opponent System*. Investigative Ophthalmology and Visual Science, vol. 31, page 262, 1990. Cité page 147.

[Alleysson 03] D. Alleysson, S. Süsstrunk & J. Marguier. *Influence of spectral sensitivity functions on color demosaicing*. In Color Imaging Conference, pages 351–357, November 2003. Cité page 52.

[Alleysson 05] David Alleysson, Sabine Süsstrunk & Jeanny Hérault. *Linear color demosaicing inspired by the human visual system*. IEEE Transactions on Image Processing, vol. 14, pages 439–449, April 2005. Cité pages 37, 44, 50, 52, 61, 62, 69, 70, 106 et 115.

[Alonso 06] J-M. Alonso, C-I. Yeh, C. Weng & C. Stoelzel. *Retinogeniculate connections : A balancing act between connection specificity and receptive field diversity.* Prog Brain Res, vol. 154, pages 3–13, 2006. Cité page 23.

[Anderson 91] S. J. Anderson, K. T. Mullen & R. F. Hess. *Human peripheral spatial resolution for achromatic and chromatic stimuli : limits imposed by optical and retinal factors.* J Physiol, vol. 442, pages 47–64, Oct 1991. Cité pages 22, 39, 40 et 41.

[Awatramani 00] G. B. Awatramani & M. M. Slaughter. *Origin of transient and sustained responses in ganglion cells of the retina.* J Neurosci, vol. 20, no. 18, pages 7087–7095, Sep 2000. Cité page 31.

[Barstow 53] L. M. Barstow & H. N Chistopher. *The measurement of random monochrome video interference*. Transactions of the American Institute of Electrical Engineers, vol. 72, pages 735–741, 1953. Cité page 79.

[Bayer 76] B.E. Bayer. *Color imaging array*. US patent 3,971,065, to Eastman Kodak Company, 1976. Cité page 52.

[Beaudot 94] William Beaudot. *Le traitement neuronal de l'information dans la rétine des vertébrés : un creuset d'idées pour la vision artificielle*. PhD thesis, Institut National Polytechnique de Grenoble, 1994. Cité pages 35 et 135.

[Berson 03] David M Berson. *Strange vision : ganglion cells as circadian photoreceptors.* Trends Neurosci, vol. 26, no. 6, pages 314–320, Jun 2003. Cité page 34.

[Bland 97] Denise M. Bland & Andrzej Tarczynski. *The Effect Of Sampling Jitter In A Digitized Signal*. In IEEE International Symposium on Circuits and Systems, Hong Kong, June 1997. Cité page 83.

[Brainard 00] D. H. Brainard, A. Roorda, Y. Yamauchi, J. B. Calderone, A. Metha, M. Neitz, J. Neitz, D. R. Williams & G. H. Jacobs. *Functional consequences of the relative numbers of L and M cones.* J Opt Soc Am A Opt Image Sci Vis, vol. 17, no. 3, pages 607–614, Mar 2000. Cité page 30.

[Buck 00] S. L. Buck, R. F. Knight & J. Bechtold. *Opponent-color models and the influence of rod signals on the loci of unique hues.* Vision Res, vol. 40, no. 24, pages 3333–3344, 2000. Cité page 32.

[Calkins 96] D. J. Calkins & P. Sterling. *Absence of spectrally specific lateral inputs to midget ganglion cells in primate retina.* Nature, vol. 381, no. 6583, pages 613–615, Jun 1996. Cité pages 32, 33 et 35.

[Calkins 98] D. J. Calkins, Y. Tsukamoto & P. Sterling. *Microcircuitry and mosaic of a blue-yellow ganglion cell in the primate retina.* J Neurosci, vol. 18, no. 9, pages 3373–3385, May 1998. Cité pages 27 et 34.

[Calkins 99] D. J. Calkins & P. Sterling. *Evidence that circuits for spatial and color vision segregate at the first retinal synapse.* Neuron, vol. 24, no. 2, pages 313–321, Oct 1999. Cité page 45.

[Campbell 66] F. W. Campbell & R. W. Gubisch. *Optical quality of the human eye.* J Physiol, vol. 186, no. 3, pages 558–578, Oct 1966. Cité page 22.

[Chatterjee 03] Soumya Chatterjee & Edward M Callaway. *Parallel colour-opponent pathways to primary visual cortex.* Nature, vol. 426, no. 6967, pages 668–671, Dec 2003. Cité page 37.

[Clark 85] J.J. Clark, M.R. Palmer & P.D.Lawrence. *A Transformation Method for the Reconstruction of Functions from Nonuniformly Spaced Samples*. IEEE Transaction on Acoustics, Speech and Signal Processing, vol. 5, pages 1151–1165, 1985. Cité pages 77 et 79.

[Cok 87] D. R. Cok. *Signal processing method and apparatus for producing interpolated chrominance values in a sampled color image signal*. US patent 4,642,678, to Eastman Kodak Company, 1987. Cité page 58.

[Cole 90] G. R. Cole, C. F. Stromeyer & R. E. Kronauer. *Visual interactions with luminance and chromatic stimuli.* J Opt Soc Am A, vol. 7, no. 1, pages 128–140, Jan 1990. Cité page 41.

[Condat 05] L. Condat, T. Blu & M. Unser. *Beyond interpolation : optimal reconstruction by quasi-interpolation*. In Image Processing, 2005. ICIP 2005. IEEE International Conference on, volume 1, pages I–33–6, 11-14 Sept. 2005. Cité pages 98 et 99.

[Condat 06] Laurent Condat. *Méthodes d'approximation pour la reconstruction de signaux et le redimensionnement d'images*. PhD thesis, INPG, 2006. Cité page 98.

[Cottaris 03] Nicolas P Cottaris. *Artifacts in spatiochromatic stimuli due to variations in preretinal absorption and axial chromatic aberration : implications for color physiology.* J Opt Soc Am A Opt Image Sci Vis, vol. 20, no. 9, pages 1694–1713, Sep 2003. Cité page 22.

[Coulombe 96] S. Coulombe. *Transmultiplexage multidimensionnel à reconstruction parfaite et son application aux systèmes de télévision améliorés*. PhD thesis, INRS-Télécommunications, Université du Québec, 1996. Cité page 63.

[Curcio 90] C. A. Curcio, K. R. Sloan, R. E. Kalina & A. E. Hendrickson. *Human photoreceptor topography.* J Comp Neurol, vol. 292, no. 4, pages 497–523, Feb 1990. Cité pages 26 et 28.

[Curcio 91] C. A. Curcio, K. A. Allen, K. R. Sloan, C. L. Lerea, J. B. Hurley, I. B. Klock & A. H. Milam. *Distribution and morphology of human cone photoreceptors stained with anti-blue opsin.* J Comp Neurol, vol. 312, no. 4, pages 610–624, Oct 1991. Cité page 29.

[Dacey 94] D. M. Dacey & B. B. Lee. *The 'blue-on' opponent pathway in primate retina originates from a distinct bistratified ganglion cell type.* Nature, vol. 367, no. 6465, pages 731–735, Feb 1994. Cité pages 31 et 34.

[Dacey 96a] D. M. Dacey. *Circuitry for color coding in the primate retina.* Proc Natl Acad Sci U S A, vol. 93, no. 2, pages 582–588, Jan 1996. Cité pages 32, 33 et 35.

[Dacey 96b] D. M. Dacey, B. B. Lee, D. K. Stafford, J. Pokorny & V. C. Smith. *Horizontal cells of the primate retina : cone specificity without spectral opponency.* Science, vol. 271, no. 5249, pages 656–659, Feb 1996. Cité pages 21 et 35.

[Dacey 99] D. M. Dacey. *Primate retina : cell types, circuits and color opponency.* Prog Retin Eye Res, vol. 18, no. 6, pages 737–763, Nov 1999. Cité pages 33 et 34.

[Dacey 00a] D. M. Dacey. *Parallel pathways for spectral coding in primate retina.* Annu Rev Neurosci, vol. 23, pages 743–775, 2000. Cité pages 24, 30, 33 et 34.

[Dacey 00b] D. M. Dacey, L. C. Diller, J. Verweij & D. R. Williams. *Physiology of L- and M-cone inputs to H1 horizontal cells in the primate retina.* J Opt Soc Am A Opt Image Sci Vis, vol. 17, no. 3, pages 589–596, Mar 2000. Cité page 30.

[de Lavarène 07a] Brice Chaix de Lavarène, David Alleysson, Barthélémy Durette & Jeanny Hérault. *Efficicent demosaicing through recursive filtering*. In IEEE International Conference on Image Processing, 2007. Cité page 134.

[de Lavarène 07b] Brice Chaix de Lavarène, David Alleysson & Jeanny Hérault. *Practical Implementation of LMMSE Demosaicing Using Luminance and Chrominance Spaces*. Computer Vision and Image Understanding, Special Issue on Color Image Processing, In Press 2007. Cité page 107.

[de Monasterio 78] F. M. de Monasterio. *Properties of ganglion cells with atypical receptive-field organization in retina of macaques.* J Neurophysiol, vol. 41, no. 6, pages 1435–1449, Nov 1978. Cité page 45.

[De Valois 93] R. L. De Valois & K. K. De Valois. *A multi-stage color model.* Vision Res, vol. 33, no. 8, pages 1053–1065, May 1993. Cité pages 42, 49 et 56.

[DeVries 00] S. H. DeVries. *Bipolar cells use kainate and AMPA receptors to filter visual information into separate channels.* Neuron, vol. 28, no. 3, pages 847–856, Dec 2000. Cité page 31.

[Dippé 85] M.A.Z. Dippé & E.H. Wold. *Antialiasing through Stochastic Sampling*. In 12th annual conference on Computer graphics and interactive techniques, pages 69–78, 1985. Cité pages 83 et 85.

[Dubois 05] Eric Dubois. *Frequency-domain methods for demosaicking of Bayer-sampled color images*. IEEE Signal Processing Letters, vol. 12, pages 847–850, December 2005. Cité pages 65, 69 et 119.

[Ferreira 94] P.J. Ferreira. *Interpolation and the discrete papoulis-Gerchberg algorithm*. IEEE Transactions on Signal Processing, vol. 42, no. 10, pages 2596–2606, October 1994. Cité page 91.

[Fossum 93] E.R. Fossum. *Active Pixel Sensors - Are CCDs Dinosaurs ?* In SPIE, volume 1900, 1993. Cité page 50.

[Franze 07] Kristian Franze, Jens Grosche, Serguei N Skatchkov, Stefan Schinkinger, Christian Foja, Detlev Schild, Ortrud Uckermann, Kort Travis, Andreas Reichenbach & Jochen Guck. *Muller cells are living optical fibers in the vertebrate retina.* Proc Natl Acad Sci U S A, vol. 104, no. 20, pages 8287–8292, May 2007. Cité page 26.

[Glotzbach 01] J.W. Glotzbach, R.W. Schafer & K. Illgner. *A method of color filter array interpolation with alias cancellation properties*. In IEEE International Conference on Image Processing, pages 141–144, Thessaloniki, Greece, October 2001. Cité page 59.

[Goethe 10] Johann Wolfgang Goethe. Zur farbenlehre. 1810. Cité page 15.

[Gooley 03] Joshua J Gooley, Jun Lu, Dietmar Fischer & Clifford B Saper. *A broad role for melanopsin in nonvisual photoreception.* J Neurosci, vol. 23, no. 18, pages 7093–7106, Aug 2003. Cité page 34.

[Gopalan 01] A. Gopalan, P. Kulkarini & A.H. Titus. *An algorithm for the AVLSI implementation of the cephalopod retinaand visual system*. In International Joint Conference on Neural Networks, 2001. Cité page 30.

[Gunturk 02] Bahadir Gunturk, Yucel Altunbasak & Russell Mersereau. *Color plane interpolation using alternating projections*. IEEE Transactions on Image Processing, vol. 11, no. 9, pages 997–1013, September 2002. Cité pages 59, 67 et 69.

[Gunturk 05] B. K. Gunturk, J. Glotzbach, Y. Altunbazak, R. W. Schafer & R. M. Mersereau. *Demosaicking : Color filter array interpolation in single-chip digital cameras*. IEEE Signal Processing Mag., vol. 22, no. 1, pages 44–54, January 2005. Cité page 59.

[Hamilton 97] J. F. Hamilton & J.E. Adams. *Adaptive color plane interpolation in single sensor color electronic camera*. US Patent 5,629,734, to Eastman Kodak Company, May 1997. Cité page 67.

[Hecht 28] S. Hecht. *On the Binocular Fusion of Colors and Its Relation to Theories of Color Vision.* Proc Natl Acad Sci U S A, vol. 14, no. 3, pages 237–241, Mar 1928. Cité page 17.

[Hel-Or 05] Y. Hel-Or. *The Impulse Responses of Block Shift-Invariant Systems and their Use for Demosaicing Algorithms*. In IEEE International Conference on Image Processing, volume 2, pages 1006–1009, 2005. Cité page 108.

[Hering 74] E. Hering. Outline of a theory of the light sense. translated from the German by L. Hurvich and D. Jameson (1964), Harvard University, Cambridge, 1874. Cité page 15.

[Hibbard 95] R.H. Hibbard. *Apparatus and method for adaptively interpolating a full color image utilizing luminance gradients*. US patent 5,382,976, to Eastman Kodak Company, 1995. Cité pages 58 et 65.

[Hilz 70] R. Hilz & C. R. Cavonius. *Wavelength discrimination measured with square-wave gratings.* J Opt Soc Am, vol. 60, no. 2, pages 273–277, Feb 1970. Cité page 41.

[Hilz 74] R. L. Hilz, G. Huppmann & C. R. Cavonius. *Influence of luminance contrast on hue discrimination.* J Opt Soc Am, vol. 64, no. 6, pages 763–766, Jun 1974. Cité page 41.

[Hirakawa 05] Keigo Hirakawa & Thomas W. Parks. *Adaptive Homogeneity-Directed Demosaicing Algorithm*. IEEE Transactions on Image Processing, vol. 14, no. 3, pages 360–369, December 2005. Cité pages 67 et 69.

[Hofer 05] Heidi Hofer, Joseph Carroll, Jay Neitz, Maureen Neitz & David R Williams. *Organization of the human trichromatic cone mosaic*. J Neurosci, vol. 25, no. 42, pages 9669–9679, Oct 2005. Cité pages 29 et 30.

[Huang 65] Thomas S. Huang. *The subjective effect of two-dimensional pictorial noise*. IEEE Transactions on Information Theory, vol. 11, pages 43–53, 1965. Cité page 79.

[Hurvich 57] L. M. Hurvich & D. Jameson. *An opponent-process theory of color vision*. Psychol Rev, vol. 64, Part 1, no. 6, pages 384–404, Nov 1957. Cité page 17.

[Ichida 02] Jennifer M Ichida & Vivien A Casagrande. *Organization of the feedback pathway from striate cortex (V1) to the lateral geniculate nucleus (LGN) in the owl monkey (Aotus trivirgatus)*. J Comp Neurol, vol. 454, no. 3, pages 272–283, Dec 2002. Cité page 23.

[Ingling 83] C.R. Ingling & E. Martinez-Uriegas. The spatiochromatic signal of the r-g channel. Academic Press. London., 1983. Cité page 43.

[Jacobs 93] G. H. Jacobs. *The distribution and nature of colour vision among the mammals*. Biol Rev Camb Philos Soc, vol. 68, no. 3, pages 413–471, Aug 1993. Cité page 31.

[Jiménez 02] J. R. Jiménez, J. M. Medina, L. Jiménez del Barco & J. A. Díaz. *Binocular summation of chromatic changes as measured by visual reaction time*. Percept Psychophys, vol. 64, no. 1, pages 140–147, Jan 2002. Cité page 47.

[Johnson 01] E. N. Johnson, M. J. Hawken & R. Shapley. *The spatial transformation of color in the primary visual cortex of the macaque monkey*. Nat Neurosci, vol. 4, no. 4, pages 409–416, Apr 2001. Cité pages 36 et 146.

[Kelly 83] D. H. Kelly. *Spatiotemporal variation of chromatic and achromatic contrast thresholds*. J Opt Soc Am, vol. 73, no. 6, pages 742–750, Jun 1983. Cité page 40.

[Kimmel 99] R. Kimmel. *Demosaicing : Image Reconstruction from Color Samples*. IEEE Transactions on Image Processing, vol. 8, pages 1221–1228, 1999. Cité pages 58 et 68.

[Kitaoka 01] A. Kitaoka. *Illusion designology (7) : Illusory light perception that cannot be explained by the Fourier analysis*. Nikkei Science (in Japanese), vol. 31(8), pages 66–68, 2001. Cité page 153.

[Knutsson 93] H. Knutsson & C.-F. Westin. *Normalized and differential convolution*. In Computer Vision and Pattern Recognition, 1993. Proceedings CVPR '93., 1993 IEEE Computer Society Conference on, pages 515–523, 15-17 June 1993. Cité page 93.

[Kolb 97] H. Kolb, P. Goede, S. Roberts, R. McDermott & P. Gouras. *Uniqueness of the S-cone pedicle in the human retina and consequences for color processing.* J Comp Neurol, vol. 386, no. 3, pages 443–460, Sep 1997. Cité page 31.

[Kouyama 92] N. Kouyama & D. W. Marshak. *Bipolar cells specific for blue cones in the macaque retina.* J Neurosci, vol. 12, no. 4, pages 1233–1252, Apr 1992. Cité page 31.

[Lankheet 98] M. J. Lankheet, P. Lennie & J. Krauskopf. *Temporal-chromatic interactions in LGN P-cells.* Vis Neurosci, vol. 15, no. 1, pages 47–54, 1998. Cité page 33.

[Laroche 94] C.A. Laroche & M. A. Prescott. *Apparatus and method for adaptively interpolating a full color image utilizing chrominance gradients*. US patent 5,373,322, to Eastman Kodak Company, 1994. Cité pages 58 et 67.

[Le Floch 97] Hervé Le Floch. *Echantillonnage irregulier et interpolation : application a la representation d'images fixes et animees*. PhD thesis, Université de Rennes 1, 1997. Cité page 79.

[Lee 96] B. B. Lee. *Receptive field structure in the primate retina.* Vision Res, vol. 36, no. 5, pages 631–644, Mar 1996. Cité page 34.

[Lee 97] S. Lee, G. Wolberg & S. Y. Shin. *Scattered Data Interpolation with Multilevel B-Splines*. IEEE Transactions on Visualization and Computer Graphics, vol. 3, no. 3, pages 228–244, 1997. Cité page 99.

[Lee 07] Barry B. Lee. *Psychophysical models, physiology reality and the specificity of retinal signals.* In Symposium of the International Colour Vision Society, 2007. Cité page 21.

[Lennie 90] P. Lennie, J. Krauskopf & G. Sclar. *Chromatic mechanisms in striate cortex of macaque.* J Neurosci, vol. 10, no. 2, pages 649–669, Feb 1990. Cité page 37.

[Lennie 93] P. Lennie, J. Pokorny & V. C. Smith. *Luminance.* J Opt Soc Am A, vol. 10, no. 6, pages 1283–1293, Jun 1993. Cité pages 27 et 30.

[Li 05] X. Li. *Demosaicing by successive approximations*. IEEE Transactions on Image Processing, vol. 14, no. 3, pages 370–379, March 2005. Cité page 69.

[Lian 05] Naixiang Lian, Lanlan Chang & Yap-Peng Tan. *Improved color filter array demosaicking by accurate luminance estimation*. In IEEE International Conference on Image Processing, volume 1, pages 41–44, September 2005. Cité pages 65, 68, 69, 115 et 129.

[Livingstone 02] Margaret B. Livingstone. Vision and art : The biology of seeing. Harry N. Abrams Inc, 2002. Cité pages 145 et 154.

[Lukac 05a] R. Lukac & K. N. Plataniotis. *Color filter arrays : Design and performance analysis*. IEEE Transactions on Consumer Electronics, vol. 51, no. 4, pages 1260–1267, November 2005. Cité pages 117 et 138.

[Lukac 05b] Ratislav Lukac & Konstantinos N. Plataniotis. *Universal demosaicking for imaging pipelines with an RGB color filter array*. Pattern Recognition, vol. 38, pages 2208–2212, November 2005. Cité pages 53, 63 et 68.

[MacGill 02] Université MacGill. *Le cerveau à tous les niveaux, http ://lecerveau.mcgill.ca/*. Web site, copyleft, 2002. Cité pages 21 et 22.

[Magnan 03] Pierre Magnan. *Capteurs d'images CCD et CMOS : comparaison des technologies et perspectives*. Jautomatise, vol. 27, pages 66–75, 2003. Cité page 50.

[Malvar 04] H.S. Malvar, L.-W. He & R. Cutler. *High-Quality Linear Interpolation for Demosaicing of Bayer-Patterned Color Images*. In IEEE International Conference on Acoustics, Speech and Signal Processing, volume 3, pages 485–488, Montréal, Canada, May 2004. Cité pages 59 et 69.

[Mangel 87] S. C. Mangel & R. F. Miller. *Horizontal cells contribute to the receptive field surround of ganglion cells in the rabbit retina.* Brain Res, vol. 414, no. 1, pages 182–186, Jun 1987. Cité page 35.

[Marmorstein 07] Alan D Marmorstein & Lihua Y Marmorstein. *The challenge of modeling macular degeneration in mice.* Trends Genet, vol. 23, no. 5, pages 225–231, May 2007. Cité pages 27 et 28.

[Martin 01] P. R. Martin, B. B. Lee, A. J. White, S. G. Solomon & L. Rüttiger. *Chromatic sensitivity of ganglion cells in the peripheral primate retina.* Nature, vol. 410, no. 6831, pages 933–936, Apr 2001. Cité pages 34, 41 et 142.

[Martinez-Conde 06] Susana Martinez-Conde, Stephen L Macknik, Xoana G Troncoso & Thomas A Dyar. *Microsaccades counteract visual fading during fixation.* Neuron, vol. 49, no. 2, pages 297–305, Jan 2006. Cité page 35.

[Martinez-Uriegas 94] E. Martinez-Uriegas. Visual science and engineering, models and applications, chapitre Chromatic-Achromatic Multiplexing in Human Color Vision, pages 117–187. 1994. Cité page 43.

[Masland 01] R. H. Masland. *The fundamental plan of the retina.* Nat Neurosci, vol. 4, no. 9, pages 877–886, Sep 2001. Cité pages 23, 25, 31, 32 et 35.

[Medina 06] José M Medina. *Binocular interactions in random chromatic changes at isoluminance.* J Opt Soc Am A Opt Image Sci Vis, vol. 23, no. 2, pages 239–246, Feb 2006. Cité page 46.

[Meylan 07] Laurence Meylan, David Alleysson & Sabine Süsstrunk. *Model of retinal local adaption for the tone mapping of color filter array images*. Journal of Optical Society of America, 2007. Cité pages 35 et 140.

[Mullen 91] K. T. Mullen. *Colour vision as a post-receptoral specialization of the central visual field.* Vision Res, vol. 31, no. 1, pages 119–130, 1991. Cité page 34.

[Mullen 96] K. T. Mullen & F. A. Kingdom. *Losses in peripheral colour sensitivity predicted from "hit and miss" post-receptoral cone connections.* Vision Res, vol. 36, no. 13, pages 1995–2000, Jul 1996. Cité pages 34 et 142.

[Okano 92] T. Okano, D. Kojima, Y. Fukada, Y. Shichida & T. Yoshizawa. *Primary structures of chicken cone visual pigments : vertebrate rhodopsins have evolved out of cone visual pigments.* Proc Natl Acad Sci U S A, vol. 89, no. 13, pages 5932–5936, Jul 1992. Cité pages 26 et 27.

[Oliva 06] A. Oliva, A. Torralba & P.G. Schyns. *Hybrid Images*. ACM Transactions on Graphics, vol. 25-3, pages 527–532, 2006. Cité page 79.

[Packer 96] O. S. Packer, D. R. Williams & D. G. Bensinger. *Photopigment transmittance imaging of the primate photoreceptor mosaic.* J Neurosci, vol. 16, no. 7, pages 2251–2260, Apr 1996. Cité page 29.

[Parker 99] K.J. Parker, Q. Yu & M. Yao. *On filter techniques for generating blue noise mask*. In Recent Progress in Digital Halftoning II, pages 208–212, march 1999. Cité page 102.

[Pei 00] S.C. Pei & I.K. Tam. *Effective Color Interpolation in CCD Color Filter Array Using Signal Correlation*. In IEEE International Conference on Image Processing, 2000. Cité page 58.

[Pei 03] S.-C. Pei & I.-K. Tam. *Effective Color Interpolation in CCD Color Filter Arrays Using Signal Correlation*. IEEE Transactions on Circuits and Systems for Video Technology, vol. 13, no. 6, June 2003. Cité pages 58 et 69.

[Perry 84] V. H. Perry, R. Oehler & A. Cowey. *Retinal ganglion cells that project to the dorsal lateral geniculate nucleus in the macaque monkey.* Neuroscience, vol. 12, no. 4, pages 1101–1123, Aug 1984. Cité pages 32 et 34.

[Pinna 01] B. Pinna, G. Brelstaff & L. Spillmann. *Surface color from boundaries : a new 'watercolor' illusion*. Vision Research, vol. 41, pages 2669–2676, 2001. Cité page 41.

[Portilla 05] J. Portilla, D. Otaduy & C. Dorronsoro. *Low-Complexity Linear Demosaicing Using Joint Spatial-Chromatic Image Statistics*. In IEEE International Conference on Image Processing, volume 1, pages 61–64, Genoa, Italy, September 2005. Cité pages 108 et 111.

[Reid 92] R. C. Reid & R. M. Shapley. *Spatial structure of cone inputs to receptive fields in primate lateral geniculate nucleus.* Nature, vol. 356, no. 6371, pages 716–718, Apr 1992. Cité page 45.

[Rodieck 91] R. Rodieck. Which cells code for color ?, in from pigments to perception. Plenum Press, 1991. Cité page 45.

[Roorda 01] A. Roorda, A. B. Metha, P. Lennie & D. R. Williams. *Packing arrangement of the three cone classes in primate retina.* Vision Res, vol. 41, no. 10-11, pages 1291–1306, 2001. Cité pages 29 et 30.

[Schwartz 80] E. L. Schwartz. *Computational anatomy and functional architecture of striate cortex : a spatial mapping approach to perceptual coding.* Vision Res, vol. 20, no. 8, pages 645–669, 1980. Cité page 37.

[Smith 92] V. C. Smith, B. B. Lee, J. Pokorny, P. R. Martin & A. Valberg. *Responses of macaque ganglion cells to the relative phase of heterochromatically modulated lights.* J Physiol, vol. 458, pages 191–221, Dec 1992. Cité pages 32, 33, 34 et 146.

[Solomon 07] Samuel G Solomon & Peter Lennie. *The machinery of colour vision.* Nat Rev Neurosci, vol. 8, no. 4, pages 276–286, Apr 2007. Cité page 27.

[Stockman 00] A. Stockman & L. T. Sharpe. *The spectral sensitivities of the middle- and long-wavelength-sensitive cones derived from measurements in observers of known genotype.* Vision Res, vol. 40, no. 13, pages 1711–1737, 2000. Cité page 27.

[Sun 06] Hao Sun, Hannah E. Smithson, Qasim Zaidi & Barry B. Lee. *Do magnocellular and parvocellular ganglion cells avoid short-wavelength cone input ?* Vis Neurosci, vol. 23, no. 3-4, pages 441–446, 2006. Cité page 34.

[Taubman 00] David Taubman. *Generalized Wiener Reconstruction Of Images From Colour Sensor Data Using A Scale Invariant Prior*. In IEEE International Conference on Image Processing, volume 3, pages 801–804, 2000. Cité pages 108 et 109.

[Thibos 87] L. N. Thibos, D. J. Walsh & F. E. Cheney. *Vision beyond the resolution limit : aliasing in the periphery.* Vision Res, vol. 27, no. 12, pages 2193–2197, 1987. Cité page 38.

[Thibos 91] Larry Thibos & Arthur Bradley. *The Limits to Performance in Central and Peripheral Vision*. Society for Information Display, Digest of Technical Papers, vol. 22, pages 301–303, 1991. Cité page 39.

[Thibos 00] Larry Thibos. Handbook of perception and cognition, chapitre Formation and Sampling of the Retinal Image, pages 1–52. Academic Press, 2000. Cité pages 22 et 39.

[Thorell 84] L. G. Thorell, R. L. De Valois & D. G. Albrecht. *Spatial mapping of monkey V1 cells with pure color and luminance stimuli.* Vision Res, vol. 24, no. 7, pages 751–769, 1984. Cité page 37.

[Trussell 02] H. J. Trussell & R. E. Hartwig. *Mathematics For Demosaicking*. IEEE Transactions on Image Processing, vol. 11, no. 4, April 2002. Cité pages 108, 109 et 110.

[Unser 99] M. Unser. *Splines : A Perfect Fit for Signal and Image Processing*. IEEE Signal Processing Magazine, vol. 16, no. 6, pages 22–38, November 1999. IEEE Signal Processing Society's 2000 magazine award. Cité pages 75 et 98.

[von Helmholtz 67] Hermann von Helmholtz. Handbuch der physiologischen optik. Leipzig, Voss, 1867. Cité page 14.

[Wang 97] Y. Z. Wang, A. Bradley & L. N. Thibos. *Aliased frequencies enable the discrimination of compound gratings in peripheral vision.* Vision Res, vol. 37, no. 3, pages 283–290, Feb 1997. Cité page 40.

[Wernet 06] Mathias F Wernet, Esteban O Mazzoni, Arzu Celik, Dianne M Duncan, Ian Duncan & Claude Desplan. *Stochastic spineless expression creates the retinal mosaic for colour vision.* Nature, vol. 440, no. 7081, pages 174–180, Mar 2006. Cité page 147.

[Wiesel 59] T. N. Wiesel. *Recording inhibition and excitation in the cat's retinal ganglion cells with intracellular electrodes.* Nature, vol. 183, no. 4656, pages 264–265, Jan 1959. Cité page 33.

[Williams 91] D. Williams, N. Sekiguchi, W. Haake, D. Brainard & O. Packer. Advances in understanding visualprocesses : Convergence of neurophysiological and psychophysical evidence, chapitre The cost of trichromacy for spatial vision. Plenum Press, 1991. Cité page 40.

[Williams 07] David Williams. *Imaging Retinal Mosaics in the Living Eye*. In 19th Symposium of the International Colour Vision Society, 2007. Cité page 29.

[Yellott 90] John I. Yellott. *The Photoreceptor Mosaic as an Image Sampling Device*. In Advances in Photoreception : Proceedings of a Symposium on Frontiers of Visual Science, pages 117–134, 1990. Cité pages 28 et 37.

[Zhu 98] W. A. Zhu, K. J. Parker & M. A. Kriss. *Color filter arrays based on mutually exclusive blue noise patterns*. In G. B. Beretta & R. Eschbach, editeurs, Proc. SPIE Vol. 3300, p. 207-218, Color Imaging : Device-Independent Color, Color Hardcopy, and Graphic Arts III, Giordano B. Beretta ; Reiner Eschbach ; Eds., pages 207–218, January 1998. Cité page 99.

L'échantillonnage spatio-chromatique dans la rétine humaine et les caméras numériques

Résumé :

Il est aujourd'hui établi que la vision des couleurs chez l'homme est permise par la présence de trois types de photorécepteurs, sensibles respectivement aux courtes, moyennes et grandes longueurs d'onde de la lumière. La présence de ces trois classes d'échantillonneurs induit un espace perceptif des couleurs tridimensionnel. Par ailleurs, les expériences psycho-physiques montrent que le système visuel humain construit trois voies perceptives à partir de l'image échantillonnée au niveau des cônes. L'une contient l'information achromatique, appelée luminance, et les deux autres véhiculent les informations d'oppositions de couleurs rouge moins vert et bleu moins jaune.

Cependant, le système visuel est confronté à un problème de taille puisqu'à chaque position spatiale de la rétine, un seul cône est présent. Chaque cône mesure, par conséquent, à la fois l'information de luminance et l'information de couleur. Le mécanisme permettant le passage de cette information rétinienne ambiguë à trois voies perceptives distinctes est encore mal compris, en particulier dans le cas d'un arrangement irrégulier des cônes, tel que c'est le cas dans la rétine.

Nous proposons un modèle d'estimation de la luminance et des oppositions de couleurs s'appliquant quel que soit l'arrangement spatial des photorécepteurs, et discutons de ses implications pour le traitement rétinien de la couleur. Ce modèle s'applique également au domaine des caméras numériques dites "mono-capteur", ces dernières étant confrontées à la même problématique d'échantillonnage spatio-chromatique que la rétine.

Abstract :

It is now well established that the human color vision is allowed by the presence of three types of photoreceptors, sensitive respectively to the short, middle and long wavelengths of light. The presence of these three classes of samplers inducts a perceptive colour space of three dimensions. It is for instance possible to repoduce any colour using three basic colours, called primary colours. Besides, psycho-physics experiments show that the human visual system builds three perceptive channels from the retinal image sampled by the cones. One conveys the achromatic information, called luminance, and the two others convey the colour oppositions red minus green and blue minus yellow.

However, the visual system faces a dramatic problem since at each spatial position in the retina, only one sort of cone is present. Consequently, each cone measures both luminance and color information. The mecanism transforming this ambiguous retinal image into three disctinct channels is still not clear, especially when considering the random arrangement of the cones.

We propose a model of estimation of the luminance and colour opposition channels. This models applies to any spatial arrangement of the photoreceptors. We also discuss its implications for the retinal processing of colour. This model also applies, in an effective manner, to single-sensor cameras, which are confronted to the same spatio-chromatic sampling issue.