

Traitement non linéaire uniforme pour la reconstruction d'une mosaïque chromatique répartie aléatoirement

Brice CHAIX DE LAVARENÈ¹, David ALLEYSSON², Jeanny HÉRAULT¹

¹Laboratoire des Images et des Signaux
46 Avenue Félix Viallet, 38031 Grenoble Cedex, France

²Laboratoire de Psychologie et NeuroCognition
Université Pierre Mendès France, BP 47, 38040 Grenoble Cedex 9, France

brice.chaix@lis.inpg.fr,
david.alleysson@upmf-grenoble.fr,
jeanny.herault@lis.inpg.fr

Résumé – Dans la mosaïque des cônes de la rétine, les informations spatiale et chromatique sont mélangées, c'est-à-dire que la carte achromatique (définie comme l'information spatiale sans contenu chromatique) est confondue avec la carte chromatique (définie comme la distribution de l'information chromatique à luminance constante). Nous avons précédemment montré [1] que dans le cas d'un arrangement régulier des photorécepteurs dans la mosaïque - tel est le cas des appareils photographiques numériques -, un algorithme linéaire est capable d'extraire l'information spatiale achromatique et d'interpoler l'information chromatique. En effet, la régularité de la disposition des photorécepteurs permet la localisation dans le plan de Fourier de la luminance et des signaux d'opposition de couleur, ce qui en retour permet leur séparation par un filtrage linéaire spatialement invariant. Cependant dans le cas d'un arrangement irrégulier, tel qu'on le trouve dans la rétine, la luminance et la chrominance sont également mélangées dans le plan de Fourier, et un algorithme non linéaire est alors nécessaire. Quel algorithme est le meilleur? Nous montrons qu'un filtrage linéaire uniforme suivi d'une normalisation non linéaire de chacun des canaux est capable d'estimer précisément l'information achromatique de la mosaïque des cônes.

Abstract – In the cone mosaic, spatial and chromatic information is mixed, i.e. the achromatic map (defined as the spatial information without chromatic content) is confounded with the chromatic map (distribution of chromatic information at constant luminance). We have previously shown [1] that in case of a regular arrangement of chromatic samples in the mosaic, a linear algorithm is able to extract achromatic spatial information and to interpolate chromatic information. Indeed, the regularity of chromatic samples permits the localization of luminance and chromatic-opponent signals in the Fourier spectrum which in turn permits their separation using uniform linear filtering. However, in the case of an irregular arrangement, luminance and chrominance are no longer localized and a non linear algorithm is needed. Which non linear algorithm is best? We show that a uniform linear filtering followed by a non linear normalisation of each channel is able to estimate accurately the achromatic information in the cone mosaic.

1 Introduction

L'acquisition des images dans la rétine humaine et dans la plupart des caméras numériques fait intervenir une représentation spatio-chromatique réduite pour laquelle une seule composante chromatique est définie à chaque position spatiale (Figure 1). Dans la rétine humaine trois types de cônes sensibles à trois gammes de longueurs d'onde, tapissent la rétine et il ne peut y avoir plusieurs cônes au même endroit. Dans les caméras numériques avec un seul capteur on applique généralement une matrice de filtres couleurs devant la caméra, permettant d'allouer une sensibilité chromatique particulière à chacune des cellules photosensible. On dit que cette image est une mosaïque (on parle aussi de la mosaïque des cônes dans la rétine) puisque toute la variété chromatique n'est pas accessible à chaque position spatiale.

Cette représentation implique un mélange de l'information spatiale et chromatique de l'image échantillonnée. En effet, si l'on constate une variation de la réponse d'un photorécepteur, on ne peut pas dire si elle résulte d'une variation de l'intensité lumineuse de la source de lumière ou d'une variation de ses

composantes chromatiques. Dans une image couleur normale définie avec trois composantes chromatiques par position spatiale on peut déterminer l'origine de la variation. Si la moyenne (ou moyenne pondérée) des trois composantes chromatiques varie on dit que c'est l'intensité lumineuse qui a changé et dans les autres cas on dit que c'est la variété chromatique. Dans une image avec une seule composante chromatique par position spatiale, il n'est pas possible de calculer la moyenne des variations chromatique en un seul point, on utilise les échantillons voisins pour réaliser cette estimation. On peut alors penser que cette estimation sera effectuée avec une résolution réduite car sur une portion plus grande de l'espace. De plus, si l'on considère chacun des canaux chromatiques séparément, la limite théorique de fréquence spatiale qui permet cette estimation sans erreur (i.e. sans repliement spectral) est celle fixée par le théorème d'échantillonnage et correspond à la moitié de la fréquence d'échantillonnage de chacun des canaux chromatiques.

En fait, on peut montrer que cette limite est franchissable du fait de la corrélation qui existe entre les trois canaux de couleur. La représentation des images en mosaïque peut se décomposer

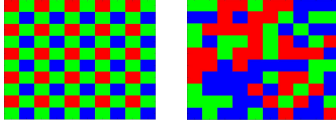


FIG. 1 – Mosaïques régulière (de Bayer) et irrégulière.

en une somme de l'information spatiale achromatique - la luminance - et de l'information spatiale chromatique - la chrominance [2]. De plus, dans le cas d'une répartition régulière (périodique) des composantes chromatiques, l'information de chrominance est localisée dans le spectre de Fourier permettant une estimation de la luminance et de la chrominance par filtrage linéaire uniforme (Figure 2) [1]. Grâce à la corrélation entre les canaux, la largeur de bande de la chrominance est inférieure à la largeur de bande de la luminance, permettant une représentation de l'information spatiale plus grande que celle imposée par le théorème d'échantillonnage définie sur chacun des canaux. On peut également trouver les conditions de fréquences spatiales de l'information achromatique et de la corrélation entre les canaux qui optimise la représentation.

En revanche, dans le cas où la répartition n'est pas périodique, la chrominance n'est pas localisée mais diluée dans le spectre de Fourier (Figure 2). Dans ce cas l'information spatiale et chromatique ne sont pas à supports disjoints dans le spectre de Fourier mais se recouvrent mutuellement. Un filtrage linéaire uniforme ne peut s'appliquer dans ce cas, puisque ce dernier est seulement capable de sélectionner des bandes de fréquences spatiales, mais pas de séparer des informations qui occupent les mêmes composantes de fréquences. Un filtrage non linéaire et/ou non uniforme est nécessaire. Dans cet article nous montrons que l'utilisation d'une non linéarité après un filtrage uniforme permet d'estimer de manière efficace l'information spatiale.

2 Informations spatiale et chromatique

2.1 Modèle d'échantillonnage

Une image couleur $\vec{T}(x, y)$ est une matrice de vecteurs à trois dimensions dans l'espace couleur correspondant. On peut donc écrire :

$$\vec{T}(x, y) = \{C_i(x, y)\}, \text{ avec } i \in \{R, G, B\} \quad (1)$$

$C_i(x, y)$ représente chacun des canaux couleurs. Dans tout ce qui suit, l'indice i représentera un canal chromatique (R, G ou B). Cette forme de représentation ne permet pas d'isoler directement les propriétés spatiales et chromatiques de la scène originale. En effet, une variation de $C_i(x, y)$ peut-être due à une variation d'intensité lumineuse ou une variation de la diversité chromatique de la scène. D'une manière générale, la reconstruction de l'information lumineuse est déterminée par l'estimation de la moyenne entre les canaux couleurs, supposant que la variation en phase des informations chromatiques est due à une augmentation de l'intensité lumineuse, alors qu'une variation en contre phase correspond à une variation de teinte. Parfois l'estimation de la moyenne est pondérée, permettant de donner une importance proportionnellement plus grande à une des gammes de longueurs d'onde comparée aux autres. Nous

écrivons donc la luminance comme :

$$\phi(x, y) = \sum_i \alpha_i C_i(x, y) \quad (2)$$

avec $\alpha_i \in [0, 1]$ la pondération de chacun des canaux de couleur à la contribution de la luminance et tels que leur somme soit égale à 1. Nous remarquons que si la moyenne pondérée ϕ ne change pas, alors que les valeurs individuelles des canaux sont modifiées, c'est une modification de teinte qui a eu lieu. De manière complémentaire, on définit la chrominance $\vec{\psi}(x, y)$ comme la différence entre l'image originale et la luminance. Il suit par définition que :

$$\vec{\psi}(x, y) = \underbrace{\{C_i(x, y) - \phi(x, y)\}}_{\psi_i(x, y)} \quad (3)$$

soit

$$\begin{aligned} \psi_i(x, y) &= C_i(x, y) - \sum_i \alpha_i C_i(x, y) \\ &= (1 - \alpha_i)C_i(x, y) - \sum_{(j \neq i)} \alpha_j C_j(x, y) \end{aligned} \quad (4)$$

Puisque l'image couleur est un vecteur à trois dimensions et que la luminance est un scalaire à une seule dimension, la chrominance est définie comme un vecteur à trois dimensions¹. Chacune des composantes $\psi_i(x, y)$ de ce vecteur s'exprime par une différence entre les canaux chromatiques, donc une variation conjointe des canaux de chromatique ne modifiera pas le signal de chrominance. La chrominance porte donc l'information de variation chromatique indépendamment de l'intensité lumineuse de la lumière incidente. Dans le cas d'une image avec une seule composante chromatique par position spatiale, cette décomposition ne peut pas avoir lieu car elle fait intervenir trois composantes pour l'estimation de la luminance et de la chrominance à chacune des positions. L'utilisation des valeurs voisines est donc indispensable. Mais, la représentation en luminance et chrominance a-t-elle un sens lorsque la répartition des valeurs chromatiques est aléatoire ? En fait, on peut écrire une image avec une seule composante chromatique par position spatiale comme une image couleur dans laquelle les informations chromatiques sont sous échantillonnées [1] :

$$\vec{T}_{mo}(x, y) = C_i(x, y)m_i(x, y) \quad (5)$$

Avec $m_i(x, y)$ les fonctions de sous échantillonnage aléatoire de chacun des canaux chromatiques. Ces fonctions sont mutuellement exclusives et prennent les valeurs 1 ou 0 suivant que la position (x, y) contient le photorécepteur i ou non. On peut également écrire cette image comme un scalaire :

$$I_{Mo}(x, y) = \sum_i C_i(x, y)m_i(x, y) \quad (6)$$

Les fonctions de modulation peuvent se décomposer en une partie constante qui correspond à la proportion de photorécepteur de type i présente dans la mosaïque et une partie variable à moyenne nulle de la manière suivante :

$$m_i(x, y) = p_i + \tilde{m}_i(x, y) \quad (7)$$

¹en fait deux dimensions d'un point de vue algébrique, puisque (2) et (3) impliquent que $\sum_i \alpha_i \psi_i(x, y) = 0$ pour toute position spatiale (x, y) .

Cette décomposition des fonctions d'échantillonnage se retrouve au niveau de l'image en mosaïquée :

$$I_{Mo}(x, y) = \underbrace{\sum_i p_i C_i(x, y)}_{\phi(x, y)} + \underbrace{\sum_i C_i \tilde{m}_i(x, y)}_{\psi_{mod}(x, y)} \quad (8)$$

Apparaît alors un terme de luminance cohérent avec la définition donnée en (2) pour une image couleur pleinement échantillonnée, avec $\alpha_i = p_i$ la proportion de chacun des photorécepteurs dans la mosaïque. Cela signifie que la luminance n'est pas affectée par le sous-échantillonnage induit par la mosaïque. En revanche le terme correspondant à la chrominance se voit modulé par les fonctions de modulation \tilde{m}_i . Même s'il existe de nombreuses façons de décomposer les fonctions de modulation $m_i(x, y)$ en somme de deux éléments, celle qui est proposée ici (équations (7),(8)) garantie que l'information de luminance est achromatique et que l'information de chrominance est à luminance constante, soit sans composante d'intensité lumineuse.

2.2 Point de vue spectral

De plus cette décomposition a une conséquence sur la représentation en fréquences spatiales des informations de luminance et de chrominance. Puisque la composante de luminance résulte de l'estimateur p_i , elle ne dépend pas de la modulation. Par conséquent, sa représentation en fréquence n'est pas modifiée. Par contre, la chrominance a une valeur moyenne nulle car $\sum_{x,y} \tilde{m}_i(x, y) = 0$ et se retrouve modulée par les fonctions m_i .

Dans le cas où la mosaïque est régulière, les transformées de Fourier des fonctions de modulation peuvent être explicitement calculées [1] et correspondent à des impulsions de dirac situées sur les milieux des bords et dans les coins du plan de Fourier. La chrominance est par conséquent modulée en haute fréquence (Fig.2). Ainsi la fréquence limite imposée par le théorème d'échantillonnage pour la luminance n'est pas $f_e - f_e/2$ mais $(f_e - B_{chr}/2)$, avec f_e la fréquence d'échantillonnage d'un canal chromatique isolé et B_{chr} la largeur de bande de la chrominance. Comme la chrominance possède la propriété d'être à bande étroite dans les images naturelles, une plus grande bande passante est allouée à la luminance. Ainsi l'algorithme d'interpolation est très efficace tant du point de vue de la qualité de l'interpolation (acuité spatiale sous-réserve de non repliement spectral luminance/chrominance) que du point de vue de la complexité algorithmique (linéarité).

En revanche dans le cas où la mosaïque est irrégulière, l'information de chrominance se retrouve modulée non pas par des porteuses hautes fréquences, mais par des porteuses aléatoires. On peut montrer, en passant par le calcul de l'autocorrélation, que ces porteuses possèdent une Densité Spectrale de Puissance moyenne (DSPm) répartie sur l'ensemble du spectre, et dont l'allure globale dépend du type d'aléatoire utilisé dans la construction de la mosaïque. Par exemple dans le cas où les échantillons sont répartis en suivant une loi de probabilité uniforme sur la surface de l'image, la DSPm possède un spectre plat, celui d'un bruit blanc. Cela signifie que la chrominance modulée va se répartir sur l'ensemble du plan de Fourier. On peut en fait maîtriser l'allure spectrale des fonctions de modulation en apportant des contraintes sur la position des photorécepteurs. Par exemple, insérer une contrainte de distance minimale

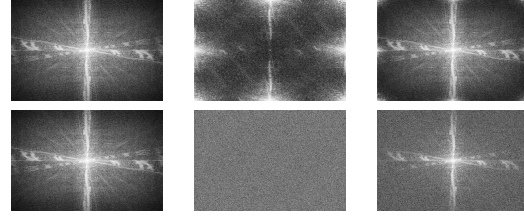


FIG. 2 – Spectres d'amplitude de la luminance, de la chrominance modulée et de l'image en mosaïque. La 1^{re} ligne correspond au cas de la mosaïque régulière, la 2^{de} à celui de la mosaïque irrégulière. Dans les deux cas on constate que la mosaïque correspond à la somme de la luminance et de la chrominance modulée.

entre deux récepteurs permet de creuser le spectre au niveau des basses fréquences. Yellott a montré que c'est ce type d'échantillonnage qui est présent dans la rétine [3]. On peut légitimement se demander pour quelle(s) raison(s) il serait intéressant d'utiliser une mosaïque irrégulière plutôt qu'une mosaïque régulière, étant donné que d'un point de vue computationnel les algorithmes linéaires sont plus efficaces. En fait la différence réside au niveau des artefacts que produisent ces deux types d'échantillonnage. Dans le cas d'une mosaïque régulière, les artefacts sont causés par le repliement spectral entre la chrominance et la luminance. Leur énergie est donc localisée en un point précis du spectre, se manifestant par des structures très régulières auxquelles le système visuel humain est sensible. Au contraire dans le cas d'une mosaïque irrégulière les artefacts de chrominance auront une énergie étalée et se retrouveront sous forme de «bruit», c'est-à-dire une information non cohérente et moins perceptible. La luminance n'est ainsi plus affectée par le problème de repliement spectral avec la chrominance, mais est affectée par un bruit de chrominance.

3 Reconstruction de l'image couleur

3.1 Principe

Dès lors que la luminance et la chrominance occupent la même région dans le plan de Fourier, on ne peut plus sélectionner la luminance par filtrage passe-bas uniforme. Remarquons que la première approximation de l'équation (7) n'est valable que si l'on considère l'image dans sa globalité. En effet si on calcule les densités p_i localement (c'est-à-dire sur un voisinage autour de chaque pixel), alors les p_i ne sont pas constants sur l'ensemble de l'image mais ils fluctuent autour de p_i , du fait du caractère aléatoire. Si bien que la luminance estimée par filtrage passe-bas uniforme est :

$$\tilde{\phi}(x, y) = \sum_i p_i^{loc}(x, y) C_i(x, y) \quad (9)$$

avec p_i^{loc} dépendant de la position spatiale, de moyenne p_i . Mais dans notre modèle la définition de la luminance est basé sur une contribution constante des canaux chromatiques sur l'ensemble de l'image (équation (8)), cela signifie que $\tilde{\phi}(x, y)$ contient de l'information de chrominance. Ainsi pour retrouver une luminance cohérente avec notre modèle, c'est-à-dire dont la contribution des canaux chromatiques est indépendante de la position, une normalisation locale de la contribution des diffé-



FIG. 3 – Exemple d’images reconstruites à partir de la grille régulière de Bayer (à gauche) et d’une grille irrégulière (à droite). On relève de types d’artefacts différents.

rents canaux est nécessaire :

$$\phi(x, y) = \sum_i \frac{p_i}{p_i^{loc}(x, y)} (p_i^{loc}(x, y) C_i(x, y)) \quad (10)$$

Cette normalisation correspond au rapport entre p_i la contribution du canal i à l’estimation de la luminance prévue par le modèle et $p_i^{loc}(x, y)$ la contribution locale du canal i à la position (x, y) . La contribution locale est estimée par filtrage passe-bas de la grille d’échantillonnage comportant uniquement les récepteurs de type i . Il faut bien entendu veiller à choisir un filtre dont la fréquence de coupure est suffisamment basse pour que la contribution locale ne soit pas nulle.

Une fois la luminance estimée, celle-ci est soustraite à la mosaïque afin d’obtenir la chrominance sous-échantillonnée et spatialement multiplexée. La chrominance étant à bande étroite dans les images naturelles, elle est facilement interpolée en utilisant de nouveau une convolution normalisée, c’est-à-dire un filtrage passe-bas normalisé localement par la densité en échantillons. La figure 3 montre un exemple de reconstruction utilisant cette méthode.

Il est intéressant de noter qu’il existe une analogie dans l’estimation de la luminance entre le cas d’une répartition régulière et celui d’une répartition irrégulière des récepteurs. Dans le premier cas, l’algorithme par sélection de fréquence utilise un filtrage passe-bas, technique utilisée pour interpoler des signaux échantillonnés par un pas régulier. Dans le second cas, l’estimation de la luminance est une sorte de «convolution normalisée» [4], technique adaptée quant à elle à l’interpolation de signaux échantillonnés aléatoirement.

3.2 Comparaison

En pratique, nous limitons notre étude au cas où la mosaïque est une grille régulière, mais sur laquelle les photorécepteurs sont disposés aléatoirement, avec les probabilités p_i associées. Les tests ont été effectués pour deux types de grilles : une où les choix des récepteurs sont indépendants des voisins (bruit blanc), et une où une contrainte de distance minimale entre deux récepteurs de même type est imposée (pseudo bruit bleu).

Il existe dans la littérature à notre connaissance un seul article [5] faisant état d’un algorithme d’interpolation spécifiquement adapté à une mosaïque chromatique irrégulière. Cet algorithme est basé sur une détection de contour qui permet d’adapter l’interpolation suivant que l’on se trouve sur un contour ou non. Il existe par ailleurs de nombreux algorithmes d’interpolation de signaux échantillonnés aléatoirement. Nous avons

également choisi de comparer la présente méthode avec l’interpolation disjointe de chaque canal chromatique par convolution normalisée [4] ainsi que par l’algorithme de Papoulis-Gerchberg [6] adapté à la dimension deux. Ce dernier étant un algorithme itératif de reconstruction de signaux à bande limitée, de haute complexité en terme computationnel. Les résultats moyennés sur une base de 23 images sont consignés dans le tableau 1. Il est à noter que dans le cas de l’algorithme de Papoulis-Gerchberg, le choix des paramètres optimaux en terme de PSNR ne correspond pas à celui en terme de qualité visuelle. Nous avons retenu le meilleur PSNR obtenu.

TAB. 1 – Comparaison des méthodes

| PSNR (dB) | [4] | [5] | [6] | Proposée |
|---------------------|------|------|-------|----------|
| bruit blanc | 27,9 | 28,2 | 27,84 | 30,7 |
| (pseudo) bruit bleu | 28,9 | 29,0 | 29,75 | 32,4 |

4 Conclusion

Ce travail fait apparaître que la reconstruction d’image définie par une mosaïque d’échantillons chromatiques arrangés aléatoirement peut se faire en utilisant une normalisation dépendant du nombre d’échantillon de chacune des composantes chromatiques dans un voisinage local après avoir fait un filtrage uniforme sur chaque canal chromatique. C’est-à-dire que cette reconstruction ne nécessite pas forcément un algorithme non linéaire et non uniforme complexe. Ce résultat est par ailleurs compatible en première approximation avec les fonctions de filtrage (connexions latérales fixes entre neurones voisins dans la rétine) et de normalisation (loi non-linéaire adaptative dans la rétine).

Références

- [1] D. Alleysson, S. Süsstrunk et J. Héroult. *Linear color demosaicing inspired by the human visual system*, IEEE Transactions on Image Processing, avril 2005.
- [2] D. Alleysson et J. Héroult. *Interpolation d’images couleurs sous échantillonnées par un modèle de perception*, GRETSI’2001, Toulouse, 2001.
- [3] J.I. Yellott Jr. *Spectral analysis of spatial sampling by photoreceptors : topological disorder prevents aliasing*, Vision Research, 22, pp.1205-1210, 1982.
- [4] F.A. Marvasti, C. Liu et G. Adams. *Analysis and recovery of multidimensional signals from irregular samples using nonlinear and iterative techniques*, Signal Processing, 36 :13–30, 1994.
- [5] W. Zhu, K. Parker et M.A. Kriss. *Color Filter Arrays Based on Mutually Exclusive Blue Noise Patterns*, Journal of Visual Communication and Image Representation, vol.10, pp.245-267, 1999.
- [6] P.J. Ferreira/ *Interpolation and the Discrete Papoulis-Gerchberg Algorithm* IEEE Transactions on Signal Processing, vol.42, n.10, octobre 1994