

Composantes principales spatio-chromatiques des images naturelles couleur

Edoardo PROVENZI¹, Julie DELON¹, Yann GOUSSEAU², Baptiste MAZIN²

¹Université Paris Descartes, Sorbonne Paris Cité, Laboratoire MAP5,
UMR CNRS 8145, 45 rue des Saints Pères, 75006, Paris, France

²Télécom ParisTech, LTCI, CNRS, 46, rue Barrault, 75013, Paris, France

Résumé – Dans cet article, nous montrons que deux hypothèses très simples permettent d’expliquer pourquoi les matrices de covariance des images en couleur se décomposent dans une base correspondant au produit tensoriel entre une base de Fourier et des canaux couleur antagonistes. La première hypothèse est la stationnarité au second ordre, la deuxième est la commutativité entre les matrices de corrélation couleur. Nous étudions la validité de ces hypothèses sur deux grandes bases de données d’images.

Abstract – In this contribution we show that two basic assumptions on the covariance matrices of color images are enough to explain the appearance of spatiochromatic features given by Fourier descriptors in the luminance plus color opponent channels. The first of these assumptions is second order stationarity while the second one is commutativity between color correlation matrices. The validity of these assumptions are experimentally studied on two large image databases.

1 Introduction

Le système visuel humain (SVH) a évolué en direction d’une optimisation du codage et de la transmission des signaux visuels, selon un processus complexe et encore mal connu qui repose sur les particularités statistiques des signaux captés par le SVH à partir des scènes naturelles. Dans ce papier, nous nous intéressons à un aspect relativement simple de la structure statistique de ces signaux, à savoir les relations de dépendance spatio-chromatiques du second ordre. Deux types de redondances sont facilement identifiées. Tout d’abord, les points voisins d’une scène ont des valeurs de luminance semblables, ce qui implique une forte *corrélacion spatiale*. Deuxièmement, les signaux lumineux sont captés par les cônes de type L , M et S , dont les sensibilités ne sont pas indépendantes (leurs courbes de sensibilité spectrale se superposent largement), ce qui implique une forte *corrélacion chromatique*. Les deux effets ont lieu conjointement, ce qui conduit à une *corrélacion spatio-chromatique*.

Parmi les résultats les plus simples sur les statistiques du second ordre des images numériques en niveau de gris, nous pouvons citer la stationnarité spatiale. Une conséquence de ce résultat est qu’une analyse en composantes indépendantes mène à des structures de type base de Fourier. Un autre résultat très classique est que le spectre des images décroît selon une loi en puissance, en lien avec les nombreuses invariances aux changements d’échelle observées sur les images naturelles. Les statistiques d’ordre supérieur ont bien entendu été largement étudiées, par exemple par l’intermédiaire de décompositions en ondelettes ou parcimonieuses [7]. En ce qui concerne les redon-

dance chromatiques, plusieurs travaux ont établi des liens entre les statistiques du second ordre et des espaces de couleurs antagonistes [11, 2].

L’une des observations les plus surprenantes dans ce domaine est que les matrices de covariance spatio-chromatiques ressemblent au produit tensoriel entre une base de Fourier et des canaux couleur antagonistes. Dans cette communication, dont une version longue est disponible [9], nous nous intéressons à ce point particulier, à la fois d’un point de vue théorique (section 3) et expérimental (section 4). Nous montrons en particulier que deux hypothèses simples sur les matrices de covariance en couleur permettent d’expliquer ce phénomène.

2 Travaux connexes

Une présentation exhaustive de la littérature sur les statistiques des images naturelles n’entre pas dans le cadre de cette communication, et nous nous concentrons ici sur deux contributions, [2] et [11], essentielles à la compréhension de nos résultats.

Redondance chromatique des images naturelles Buchsbaum et Gottshalk proposent dans [2] une analyse théorique de la structure des covariances des images couleur. Ces auteurs considèrent l’ensemble de tous les stimuli visuels (radiations) c’est à dire $S \equiv \{R(\lambda), \lambda \in \mathcal{L}\}$, où \mathcal{L} est le spectre visible, ce qui conduit aux réponses des cônes : $L = \int_{\mathcal{L}} R(\lambda)L(\lambda) d\lambda$, $M = \int_{\mathcal{L}} R(\lambda)M(\lambda) d\lambda$, $S = \int_{\mathcal{L}} R(\lambda)S(\lambda) d\lambda$. En considérant le stimulus reçu (à partir d’un point d’une scène) $R(\lambda)$ comme

une variable aléatoire, la *matrice des covariances chromatiques* est :

$$C = \begin{bmatrix} C_{LL} & C_{LM} & C_{LS} \\ C_{ML} & C_{MM} & C_{MS} \\ C_{SL} & C_{SM} & C_{SS} \end{bmatrix}, \quad (1)$$

où $C_{LL} \equiv \mathbb{E}[L \cdot L] - (\mathbb{E}[L])^2$, $C_{LM} \equiv \mathbb{E}[L \cdot M] - \mathbb{E}[L]\mathbb{E}[M] = C_{ML}$, etc., \mathbb{E} désignant l'espérance mathématique.

Soit $K(\lambda, \mu) = \mathbb{E}[R(\lambda)R(\mu)] - \mathbb{E}[R(\lambda)] \cdot \mathbb{E}[R(\mu)]$, alors $C_{LL} = \iint_{\mathcal{L}^2} K(\lambda, \mu)L(\lambda)L(\mu) d\lambda d\mu$, etc. De manière à conduire des calculs explicites, la forme de K doit être spécifiée. En l'absence d'une base d'images hyperspectrales, Buchsbaum et Gottschalk font le choix de la covariance la plus simple, à savoir $K(\lambda, \mu) = \delta(\lambda - \mu)$, δ étant la fonction indicatrice de 0. Sous cette hypothèse, les termes de la matrice C sont tous positifs. De plus la matrice est symétrique réelle et peut être diagonalisée par une matrice de vecteurs propres W

$$\begin{pmatrix} A(\lambda) \\ P(\lambda) \\ Q(\lambda) \end{pmatrix} = W^t \begin{pmatrix} L(\lambda) \\ M(\lambda) \\ S(\lambda) \end{pmatrix},$$

conduisant ainsi à trois valeurs A, P, Q qui sont décorrélées. La valeur A , associée à la plus grande valeur propre, est le canal achromatique, dont les trois coordonnées sont strictement positives, tandis que P et Q sont deux canaux antagonistes.

Le point clé de la théorie de Buchsbaum et Gottschalk est l'application du théorème de Perron-Frobenius (voir par exemple [1]), qui spécifie que les matrices positives, c'est à dire dont tous les termes sont positifs, ont un seul vecteur propre à coordonnées toutes positives, ce vecteur correspondant à la plus grande valeur propre. Ainsi, seul le canal A est une combinaison positive des valeurs d'activation des cônes L, M, S , et il porte la plus grande énergie. Les canaux P et Q sont obtenus par différences des activités des cônes. Il s'agit donc d'une élégante justification du comportement post-rétinien de type "couleurs antagonistes".

Redondance spatio-chromatique des images naturelles Des observations remarquables sur la redondance spatio-chromatique des images sont présentées dans [11], où Ruderman, Cronin et Chiao proposent un codage par *patches* et testent la théorie de Buchsbaum et Gottschalk sur une base de 12 images naturelles de feuillages. Les auteurs construisent des données LMS à partir de cette base d'images puis calculent une ACP, qui conduit à $l = \frac{1}{\sqrt{3}}(\tilde{L} + \tilde{M} + \tilde{S})$, $\alpha = \frac{1}{\sqrt{6}}(\tilde{L} + \tilde{M} - 2\tilde{S})$, $\beta = \frac{1}{\sqrt{2}}(\tilde{L} - \tilde{M})$, (avec $\tilde{L} = \text{Log } L - \langle \text{Log } L \rangle$), donc à nouveau à un système intensité + couleurs antagonistes. L'espace couleur correspondant est appelé $l\alpha\beta$.

Pour étudier la structure spatio-chromatique des images, Ruderman, Cronin et Chiao considèrent des patches 3×3 en couleur et conduisent une ACP sur l'espace correspondant (en dimension 27). Les axes principaux résultant de cette analyse expérimentale sont reproduits en Figure 1. Nous pouvons voir des fluctuations sur le canal achromatique, suivies par des fluctuations bleu-jaune (direction α) et des fluctuations vert-rouge (direction β). La structure spatiale évoque la base de Fourier,

en accord avec la stationnarité spatiale des images naturelles, comme présenté dans [3]. Sur cette même figure, il n'apparaît aucun mélange entre les canaux achromatiques, α et β , indiquant donc une décorrélation des attributs couleur et spatiaux. Ces résultats ont été confirmés dans [8]. En section 4, nous montrons des résultats similaires sur une base nettement plus volumineuse.

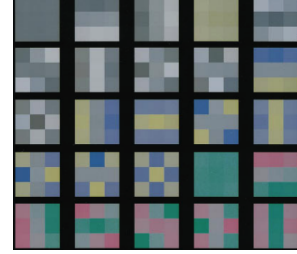


FIGURE 1 – Axes principaux des patches couleur 3×3 selon l'ordre décroissant de leurs valeurs propres, de gauche à droite et de haut en bas, adapté de [11], page 2041.

3 Stationnarité des images naturelles couleur et composantes principales

Dans cette section, nous analysons le lien entre une hypothèse de stationnarité au sens large sur les images naturelles et la structure de leurs composantes principales. Par souci de clarté, nous commençons par rappeler le cas simple des images en niveau de gris, pour lesquelles la stationnarité implique que les composantes principales sont les éléments de la base de Fourier. Nous étendons ensuite ce résultat aux images en couleur, pour lesquelles les composantes principales sont obtenues comme le produit tensoriel entre la base de Fourier et une décomposition des couleurs sous la forme achromatique + couleurs antagonistes.

3.1 Images en niveaux de gris

Soit $u : \Omega \rightarrow [0, 255]$ une image scalaire de taille $W \times H$. On note r^0, \dots, r^{H-1} les lignes et r_k^j les pixels de u . On note $c_{k,k'}^{j,j'}$ la covariance des variables aléatoires r_k^j et $r_{k'}^{j'}$. La matrice de covariance spatiale des deux vecteurs r^j et $r^{j'}$ est notée $C^{j,j'}$ et la *matrice de covariance spatiale* de l'image u est définie comme

$$C = \begin{bmatrix} C^{0,0} & C^{0,1} & \dots & C^{0,H-1} \\ C^{1,0} & C^{1,1} & \dots & C^{1,H-1} \\ \vdots & \vdots & \ddots & \vdots \\ C^{H-1,0} & \dots & \dots & C^{H-1,H-1} \end{bmatrix}. \quad (2)$$

C est une matrice de dimension $HW \times HW$, chaque matrice $C^{j,j'}$ étant de dimension $W \times W$.

A partir de maintenant, nous supposons que la covariance de u est invariante sous l'effet de translations des lignes et co-

lonnes, c'est à dire $c_{k,k'}^{j,j'} = c_{|k-k'|}^{|j-j'|}$. Il s'agit donc d'une hypothèse de stationnarité au sens large réduite aux covariances, sans hypothèse de stationnarité sur les moyennes des pixels de l'image. Cette hypothèse sera testée en section 4.

Nous supposons également que l'image a une symétrie spatiale torique, implicite dès que l'on calcule des transformées de Fourier, c'est à dire que $r_k^j = r_{k'}^{j'}$ avec $j \equiv j' \pmod{H}$ et $k \equiv k' \pmod{W}$. Il est alors immédiat que les matrices $C^{j,j'}$ sont des matrices circulantes et que C est circulante par blocs : $C = \text{circ}(C^0, C^1, \dots, C^{H-1})$. En vertu des relations classiques entre matrices circulantes et transformée de Fourier discrète (TFD), voir par exemple [4], les vecteurs propres des matrices C^j sont les vecteurs de base de Fourier : $\mathbf{e}_m = \frac{1}{\sqrt{W}} \left(1, e^{-\frac{2\pi i m}{W}}, \dots, e^{-\frac{2\pi i m(W-1)}{W}} \right)^t$ et leurs valeurs propres sont les éléments de la transformée de Fourier de la première ligne de C^j : $\hat{c}_{0,m}^{0,j} = \sum_{k=0}^{W-1} c_{0,k}^{0,j} e^{-\frac{2\pi i m k}{W}}$.

Les équations aux valeurs propres suivantes $C^j \mathbf{e}_m = \lambda_m^j \mathbf{e}_m$ peuvent être résumées sous forme matricielle par $C^j E_W = \Lambda^j E_W$, avec $\Lambda^j = \text{diag}(\hat{c}_{0,m}^{0,j}; m = 0, \dots, W-1)$ et E_W les matrices de Vandermonde :

$$E_W = [\mathbf{e}_0 | \mathbf{e}_1 | \dots | \mathbf{e}_{W-1}]. \quad (3)$$

Remarquons enfin que si une matrice M , de taille $HW \times HW$ et circulante par blocs, a la propriété que ses blocs (de taille $W \times W$) peuvent être diagonalisés dans la même base B , alors un calcul simple montre que $E_H \otimes B$ est une base de vecteurs propres de M , où \otimes est le produit de Kronecker et $E_H = [\mathbf{e}_0 | \mathbf{e}_1 | \dots | \mathbf{e}_{H-1}]$. Dans le cas de la matrice de covariance C , toutes les sous-matrices C^j ont la même base de vecteurs propres E_W et donc $E_H \otimes E_W$ fournit une base de vecteurs propres pour C . En raison de la symétrie de la matrice de covariance, les parties complexes des exponentielles de la base de Fourier s'annulent (voir par exemple [5]), de sorte que la base 2D de cosinus est également une base de vecteurs propres de C .

3.2 Le cas des images couleur

Soit $\mathbf{u} : \Omega \rightarrow [0, 255]^3$ une image RGB, et pour tout $(j, k) \in \Omega$, soit $\mathbf{u}(j, k) = (R(j, k), G(j, k), B(j, k))$ le vecteur dont les composantes sont les valeurs de rouge, vert et bleu associées au pixel de position (j, k) . Nous définissons la matrice de covariance spatio-chromatique $\mathbf{c}_{k,k'}^{j,j'}$ entre deux pixels de la manière suivante

$$\begin{bmatrix} c_{k,k'}^{j,j'}(R, R) & c_{k,k'}^{j,j'}(R, G) & c_{k,k'}^{j,j'}(R, B) \\ c_{k,k'}^{j,j'}(G, R) & c_{k,k'}^{j,j'}(G, G) & c_{k,k'}^{j,j'}(G, B) \\ c_{k,k'}^{j,j'}(B, R) & c_{k,k'}^{j,j'}(B, G) & c_{k,k'}^{j,j'}(B, B) \end{bmatrix}. \quad (4)$$

Nous conservons l'hypothèse de stationnarité réduite aux covariances : $\mathbf{c}_{k,k'}^{j,j'} = \mathbf{c}_{|k-k'|}^{|j-j'|}$. Dans le cas particulier où $j' = j$ et $k' = k$, l'autocovariance chromatique est notée simplement \mathbf{c}^0 . En procédant de la même façon que pour les images en niveau de gris mais en remplaçant $c_{k,k'}^{j,j'}$ par $\mathbf{c}_{k,k'}^{j,j'}$, nous définissons les

matrices de covariance spatio-chromatique $\mathbf{C}^{j,j'}$ entre les vecteurs r^j et $r^{j'}$, et la matrice de covariance spatio-chromatique \mathbf{C} de l'image \mathbf{u} , une matrice de dimensions $3HW \times 3HW$.

Sous l'hypothèse que tous les éléments des matrices (4) sont strictement positifs, le théorème de Perron-Frobenius implique que chacune des matrices $\mathbf{c}_{k,k'}^{j,j'}$ admet une base de vecteurs propres qui s'écrit comme une triade composante achromatique + canaux de couleurs antagonistes (A, P, Q) . En supposant que les matrices (4) peuvent être diagonalisées dans la même base de vecteur propres (ou de manière équivalente que les matrices commutent), alors de même que précédemment les vecteurs propres de la matrice de covariance spatio-chromatique \mathbf{C} peuvent s'écrire comme un produit tensoriel entre la base de Fourier et les couleurs antagonistes, c'est à dire précisément les vecteurs propres trouvés expérimentalement dans [10] ! Nous résumons ces observations dans la proposition suivante.

Proposition 1 Soit $\mathbf{u} : \Omega \rightarrow [0, 255]^3$ une image RGB, définie sur un domaine périodisé spatialement Ω . Supposons que

1. Les matrices $\mathbf{c}_{k,k'}^{j,j'}$ définies en (4) ne dépendent que des distances $|j - j'|$, $|k - k'|$, c'est à dire que la covariance de \mathbf{u} est stationnaire ;
2. Toutes les matrices $\mathbf{c}_{k,k'}^{j,j'}$ sont strictement positives, c'est à dire que leurs éléments sont strictement positifs ;
3. Pour tout (j, k) et tout (j', k') , les matrices \mathbf{c}^0 et $\mathbf{c}_{k,k'}^{j,j'}$ commutent.

Alors, les vecteurs propres de la matrice de covariance spatio-chromatique \mathbf{C} peuvent s'écrire comme le produit de Kronecker $(A, P, Q) \otimes \mathbf{e}_{m,l}$, où (A, P, Q) est la triade canal achromatique + canaux de couleurs antagonistes, et $\mathbf{e}_{m,l}$ est la base de Fourier de cosinus 2D.

4 Validation sur deux bases d'images naturelles

Afin de tester expérimentalement les hypothèses de la Proposition 1, nous considérons deux bases d'images. La première est une base d'images non calibrées issue de la base décrite dans [6], comprenant 2,3 millions d'images de taille 1024×768 , issues du site Flickr. Les images de la base ont été réordonnées aléatoirement pour limiter les biais dus aux séries d'images. Le grand nombre d'images de cette base permet en particulier de tester la stabilité des résultats. La deuxième base d'images est constituée de 7000 images RAW de scènes naturelles. Afin de dépendre aussi peu que possible du schéma de dematrigage de ces images, nous avons construit des images couleur par sous-échantillonnage d'un facteur 4, chaque pixel étant obtenu en gardant intacts les informations des canaux R et B et en moyennant les deux canaux V de la trame de Bayer. L'avantage de cette base est que ses images sont exemptes d'opération de post-processing du type correction gamma, balance des blancs, compression, et de ce fait permettent une meilleure approximation de l'information de luminance que les images de la première base.

Valeurs propres et vecteurs propres de \mathbf{c}^0 Pour les bases d'images Flickr et RAW nous avons obtenu respectivement les matrices d'autocovariance suivantes :

$$\mathbf{c}_{\text{Flickr}}^0 = \begin{bmatrix} 0.0719 & 0.0651 & 0.0612 \\ 0.0651 & 0.0713 & 0.0710 \\ 0.0612 & 0.0710 & 0.0851 \end{bmatrix}, \quad (5)$$

$$\mathbf{c}_{\text{RAW}}^0 = \begin{bmatrix} 0.0022 & 0.0021 & 0.0021 \\ 0.0021 & 0.0021 & 0.0022 \\ 0.0021 & 0.0022 & 0.0024 \end{bmatrix} \quad (6)$$

ce qui confirme l'hypothèse de positivité sur \mathbf{c}^0 . On vérifie numériquement qu'il en va de même pour les matrices \mathbf{c}^d , pour toutes les distances d . Remarquons que les covariances observées sur la base d'images RAW sont plus faibles que celles de la base Flickr. Nous pensons que cette différence provient essentiellement du contraste nettement plus important des images postées sur Flickr. Malgré cette différence, les vecteurs propres des deux matrices sont très proches :

$$\begin{cases} A_{\text{Flickr}} = (0.5483, 0.5761, 0.6061) \longleftrightarrow \lambda_1 = 0.2080, \\ P_{\text{Flickr}} = (0.7179, 0.0474, -0.6945) \longleftrightarrow \lambda_2 = 0.0170, \\ Q_{\text{Flickr}} = (0.4289, -0.8160, 0.3876) \longleftrightarrow \lambda_3 = 0.0034. \end{cases} \quad (7)$$

et

$$\begin{cases} A_{\text{RAW}} = (0.5679, 0.5683, 0.5954) \longleftrightarrow \lambda_1 = 0.0065, \\ P_{\text{RAW}} = (0.7210, 0.0055, -0.6930) \longleftrightarrow \lambda_2 = 0.0002, \\ Q_{\text{RAW}} = (0.3971, -0.8228, 0.4066) \longleftrightarrow \lambda_3 = 7.8 \cdot 10^{-7}. \end{cases} \quad (8)$$

L'estimation des commutateurs $[\mathbf{c}^0, \mathbf{c}^d]$ pour chaque distance $d > 0$ montre que ceux-ci sont très proches de 0 : la plus grande déviation de la matrice nulle que nous avons observée sur les deux bases a une norme quadratique normalisée de 0.04 pour la base Flickr et 0.025 pour la base RAW, ce qui montre que l'hypothèse de commutativité est quasiment satisfaite.

Enfin, nous montrons Figure 2 les vecteurs d'une ACP sur des patches 5×5 extraits d'images de la base Flickr, qui sont bien de la forme prédite par la Proposition 1.

5 Conclusion

Dans cet article, nous avons montré les liens théoriques entre stationnarité au second ordre et commutativité des matrices de covariance spatiochromatique d'une part, et structure spatiochromatique de ces matrices d'autre part. Notre analyse était motivée par la volonté de comprendre pourquoi une ACP appliquée à un ensemble suffisamment varié de patches ou d'images naturelles faisait systématiquement apparaître une base composée d'un produit tensoriel entre une base de Fourier et des canaux antagonistes. Les tests numériques menés sur deux grandes bases d'images montrent que nos hypothèses sont satisfaites avec un très bon degré d'approximation. Pour déduire de ces résultats des propriétés du SVH, une base importante et variée d'images multispectrales serait nécessaire, afin de minimiser l'influence des fonctions de sensibilité des capteurs numériques dans les statistiques.

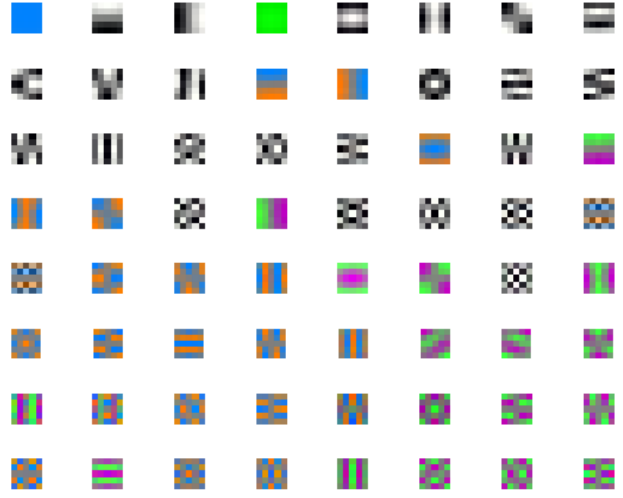


FIGURE 2 – Nous avons extrait 1.900.000 patches 5×5 d'un ensemble de 4000 images de la base Flickr. Cette Figure représente les 64 premières composantes d'une ACP sur cet ensemble de patches. Les composantes sont normalisées entre 0 et 1 pour améliorer la visualisation.

Références

- [1] A. Berman and R.J. Plemmons. *Nonnegative Matrices in the Mathematical Sciences*. SIAM, 1987.
- [2] G. Buchsbaum and A. Gottschalk. Trichromacy, opponent colours coding and optimum colour information transmission in the retina. *Proc. Royal Society of London B*, 220 :89–113, 1983.
- [3] D.J. Field. Relations between the statistics of natural images and the response properties of cortical cells. *J. Opt. Soc. Am.*, 4(12) :2379–2394, December 1987.
- [4] M.W. Frazier. *An Introduction to Wavelets Through Linear Algebra*. Springer, 2001.
- [5] R.M. Gray. *Toeplitz and Circulant Matrices : A review*. Now Publishers Inc., 2006.
- [6] J. Hays and A. Efros. Scene completion using millions of photographs. In *ACM Transactions on Graphics (SIGGRAPH)*, volume 26, 2007.
- [7] B. Olshausen and D.J. Field. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Letters to Nature*, 381 :607–609, 1996.
- [8] C. Párraga, T. Troscianko, and D. Tolhurst. Spatiochromatic properties of natural images and human vision. *Current biology*, 6(12) :483–487, 2002.
- [9] Edoardo Provenzi, Julie Delon, Yann Gousseau, and Baptiste Mazin. On the second order spatiochromatic structure of natural images. *Vision Research*, 2015. To appear.
- [10] D.L. Ruderman. Origin of scaling in natural images. *Vision Research*, 37 :3385–3398, 1996.
- [11] D.L. Ruderman, T.W. Cronin, and C. Chiao. Statistics of cone responses to natural images : implications for visual coding. *J. Opt. Soc. Am. A*, 15(8) :2036–2045, August 1998.