



CODAGE A COMPOSANTES VISUELLES D'IMAGES FIXES

A. SAADANE, H. SENANE

IRESTE/SEI/ATI La Chantrerie CP 3003 44087 NANTES CEDEX 03
Tél: (33) 40 68 30 46 Fax (33) 40 68 30 66 Email asaadane@ireste.fr

RÉSUMÉ

Cet article décrit un schéma de codage d'images fixes basé entièrement sur des aspects psychovisuels. Dans la partie analyse, le système visuel humain (SVH) est modélisé par une batterie de 17 canaux sélectifs à la fois en angulaire et en radial. Les filtres utilisés pour réaliser une telle décomposition s'inspirent dans leur principe de base des filtres "Cortex" de Watson.[4] Au niveau de la quantification, deux approches se différenciant par la manière de quantifier la luminance, ont été mises en oeuvre. Dans les deux cas les performances obtenues sont comparables à celles des schémas conventionnels mais avec une qualité visuelle des images reconstruites nettement supérieure pour l'approche psychovisuelle. L'approche est d'autant plus intéressante qu'une forte réduction de débit (avec conservation de la qualité visuelle) est possible moyennant une quantification vectorielle qui prendrait en compte les effets de masquage angulaire.

INTRODUCTION:

Cette étude s'inscrit dans le cadre du codage d'images fixes. Elle effectue un prolongement dans le sens d'une prise en compte beaucoup plus fondamentale des propriétés du système visuel humain. Il est généralement admis que ce dernier analyse les signaux d'entrée moyennant une batterie de canaux dont chacun est sensible à une gamme d'orientation et à une bande de fréquences spatiales données.

Le but de ce papier est de décrire un schéma de codage entièrement psychovisuel. Les sous bandes utilisées, au niveau de l'analyse, sont issues d'études psychophysiques qui ont été menées. Leur caractérisation est basée sur la variation de la fonction de masquage. Le sous échantillonnage optimal adapté à ce type de représentation est effectué moyennant un algorithme qui sera décrit. Les quantificateurs psychovisuels, associés à chacune des sous-bandes, utilisent le contraste local à bande limitée. On montrera en particulier l'intérêt de quantifier une telle grandeur par rapport à la quantification, au seuil différentiel de visibilité, du contraste tel que défini par Weber. Pour atteindre une qualité visuelle excellente au niveau de la reconstruction, la précision des quantificateurs est choisie de telle façon que les dégradations induites restent situées au dessous de leur seuil de perception.

ABSTRACT

In this paper, we describe a coding scheme which is entirely based on psychovisual aspects. In the analysis stage, the Human Visual System (HVS) has been modeled by a set of seventeen channels which are selectively sensitive to a restricted range of spatial frequencies and orientations. The filters used to perform this decomposition are similar to the "Cortex" filters of Watson.

At the quantization level, two kinds of quantizers have been tested. They differ in the manner that the luminance has been quantized.

The results obtained, in the two cases and in terms of signal-to-noise ratio and entropy, are similar to those of conventional schemes. However the visual quality of reconstructed images is much higher in the psychovisual approach.

As these bit rates may be reduced by using a vector quantization, this approach seems very interesting

II - COMPOSANTES VISUELLES:

Pour caractériser les canaux visuels, des études psychovisuelles basées sur le comportement de la fonction de masquage, ont été menées. Les résultats montrent que l'effet de masquage dans le cas des signaux sinusoïdaux a une portée relative assez importante (de l'ordre de 3) en moyennes et basses fréquences et que cette portée diminue (jusqu'à une octave) dans les hautes fréquences. La sélectivité angulaire par contre, croit avec la fréquence radiale et varie de 100 deg en faible fréquence à 30 deg en hautes fréquences. La synthèse de ces résultats nous ont permis de modéliser la partie périphérique du système visuel humain au moyen d'un ensemble de filtres dont une représentation est donnée en figure 1. Ce pavage fréquentiel, en concordance avec les résultats bibliographiques [1] [2], fait apparaître un canal basse-fréquence non directionnel et trois bandes de fréquences radiales (notées de III à V) décomposées elles mêmes en canaux angulaires dont le nombre dépend de la bande radiale considérée.

III - SCHEMA DE CODAGE

III - 1: Description

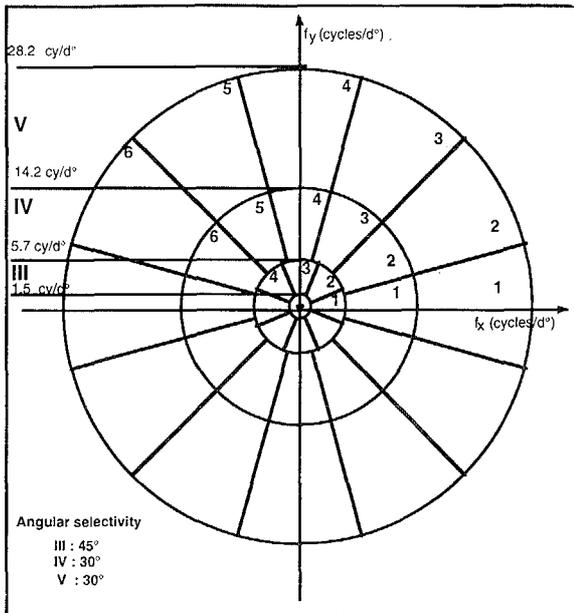


figure 1: Modélisation du système visuel humain

Le schéma fonctionnel du codeur respectant une telle décomposition est donné en figure 2. Dans ce schéma les niveaux de gris de l'image d'entrée sont transformés en luminances réelles par la modélisation de la non linéarité de l'écran basée sur des mesures photométriques. Une transformation de Fourier et un filtrage par des filtres "Cortex" sont ensuite appliqués. Une transformation de Fourier inverse est alors calculée sur chacune des sous-bandes ainsi obtenues. Compte tenu du fait que la découpage radial est non dyadique et que la sélectivité angulaire est non régulière, la surface relative occupée par les sous-bandes varie d'une couronne à l'autre. Un sous échantillonnage ad'hoc suppose donc l'utilisation de matrices ayant un déterminant approprié à chacune des couronnes. Le choix d'un tel déterminant est dépendant du rapport entre la surface totale du spectre de l'image et celle occupée par la sous-bande considérée. Pour un déterminant fixée la matrice de sous-échantillonnage qui est utilisée est celle qui minimise l'énergie totale des bandes (après sous-échantillonnage). Après le sous échantillonnage deux types de quantificateurs scalaires, décrits ci dessous, ont été mis en oeuvre. Au niveau de la reconstruction, des opérations analogues sont effectuées. L'image reconstruite est obtenue dans le domaine fréquentiel par la sommation des différentes sous bandes spectrales. Après une transformée de fourier inverse et la conversion luminance - niveaux de gris, l'image spatiale est affichée sur écran.

III - 2: Filtrés Cortex

L'avantage de ces filtres, largement décrits dans [4], réside dans le fait que leur réponse impulsionnelle, tout comme le champ réceptif de la plupart des cellules du Cortex visuel humain, ressemble à une fonction de Gabor 2-D. Ces filtres sont le résultat du produit entre des filtres qui définissent la sélectivité radiale (appelés filtres Dom) et de filtres qui caractérisent la sélectivité angulaire (appelés filtres Fan).

Les filtres Dom sont engendrés par une simple différence entre deux filtres "Mesa" d'indices

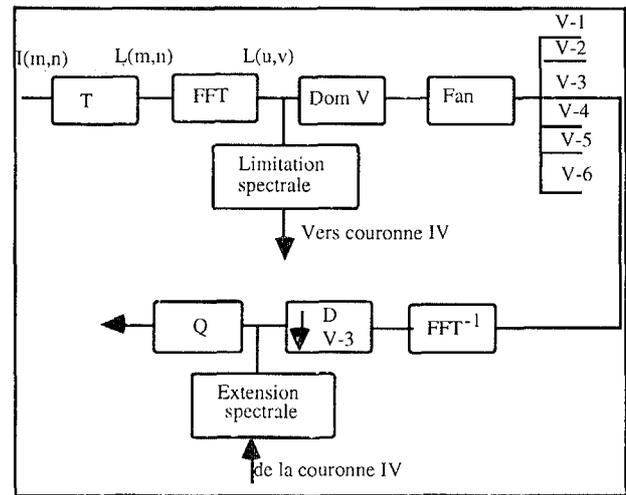


Figure 2: Ensemble des opérations nécessaires à la décomposition, le sous-échantillonnage et la quantification de chacune des sous bandes (seul est présenté ici le traitement de la bande V-3)

consécutifs (et donc de fréquence de coupure radiale différente)

$$\text{Dom}_i(u,v) = M_{i-1}(u,v) - M_i(u,v)$$

où u et v représentent les fréquences spatiales. Le filtre Mesa d'indice i est donné par

$$M_i(u,v) = M_0 \left(\left(\prod_{j=1}^{i-1} s_j \right) u, \left(\prod_{j=1}^{i-1} s_j \right) v \right)$$

$$\text{avec } M_0(u,v) = \left(\frac{\gamma}{f_0} \right)^2 \exp \left[-\pi \left(\frac{\gamma}{f_0} \right)^2 \right] \otimes \Pi \left(\frac{\omega}{2f_0} \right)$$

où $\Pi \left(\frac{\omega}{2f_0} \right)$ est la fonction porte 2-D à symétrie circulaire, centrée en zéro et de rayon f_0 .

- γ est un paramètre de raideur de l'atténuation, lié à l'écart-type σ_0 de la fonction gaussienne engendrant les zones de transition de gain des filtres.

- s est un facteur d'échelle.

Pour les filtres "Fan", l'évolution de leur gain selon les directions angulaires est également décrite par un flou gaussien sur un filtre angulaire idéal. Selon la direction horizontale cette évolution est donnée par

$$M'_0(u,v) = H(v) * \gamma_b \exp(-\pi \gamma_b^2 v^2)$$

où $H(v)$ est le filtre "échelon" qui sépare le domaine fréquentiel en deux sous domaines de gain unité pour l'un et nul pour l'autre.

- γ_b est un nouveau paramètre de raideur.

Compte tenu que le filtre échelon n'a pas de variation selon l'axe u , la convolution peut s'écrire dans ce cas

$$M'_0(u,v) = \int_{-\infty}^v \gamma_b \exp(-\pi \gamma_b^2 w^2) dw = F(\gamma_b, v)$$

L'expression du filtre échelon ayant une orientation θ est alors

$$M'_{\theta}(u,v) = F(\gamma_b(v \cos\theta - u \sin\theta))$$

Le filtre Fan correspondant à la k^{eme} direction est finalement donné par

$$\text{Fan}_k(u,v) = M'_{\theta_k}(u,v) - M'_{\theta_{k+1}}(u,v)$$

A titre d'exemple le spectre de la sous bande V-1 est donné en figure 3

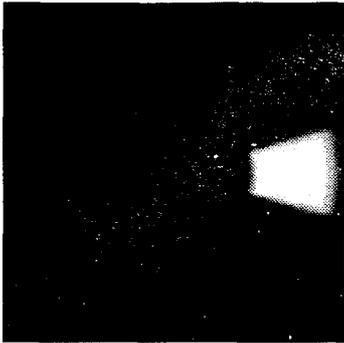


Figure 3: Le filtre cortex θ résulte de la multiplication du filtre Dom par le filtre Fan

III - 3: Quantification

Deux approches pour la quantification ont été étudiées. La première utilise comme grandeur un contraste tel que défini par le contraste de Weber et comme pas de quantification le seuil différentiel de visibilité correspondant à la sous-bande. La deuxième approche, psychovisuelle aussi, a été développée au laboratoire. Décrite dans [3], cette approche utilise, pour la menée des tests, des signaux complexes (textures filtrées) et tient compte de l'effet de cumul spatial des différentes dégradations. Rappelons que cet effet, présent dans tous les schémas de codage, est dû à la quantification simultanée des différents contrastes. Dans cette approche, trois principes ont régi la construction du quantificateur;

- Zéro est le premier niveau de reconstruction
- Les seuils de décision et les niveaux de reconstruction sont symétriques par rapport à zéro. Ceci est dû à une distribution, approximativement symétrique autour de zéro, des coefficients à l'intérieur d'une sous-bande
- La perception des dégradations décroît avec l'augmentation du contraste.

Pour tenir compte du fait que la nature locale du contraste d'une image est fortement liée à son contenu spatio-fréquentiel le contraste local à bande limitée [5] a été utilisé. Après l'avoir adapté à la décomposition de la figure 1 ce contraste est défini par

$$c_{i,j}(m,n) = \frac{L_{i,j}(m,n)}{\sum_{k=0}^{i-1} \sum_{l=0}^{j-1} L_{k,l}(m,n)}$$

où i et k sont les numéros des bandes de fréquences radiales (figure 1) et j et l ceux des fréquences

angulaires. l_k est le nombre de secteurs angulaires de la k^{eme} bande radiale. Ce contraste dépend à la fois du contenu spatial et du contenu spatio-fréquentiel de l'image.

Le dénominateur représente la luminance moyenne locale correspondante à toutes les bandes de support spectral inférieur à celui de la j^{eme} bande.

Dans cet espace de contraste, les résultats obtenus montrent que les quantificateurs suivent une loi linéaire. Cette loi est conservée d'une sous-bande à l'autre mais avec un pas de quantification différent.

L'utilisation du contraste local à bande limitée dans un schéma de codage doit donc prendre en compte la nécessité de reconstruire progressivement la basse fréquence au fur et à mesure que l'on "monte" en fréquence radiale. Cette contrainte se traduit par une procédure qui s'effectuera dans l'ordre suivant:

- quantification, codage et transmission des luminances de la basse fréquence
- quantification, codage et transmission des contrastes de chacune des bandes de la couronne III
- Construction de la basse fréquence relative à la couronne IV par sommation de la basse fréquence et de la couronne III
- Répétition du procédé pour les couronnes IV et V.

IV - RESULTATS ET CONCLUSION

Les performances de ce schéma de codage à composantes visuelles ont été évaluées en termes d'entropie, de rapport signal à bruit et de qualité visuelle. Les pas de quantification du contraste local à bande limitée utilisés sont donnés par le tableau 1.

Sous-bande	pas de quantification
III	0,01
IV.1 et IV.4	0,01
IV.2 et IV.3	0,015
IV.5 et IV.6	0,015
V.1 et V.4	0,04
V.2 et V.3	0,06
V.5 et V.6	0,06

Tableau 1: Pas de quantification utilisés

On notera que les valeurs utilisées pour les directions obliques sont légèrement supérieures à celles relatives aux directions horizontales et verticales. Ceci est due à une plus grande sensibilité du SVH à ces dernières directions.

Malgré un sous-échantillonnage non critique (accroissement du nombre d'échantillon en moyenne de 1,4) du à la non régularité des sous bandes visuelles, les résultats obtenus pour les quatre images "Port", "Couple", "Fruit" et "Clown" sont donnés par le tableau 2.

Dans les deux cas les résultats obtenus sont très probants en terme de qualité visuelle. Avec un rapport signal à bruit relativement bas, les images obtenues sont d'excellente qualité (défauts non visibles à une distance d'observation de 4H), alors qu'avec ces mêmes rapports les schémas classiques présentent des images de qualité médiocre.



Image	Quantification du contraste (Weber)		Quantification du contraste local à bande limitée	
	Entropie (bpp)	SNR (dB)	Entropie (bpp)	SNR (dB)
Port	1,72	38,53	1,17	36,54
Couple	1,5	37,00	1,13	36,14
Fruit	1,2	39,59	0,9	38,56
Clown	1,18	36,60	1,1	37,97

Tableau 2: Résultats obtenus avec une quantification du contraste tel que défini par Weber et du contraste local à bande limitée



Figure 4: Image Originale

L'utilisation du contraste local à bande limitée est pertinente pour au moins deux raisons;

- l'effet de masquage des fréquences radiales inférieures à une fréquence donnée est déjà pris en compte.
- la structure de l'image d'erreur de reconstruction se retrouve totalement corrélée avec celle de l'image d'origine.

A titre d'illustration, l'image originale "Clown", l'image reconstruite et l'image d'erreur sont données respectivement en figure 4, 5 et 6

La démarche est d'autant plus prometteuse qu'une réduction de débit très significative est possible moyennant une quantification vectorielle avec prise en compte de l'effet de masquage inter sous-bandes angulaires.

V - REFERENCE

- [1] D.C. BURR and S.A. WIJESUNDR
"Orientation discrimination depends on spatial frequency"
Vision Research Vol 31, N°7/8 pp 1449-1452, 1991
- [2] S.J. ANDERSON, D.C. BURR and M.C. MORRONE
"Two dimensional spatial and spatial frequency selectivity of motion sensitive mechanisms in human vision"
J.O.S.A. Vol 8, pp 1340-1351 1991
- [3] A. SAADANE, H. SENANE, D. BARBA
"On the design of psychovisual quantizers for a visual subband image coding"
V.C.I.P'94 Chicago
- [4] A.B. WATSON
"Efficiency of a model human image code"
J. Opt.Soc.of Am. Vol 4, n° 12, pp 2401-2417 (1987)
- [6] É. PELI
"Contrast in complex images"
J.Opt.soc.Am, A/vol.7, n°10, pp 2032-2040 (October 1990)



Figure 5: Image Reconstituée

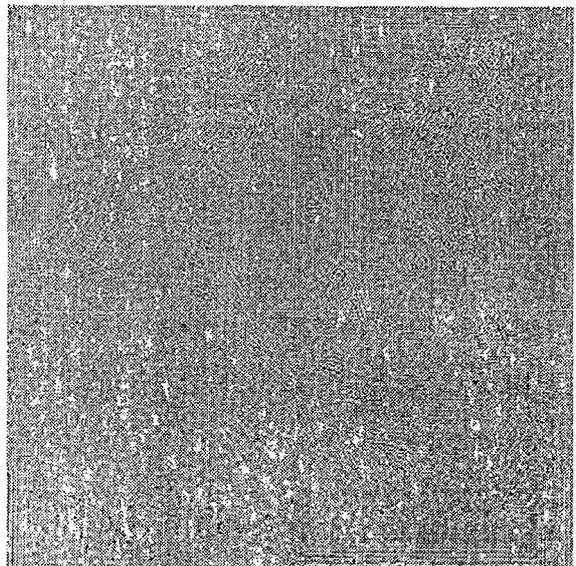


Figure 6: Image d'erreur