

# Caractérisation d'images par analyse en composantes indépendantes

Hervé Le Borgne<sup>1</sup>, Anne Guérin-Dugué<sup>2</sup>  
L.I.S – I.N.P.G.  
46 avenue Félix Viallet  
F-38000 Grenoble Cedex  
Mél : {hleborgn, guerin}@lis.inpg.fr

## *Résumé*

L'analyse en composantes indépendantes appliquée à une collection d'images permet d'obtenir des détecteurs qui ont des similarités avec les cellules simples du cortex visuel de type passe-bande orienté. Les propriétés spectrales des détecteurs obtenus sont en relation avec la cohérence spectrale des images de la collection considérée. Les détecteurs ainsi extraits offrent plusieurs avantages. Ils se « placent » dans l'espace de Fourier relativement aux spectres des images dont ils sont issus (pas de placement a priori dans tout l'espace de Fourier). Le codage des images qui en résulte, a la propriété d'être épars. Il permet une représentation parcimonieuse des images sur quelques détecteurs les plus actifs. Enfin, la propriété d'indépendance permet une simplification essentielle de l'estimation des densités de probabilité des codes caractérisant une image. Ces propriétés sont illustrées sur une application de catégorisation de scènes naturelles selon leur contexte sémantique. Les contextes choisis sont les scènes de ville, les scènes d'intérieur, les paysages « fermés » (forêts, montagnes, ...), et les paysages « ouverts » (plages, champs, déserts, ...). Une étape de modélisation statistique des distributions globales des caractéristiques extraites permet de définir d'une part la signature de l'image et d'autre part les dissimilarités entre images, implantées par la divergence de Kullback. A partir de ces mesures de dissimilarités, les expériences présentées illustrent les potentialités de la méthode en matière de catégorisation dans une base d'images.

**Mots clés :** analyse en composantes indépendantes, codage épars, divergence de Kullback, catégorisation de scènes.

---

<sup>1</sup> Hervé Le Borgne est financé par la région Rhône-Alpes dans le cadre du projet EMERGENCE 2000 « ASCII : Architecture Sémanti-Cognitive d'Indexation d'Images ».

<sup>2</sup> Anne Guérin-Dugué est actuellement Chargé de Recherche, détachée auprès de l'INRIA Rhône-Alpes.

## I. Introduction

L'interprétation d'images naturelles est un domaine de recherche très actif et prometteur, tant du point de vue de la modélisation de notre perception visuelle, que du point de vue de la recherche algorithmique en segmentation ou catégorisation d'images. Les travaux présentés ici se situent à la frontière de ces deux domaines. En amont, les détecteurs de caractéristiques sont inspirés des modèles corticaux des cellules simples du cortex visuel. En aval, les caractéristiques ainsi extraites des images servent de signatures dans une problématique de catégorisation de bases d'images.

Partant du constat que chez l'être humain, la reconnaissance visuelle des scènes, des objets, des visages est généralement rapide, automatique et fiable, il est important de s'inspirer de ces modèles de perception pour l'analyse de scènes en vision par ordinateur. De nombreuses études mettent en évidence les similitudes entre les cellules simples du cortex visuel et les détecteurs extraits des images par ACI [2,6,13,16], montrant que les détecteurs corticaux pourraient résulter de l'application d'un principe de réduction de redondance par indépendance statistique de leurs activités. De plus le codage des images dans cette base de détecteurs a la propriété d'être épars : les informations extraites des images sont concentrées sur un faible nombre de détecteurs les plus actifs.

L'application de cette technique à l'analyse de scènes [3,11,23] est une voie prometteuse. L'approche choisie est de construire automatiquement à partir des images d'une base de référence, des familles de détecteurs (jouant le rôle de filtres corticaux) aptes à organiser une base d'images suivant des contextes facilement identifiables par des sujets humains. Les contextes sémantiques ont été choisis à partir d'études psychophysiques [15] montrant la capacité des sujets humains à discerner des environnements de type « villes », « pièces d'intérieur », « paysages fermés », et « paysages ouverts », à partir d'images en basse résolution, sans avoir préalablement identifié les objets composant la scène. C'est une stratégie de type *Coarse-to-Fine*. D'autres supports de catégorisation sont également possibles. Citons par exemple, les dichotomies suivantes : paysages versus scènes urbaines [7,18,22], ou scènes d'intérieur versus scènes d'extérieur [21]. La catégorisation proposée ici est plus riche sur quatre environnements. On trouvera dans [8] une étude des propriétés spectrales des images appartenant à ces catégories.

Après avoir décrit la méthodologie et l'implantation de l'ACI sur une base d'images, les détecteurs extraits sont analysés quant à leurs propriétés fréquentielles et de codage épars. La propriété d'indépendance des codes permet une estimation simple et faible des densités de probabilité des codes utiles à l'établissement de mesures de dissimilarités entre images. Enfin les potentialités de la méthode est illustrée sur une tâche de catégorisation de scènes naturelles.

## II. Principe

Issue des travaux sur la séparation de sources [10], l'analyse en composantes indépendantes [4] est un outils d'analyse de données permettant de trouver une nouvelle base de représentation dans laquelle les projections des observations sur cette base fournissent des composantes indépendantes. En comparaison, une analyse en composantes principales fournit une base de représentation dans laquelle les composantes sont simplement décorréliées. L'ACP est employée, dans notre cas, comme prétraitement en amont de l'ACI.

Soit  $\Phi = \{\phi_i, i = 1..m\}$  cette nouvelle base. Soit  $e \in \mathfrak{R}^p$  une observation. Dans cette base, soient  $s_i = \langle e, \phi_i \rangle$ , les composantes statistiquement indépendantes, on a :

$$e = \sum_{i=1}^m s_i \cdot \phi_i . \quad (1)$$

De nombreux critères ont été proposés pour estimer les primitives  $\phi_i$ , comme par exemple (i) la minimisation de l'information mutuelle entre les sorties ( $s_1, s_2, \dots, s_m$ ) [16], (ii) l'optimisation d'une fonction de contraste, ou d'entropie utilisant les statistiques d'ordre quatre [4,9]. Nous avons utilisé ici l'algorithme appelé *Fast ICA* [9] pour ses propriétés de convergence rapide et d'extraction séquentielle des sources. Il est basé sur l'optimisation d'une fonction d'entropie (appelée entropie négative ou *negentropy*) appliquée sur les sorties. En appliquant un tel modèle (équation 1) aux images, on considère alors qu'une image  $e$  est une combinaison linéaire de fonctions de base  $\phi$ . Les coefficients de pondération forment le code ( $s=[s_1, s_2, \dots, s_m]$ ) de l'image, comme l'illustre la figure 1.

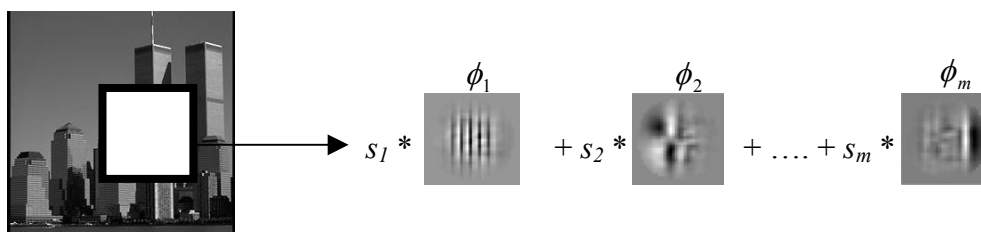


Figure 1 : illustration de la décomposition d'une image dans la base  $\phi$ .

### III. Description de la base d'images

La base d'images utilisée est constituée de 200 images 128x128 de scènes naturelles (extraction de la base COREL) se répartissant en 4 catégories sémantiques de 50 images chacune : villes, pièces d'intérieur, paysages « fermés » (forêts, montagnes, vallées, ...), paysages « ouverts » (plages, déserts, champs, ...). La base d'images est de petite taille (seulement 200 images), mais les images ont été sélectionnées pour former une base ayant une grande variabilité intra-classe.

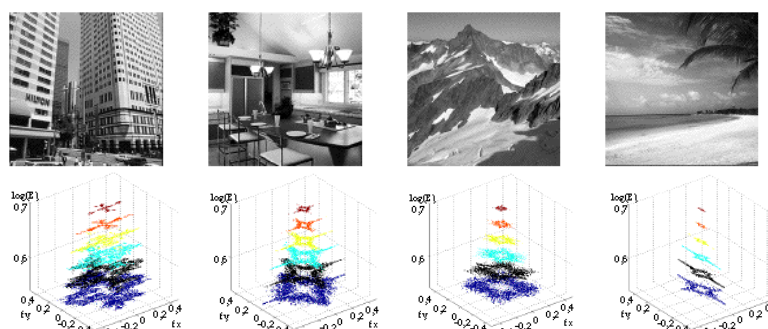


Figure 2 : formes spectrales prototypiques d'une image et son spectre d'énergie pour chaque catégorie.

Les images dans chacune des classes ont une cohérence spectrale (cf. figure 2) se modélisant par un gabarit suivant l'anisotropie/isotropie en orientation du spectre d'énergie dans l'espace de Fourier [14]. On remarque par exemple une forme spectrale en « croix » pouvant caractériser les scènes d'intérieur et aussi les scènes urbaines (présence d'horizontales et de verticales), ou une forme anisotrope sur les fréquences verticales pour les scènes « ouvertes » caractérisées par la présence d'une ligne d'horizon, ou une forme plutôt isotrope pour les paysages « fermés » (scènes plus texturées).

### IV. Méthodologie, implantation

La base d'images regroupe a priori quatre populations d'images. Nous avons réalisé une analyse en composantes indépendantes pour chacune de ces quatre populations. Le problème particulier lié à l'image réside dans la grande dimensionnalité des observations. Les données en entrée de l'algorithme *Fast-ICA* sont des imagelettes de taille  $p=l \times l$  (ici  $l=32$  donc  $p=1024$ ) extraites à des positions aléatoires dans les images d'une catégorie. On extrait 30 imagelettes dans chacune des

50 images des quatre catégories, donc nous disposons de  $n = 1500$  observations de dimension  $p = 1024$  en entrée de l'algorithme pour chacune des quatre ACI. Les imagerie sont apodisées par une fenêtre de Hamming afin de réduire l'effet de Gibbs (augmentant artificiellement les énergies des fréquences spatiales verticales et horizontales) et permettre un effet de focalisation au centre de l'imagerie. De plus, cette apodisation réduit la dimension des données en supprimant les pixels périphériques et en ne conservant que les pixels centraux ayant une variance significative (environ 600 pixels). Une analyse en composantes principales permet ensuite un blanchiment des données et la réduction de leur dimension à  $m = 30$ . Ce choix, indépendant du nombre d'imagerie extraites de chaque image, est dû à un compromis sur le temps de calcul et la complétude des caractéristiques extraites. A la convergence de l'algorithme *Fast-ICA*, on obtient au plus  $m = 30$  fonctions de base ou primitives ( $\phi_i$ ), interprétées comme la réponse impulsionnelle de filtres RIF 2D de taille  $32 \times 32$ .

L'utilisation de l'ACI se justifie à deux niveaux. D'une part, elle permet d'extraire des détecteurs similaires aux détecteurs corticaux du cortex visuel, et d'autre part, la propriété d'indépendance permet d'estimer les densités de probabilité de variables aléatoires multivariées par le produit des densités de probabilités marginales unidimensionnelles.

La méthodologie de caractérisation des images par ACI pour la catégorisation d'images se décompose en quatre étapes :

- ACI sur chaque catégorie,
- estimation des densités de probabilité des codes obtenus en filtrant une image par les filtres ACI,
- estimation des dissimilarités entre images par la divergence de Kullback entre les distributions multivariées issues des deux images,
- classement des images suivant ces dissimilarités.

## ***V. Présentation et propriétés des filtres ACI***

### ***V.1 Extraction et modélisation***

En réalisant une analyse en composantes indépendantes pour chacune des quatre catégories d'images, on spécifie ainsi la famille de filtres obtenus à sa catégorie d'origine. Conformément à la méthodologie décrite au paragraphe IV, quatre populations de 30 filtres sont obtenues.

En première approximation, ces filtres peuvent être assimilés à des filtres de type « passe-bande orienté ». Afin de quantifier l'orientation et le positionnement des filtres dans le plan fréquentiel, nous cherchons le modèle de filtre de Gabor 2D le plus proche. Nous récupérons ainsi deux paramètres de positionnement de la fréquence centrale  $(u_0, v_0)$ , et deux paramètres d'étalement spatial  $(\sigma_x, \sigma_y)$  pour chaque filtre.

On peut qualitativement repérer trois types de filtres ACI, (i) les filtres plutôt passe-bande orientés, (ii) les filtres plutôt passe-bande anisotropes et (iii) les filtres anisotropes sur deux orientations orthogonales. La modélisation est proche du filtre pour le premier type, et fournira le motif d'orientation principal pour les deux autres types. Pour trois filtres pris en exemple dans chaque catégorie, la figure 3 illustre cela par leur réponse impulsionnelle, leur spectre d'énergie et le spectre d'énergie du filtre de Gabor le plus proche.

Dans tous les cas, cette modélisation a pour but, unique, une analyse statistique des quatre ensembles de filtres à mettre en relation avec la forme spectrale moyenne de chaque catégorie [15]. Cette étude met également en relief des similitudes avec les populations de cellules simples du cortex visuel [16]. On trouvera dans [12] l'analyse statistique complète. On reporte ici uniquement

les résultats obtenus sur la localisation des filtres pour les mettre en perspective avec l'application visée.

	R. impuls	Fourier	Gabor	R. impuls	Fourier	Gabor	R. impuls	Fourier	Gabor
Ville									
Intérieur									
«Ouvert»									
«Fermé»									

Figure 3 : illustration pour chacune des catégories de 3 filtres obtenues par ACI (réponse impulsionnelle – gauche-, spectre d'énergie –milieu-, spectre d'énergie du filtre de Gabor le plus proche –droite).

La figure 4 montre pour chaque ensemble de 30 filtres, la localisation spatiale dans le plan de Fourier, du filtre de Gabor le plus proche. On remarque qu'il y a un plus grand nombre de filtres sensibles aux orientations de référence ( $0^\circ$ ,  $90^\circ$ ), par rapport aux orientations obliques. Les scènes de la base étant des scènes naturelles, elles reproduisent notre environnement structuré plutôt horizontalement et verticalement. De plus, on remarque la cohérence entre la localisation des filtres et les propriétés spectrales des images de chaque catégorie (cf. paragraphe III). Pour les catégories « ville » et « intérieur », les localisations sont très voisines : on peut déjà penser que pour ces deux catégories, la discrimination sera plus difficile.

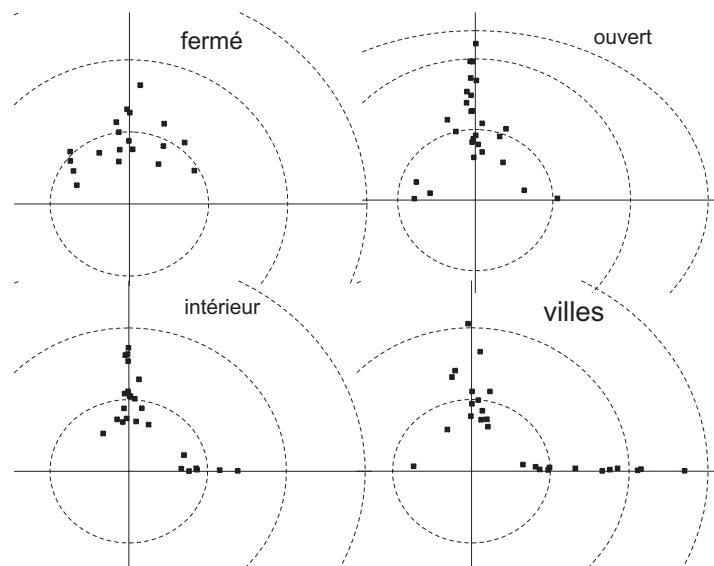


Figure 4 : localisation des fréquences centrales  $(u_0, v_0)$  des filtres dans le demi plan spectral supérieur. Graduations tous les 4 cycles par image.

## V.2 Propriétés de reconstruction et de codage épars.

Une propriété intéressante du codage par les primitives ACI est le caractère épars de celui-ci [5,16]. Cela se traduit pour les  $m$  codes représentant une image, par l'existence de peu de codes à valeur élevée par rapport à un nombre plus important de codes à valeur plus faible pouvant se négliger devant les premiers. Ainsi le caractère épars du code obtenu par la projection d'une image sur

une collection de filtres ACI, permet d'espérer sa reconstruction avec un nombre restreint de primitives. La figure 5a illustre la reconstruction d'une image complète à partir de sa décomposition dans la base des primitives ACI. La figure 5b illustre cette reconstruction en ne prenant que les filtres les plus actifs sur l'image (8 filtres sur 30).

L'écart quadratique moyen entre une image reconstruite avec toutes les primitives et celle reconstruite avec uniquement les primitives les plus actives est inférieur à 0,01. Ces reconstructions sont réalisées en appliquant l'équation (1), à partir d'une image découpée en bloc de taille  $32 \times 32$ . Les lieux de ces blocs sont espacés de façon régulière, avec plus ou moins de recouvrement. La valeur de maillage indiquée sur les figures 5 représente l'écart entre deux centres de blocs à reconstruire.

Les primitives obtenues par ACI étant des filtres de type passe-bande, la valeur moyenne de l'image ne pourra pas être restituée et on ne pourra reconstruire qu'une partie des fréquences spatiales de l'image. Les filtres ACI sont situés dans le plan de Fourier pour des fréquences spatiales moyennes jusqu'à environ 10 cycles par images (*cf.* figure 3). La reconstruction s'effectuera donc dans cette gamme de fréquences, avec une perte d'information en basses fréquences et par conséquent un rehaussement des informations de frontières entre régions.

Cette propriété est une voie d'investigation pour une représentation parcimonieuse des images avec les primitives les plus actives.

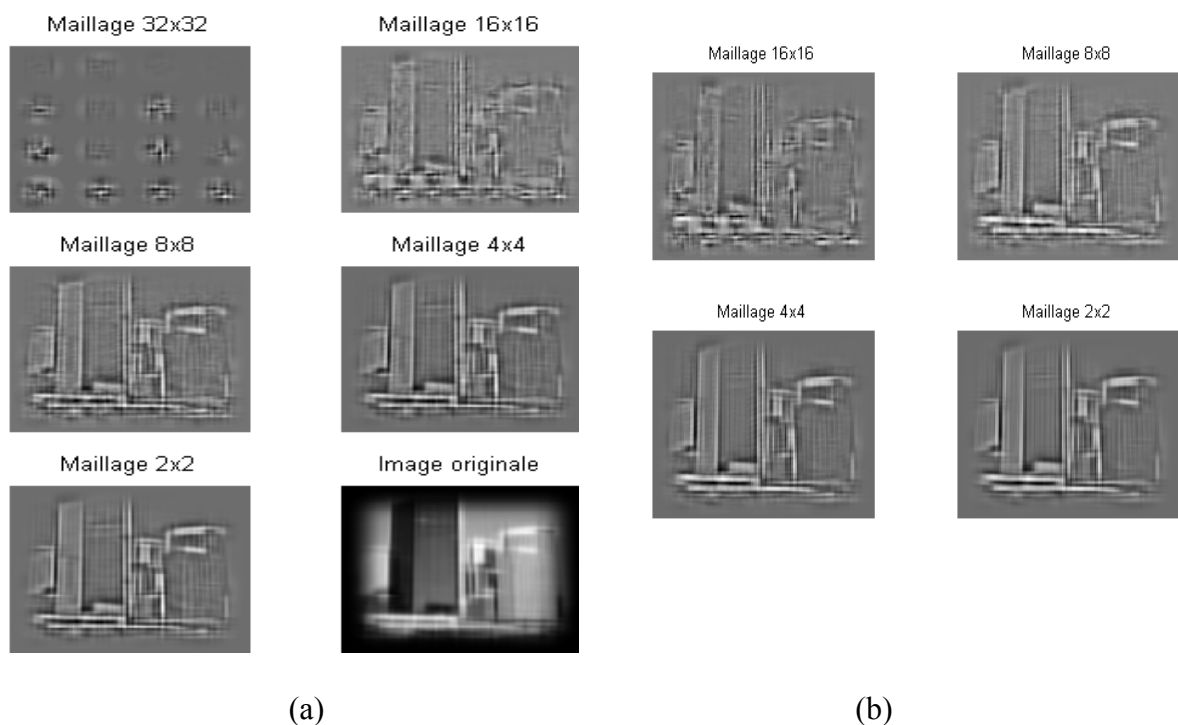


Figure 5 : (a) reconstruction d'une image avec les primitives ACI ; (b) idem avec seulement les détecteurs les plus actifs.

## VI. Signatures des images par les densités de probabilité des codes

Les images de la base sont filtrées par les quatre populations  $\phi$  de primitives. Les distributions des énergies qui en résultent constituent la signature des images. Une densité de probabilité multivariée ( $m = 30$ ) est ainsi associée à chaque image. De ce fait, la différence entre images est implantée comme une différence entre distributions.

### VI.1 Estimation des densités de probabilité multivariées

Considérons les notations suivantes. On appelle  $\Phi_v = \{\phi_{v_i}, i = 1..30\}$ ,  $\Phi_o = \{\phi_{o_i}, i = 1..30\}$ ,  $\Phi_i = \{\phi_i, i = 1..30\}$  et  $\Phi_f = \{\phi_{f_i}, i = 1..30\}$ , les ensembles de primitives issues respectivement des villes, des paysages ouverts, des scènes d'intérieur et des paysages fermés. Explicitons la suite des notations en prenant comme exemple la famille des filtres issus des scènes fermées. Soit par exemple  $I_{f_i}(x, y) = I(x, y) * \phi_{f_i}(x, y)$ , où \* désigne l'opérateur de convolution et  $\phi_{f_i} \in \Phi_f$ , l'image I filtrée par le i<sup>ème</sup> filtre de l'ensemble  $\Phi_f$  des filtres (de même pour les autres catégories). Sur l'image  $I_{f_i}(x, y)$ ,  $e_{f_i}(x, y) = [I_{f_i}(x, y)]^2$  définit une variable aléatoire d'énergie. Soit  $e_f$  la variable aléatoire multivariée de dimension 30 générée à partir de la projection de l'image sur les 30 filtres de  $\Phi_f$  ( $e_f \in \mathcal{R}^{30}$ ). Les composantes de  $e_f$  sont supposées indépendantes (de même que les composantes de  $e_v$ ,  $e_o$  et de  $e_i$ ), se traduisant par :

$$h_f(e_f) = h_f(e_{f_1}, \dots, e_{f_i}, \dots, e_{f_{30}}) = \prod_{i=1}^{30} h_{f_i}(e_{f_i}), \quad (2)$$

avec  $h_f$  la densité de probabilité de  $e_f$  et  $h_{f_i}$  la densité de probabilité marginale sur l'axe  $i$ .

### VI.2 Différence entre images

Pour mesurer la différence entre images utilisant les propriétés d'indépendance des codes des images représentées par les filtres ACI, le choix s'est naturellement porté sur la divergence de Kullback, version symétrique de l'information de Kullback. Notons par ailleurs que cette mesure de dissimilarité est souvent utilisée (de même que la distance du  $\chi_2$ ) comme technique d'appariement dans les systèmes de recherche d'images [17,19].

Soient deux images,  $I_k$  et  $I_l$ . Par définition de l'information de Kullback entre deux distributions [1] et avec (2), on a :

$$K(h_f^{(k)} \| h_f^{(l)}) = \int_{D_e} h_f^{(k)}(e) \cdot \log\left(\frac{h_f^{(k)}(e)}{h_f^{(l)}(e)}\right) \cdot de = \sum_{i=1}^{30} \int_{D_e} h_{f_i}^{(k)}(e) \log\left(\frac{h_{f_i}^{(k)}(e)}{h_{f_i}^{(l)}(e)}\right) de \quad (3)$$

La divergence de Kullback s'obtient donc par :

$$Div(h_f^{(k)}, h_f^{(l)}) = (K(h_f^{(k)} \| h_f^{(l)}) + K(h_f^{(l)} \| h_f^{(k)})). \quad (4)$$

Grâce à la propriété d'indépendance (équation (2)), la divergence s'évalue comme une somme de divergences sur les densités marginales monovariées :

$$Div(h_f^{(k)}, h_f^{(l)}) = \sum_{i=1}^{30} Div(h_{f_i}^{(k)}, h_{f_i}^{(l)}). \quad (5)$$

Cette formulation montre l'avantage ici de la propriété d'indépendance où une estimation en 30 dimensions se calcule par la sommation de 30 estimations élémentaires en 1 dimension. On contourne ainsi le problème de « l'espace vide » pour les estimations en grande dimension. Toutes les formulations précisées ici avec la famille de filtres  $\Phi_f$  se déclinent identiquement avec les trois familles de filtres  $\Phi_v$ ,  $\Phi_o$  et  $\Phi_i$ .

### VI.3 *Implantation*

Pour une image  $I_k$  donnée (taille 128 x 128), les densités de probabilité constituent la signature de cette image. Elles sont calculées (équation (2)) à l'aide d'une technique d'estimation par noyaux gaussiens [20] discrétisée sur 256 points. Afin de limiter le temps de calcul, tout en ayant une estimation fiable, les calculs de densité de probabilité sont effectués sur 50% des pixels de l'image (soit  $(128-31)^2 / 2 = 4852$  observations).

Les distances entre images, donc entre distributions sont estimées suivant les équations (3) et (4). Nous avons ainsi généré 4 matrices de Kullback 200x200 en projetant les images sur chacune des quatre bases  $\Phi$ . Notons  $D_f, D_v, D_i$  et  $D_o$  ces quatre matrices, calculées comme somme de matrices élémentaires  $D_{f_i}$  (équation (5)).

## VII. *Expérimentations en catégorisation*

La discrimination entre images suivant le contexte est réalisée par apprentissage supervisé. L'information utile est synthétisée dans la matrice de dissimilarités entre les images, calculée par la divergence de Kullback. Ainsi il n'y a pas de représentation euclidienne des images dans un espace de caractéristiques multidimensionnelles, et l'outil le plus simple de discrimination disponible pour cela, est alors le classifieur aux K plus proches voisins. Le paramètre K et la matrice de confusion sont estimés sur l'ensemble de la base d'images par validation croisée.

Préalablement à la catégorisation en utilisant l'ensemble des filtres extraits, nous avons cherché à ordonner les filtres selon leur pouvoir de discrimination. La démarche suivie est issue de l'analyse discriminante linéaire, où le pouvoir de discrimination est évalué par le rapport de l'inertie inter-classe sur l'inertie totale. Soit par exemple  $D_{f_{i^*}}$ , la matrice élémentaire de Kullback pour les images projetées uniquement sur le filtre  $i^*$  pris arbitrairement dans la population  $\Phi_f$ . Notons  $\tau_f(i^*)$  le pouvoir discriminant du filtre  $i^*$ . Ce critère classique en analyse discriminante linéaire peut également s'évaluer ici par la connaissance de la matrice de dissimilarités. En toute rigueur, il s'agira en fait de « pseudo-inertie » puisque les mesures utilisées ne sont pas des distances euclidiennes. On appelle  $IT_{f_{i^*}}$ , la pseudo-inertie totale des images projetées sur le filtre  $i^*$  de  $\Phi_f$  et  $IW_{f_{i^*}}$ , la pseudo-inertie intra-classe. Par analogie avec les distances euclidiennes et en développant le calcul de l'inertie totale, on a :

$$IT_{f_{i^*}} = \frac{1}{2 \cdot N^2} \sum_{k=1}^N \sum_{l=1}^N [D_{f_{i^*}}(k, l)]^2, \quad \text{et} \quad IW_{f_{i^*}} = \sum_{c=1}^4 q_c \cdot IT_{f_{i^*}}(c), \quad (6)$$

avec  $IT_{f_{i^*}}(c)$ , la pseudo-inertie totale restreinte à une catégorie  $c$  donnée d'images et  $q_c$  le poids de chaque catégorie (ici =1/4). Le pseudo-pouvoir discriminant s'évalue alors par :

$$\tau_f(i^*) = 1 - \frac{IW_{f_{i^*}}}{IT_{f_{i^*}}}. \quad (7)$$

Dans chaque population  $\Phi$ , les filtres sont classés par ordre décroissant de leur pouvoir de discrimination (cf. figure 6a). Les performances en catégorisation sont analysées en cumulant progressivement les matrices élémentaires associées aux filtres ainsi triés. La figure 6b illustre



l'évolution du taux de reconnaissance en fonction du nombre de filtres ainsi ordonnés pour le calcul cumulatif de la matrice de Kullback selon l'équation (5).

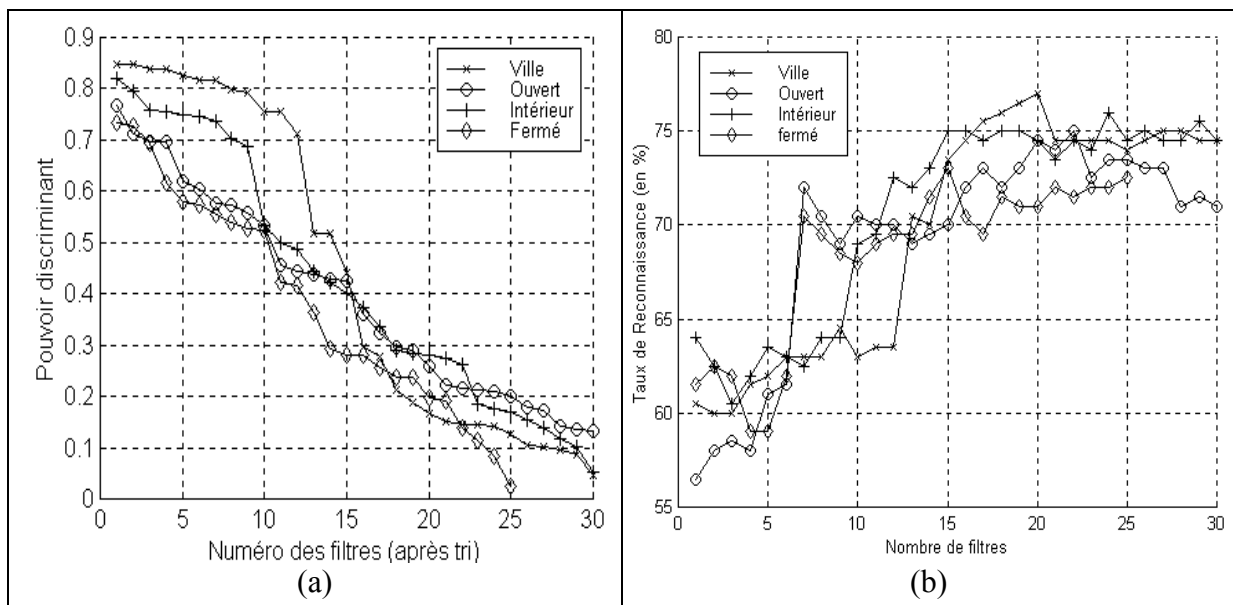


Figure 6 : (a) évolution du pouvoir discriminant des filtres ACI ordonnés, (b) évolution du taux de reconnaissance suivant le nombre de filtres ordonnés et utilisés pour la discrimination

Ces deux courbes montrent que les populations de filtres issues des villes et scènes d'intérieur sont les plus aptes à catégoriser la base d'images. Ces deux populations de filtres sont en fait assez voisines, car elles sont issues de deux classes d'images ayant des propriétés spectrales proches, comme nous l'avons signalé au paragraphe V.1. De plus, on remarque qu'à partir de vingt filtres les informations détectées n'ont plus d'apport significatif dans la discrimination. Le taux de reconnaissance maximal est de 77 % avec la matrice de confusion obtenue avec les vingt filtres de ville les plus discriminants :

%	Ville	Ouvert	Intérieur	Fermé
Ville	84	0	12	4
Ouvert	0	74	14	12
Intérieur	10	4	78	8
Fermé	0	20	8	72

### VIII. Conclusions

Nous avons montré comment des filtres obtenus par analyse en composantes indépendantes pouvaient simplement catégoriser une base d'images. La principale originalité de cette étude par rapport aux approches de type « caractérisation par banc de filtres » est que les filtres utilisés ici sont extraits directement des catégories que l'on cherche à discriminer. Ils ne sont pas imposés a priori pour couvrir tout l'espace de Fourier. Il y a une relation forte entre les caractéristiques fréquentielles des détecteurs ainsi extraits et celles des images dont ils sont issus. Il s'agit donc d'une approche flexible pouvant se généraliser naturellement pour s'adapter à d'autres classes d'images.

Suivant cette approche, l'utilisation de l'ACI est fondamentale par la propriété d'indépendance qui rend plus robuste le processus aval d'estimation des signatures par les densités de probabilité marginales. Le problème de « l'espace vide » en grande dimension est ainsi contourné en se ramenant à des estimations monodimensionnelles. Une autre conséquence de cette propriété d'indépendance est la justification naturelle d'un modèle simple additif pour l'établissement d'une mesure de dissimilarité entre images. Si ce modèle est souvent utilisé de par

sa simplicité dans les systèmes de reconnaissance, il est ici justifié théoriquement par la propriété d'indépendance statistique des codes.

Enfin, cette approche se met en perspective avec les modèles de perception visuelle par la similitude entre les détecteurs extraits par ACI et les cellules du cortex visuel. Ce travail se poursuivra (i) en étendant cette approche à des bases d'images plus complexes, (ii) en intégrant l'information de chrominance et (iii) en tenant compte de l'organisation spatiale dans les images par des statistiques locales, pour mettre en place des stratégies de type *Coarse to Fine*.

## Références

- [1] Basseville M., (1989) Distance measures for signal processing and pattern recognition, *Signal Processing*, vol. 18, pp. 349-369.
- [2] Barlow, H.B. (1989) Unsupervised Learning. *Neural Computation*, vol. 1, pp. 295-315.
- [3] Bell A.J, Sejnowsky T.J. (1997) The Independent Component of Natural Scenes are Edge Filter, *Vision Research*, vol 37, n° 23, pp 3327-3338.
- [4] Comon (1994) Independent Component Analysis – a new concept ?, *Signal processing*, vol 36, pp 287-314.
- [5] Donoho D.L. (2000) Nature vs. Math : Interpreting Independent Component Analysis in light of computational harmonic analysis, ICA'2000, Helsinki, Finlande, juin 2000, pp. 459-470.
- [6] Field, D.J. (1994). What is the Goal of Sensory Coding ?, *Neural Computation*, vol. 6, pp. 559-601.
- [7] Gorkani M.M., Picard R.W., (1994) Texture Orientation for Sorting Photos at Glance, *IEEE conference on Pattern Recognition*, octobre 1994.
- [8] Guérin-Dugué A., Oliva A., (2000) Classification of Scene Photographs from Local Orientations Features, *Pattern Recognition Letters*, sous presse, 2000.
- [9] Hyvärinen A., Oja.E. (1997) A Fast fixed-point algorithm for Independent Component Analysis, *Neural Computation*, vol 9, no 7, pp. 1483-1492.
- [10] Jutten C., Héroult J. (1991) Blind separation of sources : an adaptive algorithm based on neuromimetic architecture, *Signal processing*, vol. 24, pp. 1-10.
- [11] Labbi A., Bosch H., Pellegrini, Ch. (1999). Image Categorization using Independent Component Analysis, *Workshop on Biologically Inspired Machine Learning, BIML'99*, 14 juillet (conférence invitée), Crète, Grèce.
- [12] Le Borgne H., Guérin-Dugué A.(2000) Propriétés des détecteurs corticaux extraits des Scènes Naturelles par Analyse en Composantes Principales, disponible sur [http://www.supelec-rennes.fr/acth/valgo/Valgo\\_Numero-01-01.html#LeBorgne2001](http://www.supelec-rennes.fr/acth/valgo/Valgo_Numero-01-01.html#LeBorgne2001), revue Valgo, 2001.
- [13] Li, Z., Attick, J.J. (1994) . Toward a Theory of Striate Cortex, *Neural Computation*, vol. 6, pp. 127-146.
- [14] Oliva A., Schyns P.G. (1997) Coarse blobs or fine edges ?, *Cognitive Psychology*, vol. 34, pp. 72-102.
- [15] Oliva A., Torralba A.B., Guérin-Dugué A., Héroult J., (1999) Super-Ordinate representation of scenes from power spectrum shapes, *CIR-99, The challenge of image retrieval*, Newcastle, Royaume-Uni, mars 1999.
- [16] Olshausen B.A, Field D.J (1997) Sparse coding with an overcomplete basis set : a strategy employed by V1 ?, *Vision Research*, vol 37, n° 23, pp. 3311-3325.
- [17] Pusicha J., Bushmann J.M., Rubner Y., et Tomasi C. (1999) Empirical Evaluation of Dissimilarity Measures for Color and Texture. *ICCV'99*, pages 1165-1172, Kerkyra, Grèce, 1999.
- [18] Rogowitz B.E., Frese T., Smith J.R., Bouman C.A., Kalin E., (1998) Perceptual Image Similarity Experiments, in *Human Vision and Electronic Imaging III*, SPIE, n° 3299, San Jose, Etat-Unis, janvier 1998.
- [19] Rubner Y., Tomasi C., Guibas L.F., (1998) A metric for distribution with application to images databases, *ICCV, Bombay, Inde*, pp. 59-66, janvier 1998.
- [20] Silverman B.W, (1986) *Density Estimation for Statistics and Data Analysis*, New York: Chapman and Hall.
- [21] Szummer M., Picard R.W., (1998) Indoor-Outdoor image classification, *IEEE Int. Workshop on Content-Based Access of Image and Video Database in conj. with ICCV'98*, janvier 1998.
- [22] Vailaya A., Jain A., Zhang H.J., (1998) On image classification : City images vs Landscapes, *Pattern Recognition*, vol. 31, n°12, pp. 1921-1935, 1998.
- [23] Van Hateren J.H, Van der Schaaf A.(1998) Independent component filters of natural images compared with simple cells in visual cortex, *Proc. R. Soc. London*, B265, pp. 359-366.