
Nouveaux résultats en classification à l'aide d'un codage par motifs fréquents

Sébastien Jouteau¹, Antoine Cornuéjols², Michèle Sebag², Philippe Tarroux³ et Jean-Sylvain Liénard³

(1)
Institut d'Informatique d'Entreprise
18, Allée Jean Rostand
F-91425 Evry
jouteau@iie.cnam.fr

(2) Équipe Inférence et Apprentissage
Laboratoire de Recherche
en Informatique (L.R.I.)
Bât 490, Université Paris-Sud
F-91405 Orsay Cedex
(antoine.sebag)@lri.fr

(3) Équipe Perception Située
Laboratoire d'Informatique
pour la Mécanique et les Sciences
de l'Ingénieur (L.I.M.S.I.)
Bât 508, Université Paris-Sud
F-91405 Orsay Cedex
(philippe.tarroux.lienard)@limsi.fr

RÉSUMÉ. Le nombre nécessairement limité de données ne permet pas de distinguer sûrement entre les régularités réelles et les régularités accidentelles lorsque les données sont décrites par un grand nombre d'attributs. Nous présentons ici une méthode radicalement nouvelle de recodage des données qui consiste à chercher des motifs fréquents (avec un taux de couverture faible) dans les données et à les utiliser pour coder les entrées. Cette méthode a été testée dans le cadre d'une application de classification de scènes naturelles. Une partie des images disponibles a été utilisée pour identifier 1000 motifs fréquents, grâce à un algorithme spécialement adapté. Ces motifs ont ensuite été employés pour coder les images. La classification des images restantes a alors été réalisée par une méthode de classification par plus proches voisins. Les résultats obtenus sur cette tâche sont très nettement supérieurs à ceux obtenus jusque là par une méthode neuronale à base de fonctions radiales.

ABSTRACT. The limited number of data points usually available does not allow us to distinguish with certainty true underlying regularities from spurious ones when the data are described using numerous attributes. In this paper, we present a new coding method where low coverage frequent item sets discovered in the training sample are used to encode the data. This method has been used in a classification task of natural scenes. Thanks to an adapted algorithm, images in the training sets have been used to find 1000 frequent items that were subsequently used to encode all the images. The classification of test images was then performed using a nearest neighbour method. The results obtained in this way far outperformed previous results arrived at with neural networks using radial basis functions.

MOTS-CLÉS : Codage, motifs fréquents

KEYWORDS: Coding, Frequent Item Sets

1. Introduction

Les données relatives à des mesures réelles sont souvent difficiles à analyser directement car elles ne sont pas décrites de manière à faire apparaître les régularités sous-jacentes pertinentes. C'est le cas lorsque l'on cherche à analyser des données décrites par un très grand nombre d'attributs et dont on soupçonne qu'elles correspondent de fait à un phénomène descriptible dans un sous-espace très limité de l'espace de description initial. Les techniques d'apprentissage artificiel et d'analyse des données qui visent à découvrir les caractéristiques de ces sous-espaces sont malheureusement souvent incapables de distinguer sûrement entre les régularités réelles et les régularités accidentelles lorsque les données d'apprentissage sont en nombre limité. C'est pourquoi on a fréquemment recours à une phase de redescription des données visant à mettre en évidence des descripteurs pertinents. Classiquement, les représentations des données cherchées (on emploiera aussi le terme de codage), doivent vérifier un certain nombre de contraintes. Parmi elles :

- L'**orthogonalité** des primitives de représentation. Cela permet en effet d'assurer une non redondance de l'information codée et, sous certaines conditions, l'unicité de la représentation obtenue. La décomposition en séries de Fourier en est un exemple.

- La **capacité d'approximation** du système de représentation. Il s'agit de pouvoir reconstruire fidèlement les données d'origine à partir de leur codage. Les mathématiciens parlent de *représentation dense presque partout* pour signifier que l'on peut reconstituer n'importe quelle information aussi précisément que souhaité. On peut ainsi reconstruire une fonction à partir de sa décomposition de Fourier.

- La **robustesse** du système de codage. Il est important qu'une petite erreur sur le codage ne conduise pas à une grosse erreur sur l'information codée.

- La **pertinence des primitives** de représentation vis-à-vis de l'objectif poursuivi. Une représentation en fréquences d'un signal périodique facilite par exemple certaines opérations sur les signaux. Il est courant que la recherche de primitives pertinentes conduise à une *réduction de l'information codée* par élimination de redondances et de variations accidentelles.

- Souvent, le système de codage est recherché indépendamment des données, mais résulte d'un choix préalable. C'est le cas par exemple de la décomposition en série de Fourier. Inversement, l'analyse en composantes principales (ACP) détermine les primitives de représentation (les axes d'inertie principaux) à partir des données.

L'objectif du codage est de rendre (plus) explicite la structure présente dans le phénomène étudié et qui est cachée sous la forme de dépendances statistiques complexes. Barlow ([BAR 61, BAR 89]) a défendu l'idée selon laquelle un prétraitement des données qui expliquerait celles-ci sous la forme de collections d'événements indépendants serait utile. Une méthode d'analyse des données récemment développée ([CAR 98, HYV 01]), l'*analyse en composantes indépendantes* (ACI), précise cette idée en cherchant un ensemble de fonctions de base tel que les données puissent s'exprimer comme combinaison linéaire de ces fonctions de base. Lorsque le nombre de fonctions de base cherchées est grand par rapport à la dimension effective des don-

nées, on obtient une base « surcomplète ». Le codage obtenu peut alors être clairsemé (*sparse coding*), c'est-à-dire que chaque donnée est décrite par un petit nombre de descripteurs parmi l'ensemble des descripteurs (fonctions de base) possibles. On peut montrer de plus qu'un codage clairsemé a des propriétés de robustesse car des variations limitées des fonctions de base peuvent être compensées par des variations limitées et locales des coefficients du codage des formes décrites. L'emploi de bases non surcomplètes, en revanche, conduit à des variations globales des coefficients de codage lorsque les fonctions de base sont modifiées. De telles instabilités sont indésirables dans des applications de reconnaissance des formes.

Pour résumer, l'analyse en composantes indépendantes produit un codage dont les caractéristiques sont les suivantes : les primitives (fonctions de base) peuvent être *non orthogonales*, la représentation obtenue *permet la reconstruction* du signal d'origine, le système de *codage dépend des données* car les fonctions de base ne sont pas déterminées *a priori*. Par ailleurs, le codage possède la propriété de robustesse.

En pratique malheureusement, l'analyse en composantes indépendantes ne peut se réaliser pour des espaces d'entrée dont le nombre de dimensions excède la centaine environ (voir le chap. 14 de [HYV 01]). Elle ne peut donc être appliquée pour des données décrites par de très nombreux attributs, comme par exemple les images qui impliquent des milliers de pixels ou les puces ADN mesurant l'activité de milliers de gènes, deux applications qui nous intéressent.

Si l'approche par l'analyse en composantes indépendantes d'un codage clairsemé est impossible, une approche directe est-elle envisageable ?

En faisant l'hypothèse que les données résultent d'une somme de formes latentes, et que ces formes pour être intéressantes doivent figurer suffisamment souvent dans les données, le rapprochement avec la technique de recherche de motifs fréquents s'impose. Les données étant décrites par un ensemble de descripteurs (attributs-valeur), on cherche les conjonctions d'attributs-valeur présentes dans un certain pourcentage des données. Grâce à certaines contraintes, on peut guider la recherche de tels motifs de manière à favoriser la découverte d'un ensemble de primitives (les motifs) permettant un codage clairsemé des données. (Voir la section 2.3 ci-dessous).

À ces contraintes peuvent s'ajouter des critères supplémentaires liées au domaine d'application permettant de sélectionner les motifs fréquents les plus intéressants.

Ce papier décrit l'adaptation de la technique de recherche de motifs fréquents à la recherche directe d'un codage clairsemé dans un espace de grande dimension et son application à une tâche de classification de scènes naturelles. La section 2 fournit les notations, les définitions et les principes de l'approche. L'algorithme développé est décrit en section 3.2. Les primitives ainsi trouvées servent à coder les entrées. L'algorithme de classification fait l'objet de la section 4. On montre dans la section 5 les résultats obtenus sur la tâche de classification de scènes naturelles et la comparaison avec d'autres techniques. Les perspectives sont soulignées dans la section 6.



Figure 1. Échantillon d'images utilisées dans cette étude. Noms des classes : avions (1), plats (2), Utah (3), minéraux (4), chiens (5), poissons (6), verres (7), papillons (8), porcelaines (9), figurines (10), voitures (11), fleurs (12). (Cette figure est reprise de [DEN 02]).

2. Le principe de l'approche

2.1. Une tâche d'apprentissage supervisé

Si les problèmes de vision par ordinateur dans les univers artificiels commencent à être bien connus, la vision dans les univers naturels n'en est encore qu'au stade de recherches préliminaires. Les caractéristiques des scènes naturelles sont en effet notablement différentes de celles des scènes en milieu artificiel et nécessitent de faire appel à de nouvelles approches. Le fait que de plus en plus de robots sont destinés à évoluer en milieu naturel rend urgent des recherches dans ce domaine. C'est dans ce cadre que se situe la tâche étudiée ici.

Le problème consiste à apprendre à reconnaître des images de scènes naturelles classées en 12 catégories (voir la figure 1). Ces images proviennent de la base COREL (http://www.corel.com/gallery_line/). Les images sont redécrites par 128×128 pixels en 128 niveaux de gris. Pour ces expériences, la base utilisée comportait 1082 images réparties également entre les 12 classes. Nous considérerons dans la suite que chaque pixel est un attribut pouvant prendre une valeur parmi 128. La dimension de l'espace d'entrée est donc dans ce cas de $32768 = 128 \times 128$.

2.2. Définitions

Les *fonctions de base* cherchées sont des conjonctions d'attributs-valeur, on parlera aussi d'atomes. Dans le cas présent, il s'agira de collections de pixels, chacun de

ceux-ci étant associé à un niveau de gris donné. Une fonction de base est ainsi une fonction booléenne prenant la valeur *vrai* si la collection d'atomes (pixel = valeur) correspondante est satisfaite dans l'image étudiée. On dira que le *support* d'une fonction de base est de $\varepsilon\%$ si cette fonction prend la valeur *vrai* dans $\varepsilon\%$ des images testées. Réciproquement, on appellera *code* d'une forme d'entrée (ici une image), l'ensemble des fonctions de base vérifiant cette forme.

2.3. Propriétés des fonctions de base recherchées

Dans ce nouveau cadre, les propriétés désirées du système de codage se traduisent comme suit :

- 1) **Représentativité.** Chaque fonction de base a un support supérieur à $\varepsilon\%$. Elle est donc suffisamment représentée dans la base d'exemples pour être utile.
- 2) **Parcimonie.** Peu de fonctions de base sont vérifiées par un exemple (e.g. image). On vérifie ainsi l'une des propriétés du codage clairsemé.
- 3) **Suffisance.** Tout exemple rend vrai un nombre minimal de fonctions de base.
- 4) **Orthogonalité.** Pour chaque paire de fonctions de base, l'intersection des exemples qui rendent *vrai* l'une et l'autre est réduite. Les exemples sont donc décrits par des codes différents.

Nous avons adapté la méthode de recherche de motifs fréquents dans une base de données pour chercher un codage tendant à vérifier les propriétés ci-dessus.

3. Algorithme de recodage non supervisé

La recherche de motifs fréquents de taux de couverture faible dans des données décrites par de nombreux attributs ne peut s'effectuer sans précautions. C'est pourquoi nous présentons rapidement l'algorithme développé à cet effet.

3.1. Échec d'une approche naïve

La première approche consiste à utiliser un algorithme existant de recherche de motifs fréquents, tel qu'APRIORI [AGR 94], pour chercher tous les motifs ayant un certain taux de couverture, puis ensuite à sélectionner parmi eux ceux qui satisfont les contraintes soulignées dans la section 2. Malheureusement cette approche s'est révélée impraticable pour la reconnaissance d'images car le nombre de motifs fréquents croît exponentiellement avec la taille des motifs. Ainsi, la table ci-dessous montre le nombre de motifs trouvés pour des images de taille 32x32 en 64 niveaux de gris (où *m* signifie mille, *M* million et *MM* milliard) :

Nb atomes par motif	1	2	3	4	5	6
Nb motifs trouvés	2m	110m	3,8M	80M	1,15MM	12,5MM

Il faut donc renoncer à une méthode de recherche exhaustive des motifs.

3.2. Approche randomisée

Nous avons développé une méthode de construction incrémentale de motifs fréquents par ajouts successifs d'atomes (ici, de pixels d'un certain niveau de gris) en les sélectionnant à chaque pas afin que le motif en construction satisfasse aux critères désirés. Par ailleurs, l'exploration des motifs fréquents est maintenant guidée mais non exhaustive et stochastique. Des essais successifs peuvent donc produire des bases de motifs différentes. (Voir la description de l'algorithme).

Algorithme 1 Recherche itérative et stochastique de motifs fréquents.

Paramètres : taux de couverture $\varepsilon\%$. Nombre de motifs cherchés = N .

Nombre de motifs trouvés = $n \leftarrow 0$.

while $n \leq N$ **do**

 Choix dans un exemple x_i encore peu couvert, d'un premier atome a_0 présent dans au moins $\varepsilon\%$ des exemples.

motif $\leftarrow a_0$

while Taux de couverture de *motif* $> \varepsilon\%$ **do**

 Tirer au hasard un atome a de x_i couvrant au moins $\varepsilon\%$ des exemples et peu utilisé dans les motifs existants et satisfaisant des contraintes additionnelles sémantiques (voir section 5).

if *motif* + a couvre au moins $\varepsilon\%$ des exemples **then**

motif \leftarrow *motif* + a

end if

end while

end while

4. Exploitation pour l'apprentissage supervisé

Une fois une base de N fonctions de base trouvée sur un ensemble d'apprentissage, les exemples sont recodés dans cette base. Chaque exemple devient un vecteur de N booléens prenant la valeur *vrai* ou *faux* selon que la fonction de base correspondante couvre l'exemple ou non. Nous avons également utilisé une formule d'appariement plus souple utilisant une fonction sigmoïde à valeur dans $[0, 1]$ qui tient compte du nombre d'atomes de la fonction de base qui couvrent l'exemple, et donnant un recodage dans $[0, 1]^N$ au lieu de $\{0, 1\}^N$.

Dans le nouvel espace d'exemples ainsi construit, il est possible d'utiliser n'importe quelle méthode d'apprentissage supervisé. Dans les expériences rapportées ici, nous avons utilisé une méthode de classification par plus proche voisin. Les exemples d'apprentissage utilisés pour la recherche de fonctions de base sont également employés comme exemples étiquetés servant à la classification des exemples testés.

5. Les résultats expérimentaux

La méthode développée a été testée sur la tâche de classification de scènes naturelles décrite en section 1. Elle implique deux phases : d'abord une étape de détermination d'une base de fonctions de base permettant de redécrire les données, ensuite l'emploi du système de codage ainsi obtenu pour classer de nouvelles formes.

5.1. Les fonctions de base obtenues par recherche de motifs fréquents

Pour l'application étudiée, nous avons fixé à 1000 le nombre de fonctions de base recherchées. Plusieurs bases ont été obtenues en faisant varier les paramètres suivants :

– *Taux de couverture* : 1%, 2%, 5% et 10%

– *Critère sémantique additionnel*. En tenant compte du domaine étudié, ici l'analyse d'images, nous avons introduit des contraintes supplémentaires sur la construction des fonctions de base afin de tester des équivalences possibles avec d'autres types de codages classiques en traitement d'images. Quatre conditions ont été testées :

1) Aucune contrainte.

2) Les fonctions de base doivent correspondre à des *régions connexes* sur l'image : un nouveau pixel n'est ajouté à la fonction de base courante que si il est contigu à un pixel déjà sélectionné.

3) Les fonctions de base doivent correspondre à des *lignes* de l'image (régions de dimension 1). L'idée ici est de voir si l'on peut forcer le système de codage à retenir des contours dans l'image.

4) Les fonctions de base doivent correspondre à des *lignes raisonnables* de l'image, c'est-à-dire plus contraintes dans les changements de directions possibles. Cette contrainte plus forte que la précédente a été imposée lorsqu'il s'est avéré que la précédente produisait des « vermiciaux » remplissant des régions et non pas des lignes.

Environ la moitié des images de la base initiale de 1082 images, soit 500, ont été utilisées pour le calcul des fonctions de base. (Note : notre algorithme calcule une base de 1000 motifs en quelques minutes sur un PC équipé d'un Pentium II à 266 Mhz et 384 Mo de RAM).

Les figures 2 et 3 permettent de se figurer le type de fonctions de base obtenues pour certaines conditions expérimentales. Des résultats plus complets sont disponibles sur le site <http://www.lri.fr/~antoine/Research/FISICA/egc-03.html>.

L'histogramme présenté dans la figure 4 permet de contrôler l'orthogonalité des fonctions de base obtenues. Ces fonctions de base sont orthogonales lorsqu'elles sont rarement vérifiées par les mêmes images. La figure montre que les différentes bases obtenues pour des conditions différentes peuvent effectivement être considérées comme orthogonales.

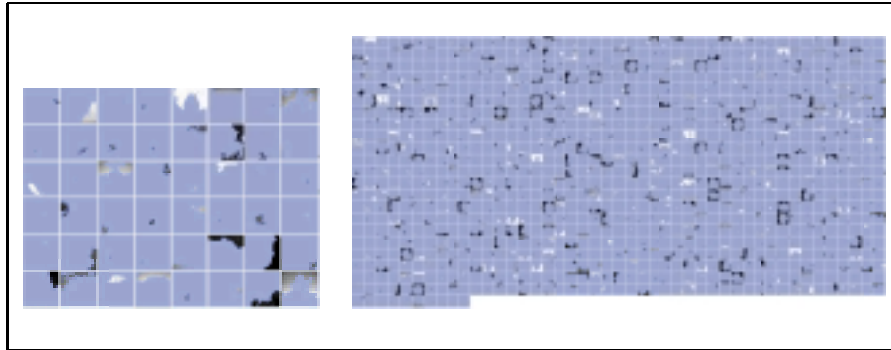


Figure 2. À gauche, le détail de quelques unes des fonctions de base obtenues sur des images 64×64 en 16 niveaux de gris avec un taux de couverture de 1%, et en cherchant des régions connexes de l'image. Dans les images accessibles sur le site internet, le fond bleu correspond aux zones qui ne font pas partie des fonctions de base, tandis que les fonctions de base sont figurés par des pixels de niveaux de gris variés. Ici, il faut se figurer que le gris moyen est le fond. À droite, figurent une base de 1000 fonctions de base.

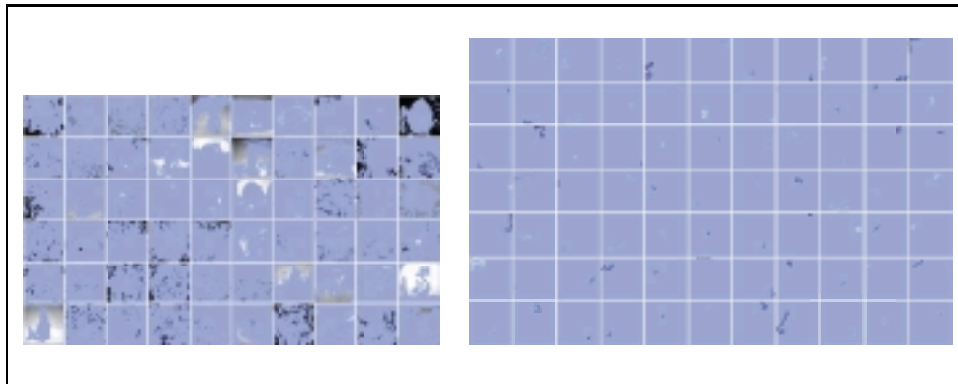


Figure 3. À gauche, des exemples de fonctions de base obtenues avec un taux de couverture de 1%, et sans contrainte. À droite, des exemples de fonctions de base obtenues avec un taux de couverture de 1%, et sous contrainte de linéarité raisonnable. L'examen des motifs trouvés montre qu'ils ne correspondent pas à des contours des images de la base d'exemples.

Inversement, l'histogramme de la figure 5 indique le nombre de fonctions de base qui sont vérifiées par les images. On constate que ce nombre varie autour d'une dizaine, ce qui traduit bien que le codage obtenu est clairsemé.

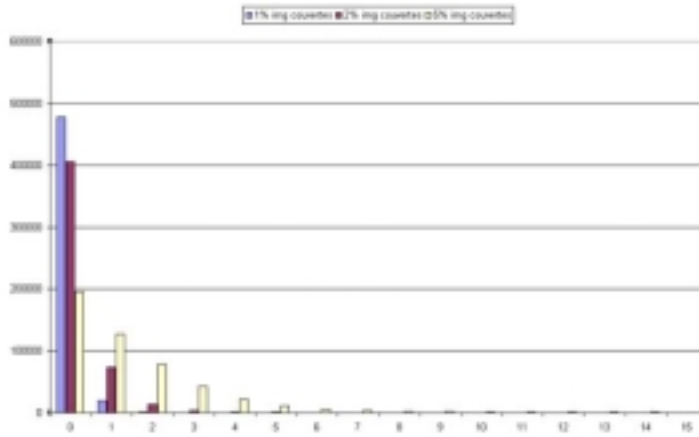


Figure 4. Histogramme représentant le nombre y de couples de fonctions (en ordonnée) ayant x images en commun (en abscisse). Ces fonctions ont été calculées à partir d’images de taille 64×64 en 16 niveaux de gris et sous la contrainte de connexité. Les résultats sont présentés pour des fonctions de base de taux de couverture 1%, 2% et 5%. Les fonctions de base obtenues pour $\varepsilon = 1\%$ sont les plus orthogonales entre elles.

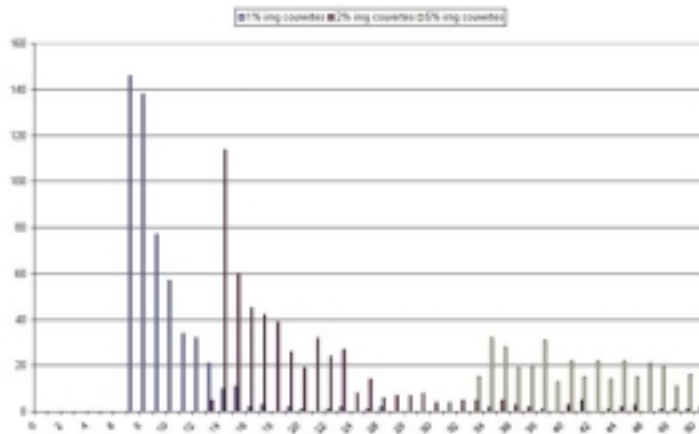


Figure 5. Histogramme représentant le nombre y d’images (en ordonnée) activant x motifs (fonctions de base) en abscisse. Ces fonctions ont été calculées à partir d’images de taille 64×64 en 16 niveaux de gris et sous la contrainte de connexité. Les résultats sont présentés pour des fonctions de base de taux de couverture 1%, 2% et 5%. Plus le taux de couverture est élevé, plus chaque exemple est couvert en moyenne par un nombre élevé de fonctions. On peut ainsi régler la parcimonie de la représentation et donc son caractère clairsemé.

5.2. Les résultats en classification

Les performances en classification ont été calculées sur les 582 images non utilisées pour déterminer la base des fonctions de base. Toutes les images sont recodées en utilisant les fonctions de base obtenues, et s'expriment donc sous la forme d'un vecteur de 1000 valeurs booléennes (en fait, dans certaines expériences, cette valeur booléenne était remplacée par une mesure plus continue d'appariement de l'image avec une fonction de base). Les 500 images employées pour la détermination des fonctions de base sont également utilisées comme base d'exemples étiquetées. Les images à classer sont alors étiquetées en utilisant une méthode de plus proche voisin. Dans les expériences rapportées ici, la distance utilisée est la distance L_1 .

La table 5.2 fournit les résultats obtenus avec une base de 1000 fonctions de base de taux de couverture de 5% soumis à la contrainte de connexité. Quoique les différents nombres puissent varier sensiblement, on observe en général que les résultats obtenus dans une grande variété de conditions sont assez similaires en moyenne. Des résultats plus complets sont accessibles à l'url :

<http://www.lri.fr/~antoine/Research/FISICA/egc-03.html>.

	Av	Pl	Ut	Mi	Ch	Po	Ve	Pa	Por	Fi	Vo	Fl
Avi	67%	2%	-	-	2%	2%	10%	10%	4%	2%	-	-
Pla	-	21%	-	2%	7%	19%	10%	12%	5%	-	19%	5%
Uta	17%	-	33%	-	7%	-	-	3%	10%	10%	13%	7%
Min	-	-	-	100%	-	-	-	-	-	-	-	-
Chi	26%	5%	7%	-	14%	9%	12%	9%	5%	2%	12%	-
Poi	5%	13%	3%	8%	-	13%	18%	21%	-	3%	10%	8%
Ver	2%	2%	-	-	10%	7%	43%	-	21%	5%	7%	2%
Pap	6%	6%	-	-	2%	14%	14%	35%	6%	-	12%	4%
Por	2%	2%	-	-	-	2%	-	12%	70%	10%	-	2%
Fig	-	-	-	-	-	-	6%	-	24%	70%	-	-
Voi	21%	6%	-	-	4%	4%	8%	4%	4%	29%	19%	-
Fle	2%	9%	-	-	-	9%	21%	14%	-	-	16%	28%

Tableau 1. Matrice de confusion obtenue avec des fonctions de base de taux de couverture $\epsilon = 5\%$ sous contrainte de connexité, en utilisant une formule d'appariement continu entre les images et les fonctions de base.

5.3. Analyse comparative

Parmi les travaux portant sur la reconnaissance de scènes naturelles, nous nous référons ici à des travaux parallèles de Nathalie Denquive et Philippe Tarroux [DEN 02] concernant la même tâche de classification que celle rapportée ici. Leur approche s'appuie sur les études de Field [FIE 94] analysant le système visuel des mammifères et mettant en avant le rôle possible d'un codage par analyse en composantes indépendantes grâce aux ondelettes de Gabor. Plusieurs travaux récents semblent indiquer que les fréquences les plus basses contenues dans l'image d'une scène suffisent à pouvoir la catégoriser rapidement [THO 96]. Les images de la base de 1082 images ont été

codées en utilisant un banc d'ondelettes de Gabor ajustées selon quatre orientations spatiales (horizontale, verticale et diagonales) et cinq fréquences couvrant quatre octaves. Ce filtrage, suivi du calcul d'une énergie moyenne, permet de coder chaque image par vingt nombres.

Ces vecteurs sont alors utilisés en entrée d'un réseau de neurones à fonctions radiales gaussiennes avec une couche de sortie pour la classification. Les expériences ont été conduites en faisant varier le nombre de fonctions radiales (de 20 à 300), ce qui correspond à la recherche d'un codage par recombinaison des informations fréquentielles de bas niveau, et en testant des méthodes de classification binaire (un neurone en sortie, soit une classe contre les autres) ou de classification multi-classes (12 neurones de sortie, un pour chaque classe).

Dans les expériences, 537 exemples ont été utilisés pour l'apprentissage, les 545 restants servant à la mesure de performance en généralisation, avec une méthode de validation croisée en répétant plusieurs fois l'expérience avec des tirages des exemples différents pour l'apprentissage et le test.

Les résultats obtenus montrent que le nombre optimal de fonctions gaussiennes est de l'ordre de 200. La table 5.3 montre la matrice de confusion obtenue avec 200 fonctions gaussiennes et une classification multi-classes. On notera que les taux de généralisation correcte varient entre 0% (pour la classe des plats (2)) à 50.7% (pour la classe des avions (1)) avec une moyenne de 23.6%.

	Av	Pl	Ut	Mi	Ch	Po	Ve	Pa	Por	Fi	Vo	Fl	Rj
Avi	50.7	-	3.3	-	-	-	-	-	-	-	1.3	-	44.7
Pla	-	-	-	6.3	0.8	-	-	-	-	-	-	1.6	91.3
Uta	1.1	-	23.3	-	3.3	1.1	-	-	2.2	-	-	-	68.9
Min	0.8	-	0.8	28.8	-	2.3	-	0.8	-	-	-	3.0	63.6
Chi	-	-	4.0	0.8	11.1	2.4	-	0.8	-	-	3.2	0.8	77
Poi	-	-	3.7	2.2	-	0.7	-	-	-	-	0.7	0.7	91.9
Ver	-	-	2.9	-	0.7	-	9.4	-	20.3	15.9	-	-	50.7
Pap	-	-	-	7.3	-	1.3	-	13.3	1.3	-	-	-	76.7
Por	2.0	-	0.7	-	-	-	0.7	-	45.3	4.7	-	-	46.7
Fig	-	-	-	-	-	-	18.7	-	6.7	42.0	0.7	-	32
Voi	4.1	-	-	-	2.0	-	-	-	-	2.0	34.0	-	57.8
Fle	-	-	-	2.1	-	0.7	-	1.4	-	-	0.7	24.8	70.2

Tableau 2. Matrice de confusion obtenue en utilisant un codage par ondelettes de Gabor et réseau de neurones avec fonctions de base radiales. La colonne R_j correspond à la classe rejet.

En comparaison, les résultats obtenus par le codage par motifs fréquents cités plus haut montrent des taux de généralisation variant de 10% (pour la classe des voitures (11)) à 100% (pour la classe des minéraux (4)), avec une moyenne de 45.8%. L'écart de moyenne est donc de 22.2% et de 49.3% sur le meilleur taux de généralisation. Ces écarts sont considérables compte tenu de l'expertise mise en jeu dans la réalisation de l'approche par précodage et par réseau de neurones.

6. Conclusion

Ce papier présente une méthode de précodage des données par l'utilisation de fonctions de base correspondant à des motifs fréquents (de faible taux de couverture) trouvés dans les données d'apprentissage. Ce codage est inhabituel : il n'est pas défini *a priori*, mais au contraire dépend des données ; il ne permet pas de reconstruire les données codées et n'a donc pas de capacité d'approximation ; finalement, l'orthogonalité des fonctions de base est définie par rapport aux données.

Initialement, c'est l'impossibilité d'utiliser un algorithme d'analyse en composantes indépendantes en grande dimension qui nous a conduit à chercher à adapter l'approche de recherche de motifs fréquents afin de trouver directement un codage clairsemé des données. Ce codage, obtenu de manière non supervisée, s'est révélé très performant, beaucoup plus que des méthodes dédiées, dans une tâche réputée difficile de reconnaissance de scènes naturelles. Ce résultat a constitué une surprise et il reste donc à explorer les possibilités, et les limites, de cette nouvelle méthode de prétraitement et de codage des données.

Nos perspectives de recherche concernent l'étude théorique des propriétés et des limites du codage non supervisé présenté, ainsi que celle des caractéristiques des problèmes supervisés pour lesquels ce codage est adapté.

7. Bibliographie

- [AGR 94] AGRAWAL R., SRIKANT R., « Fast algorithms for mining association rules », *Very Large Data Bases*, Santiago, Chile, September 1994, p. 487-499.
- [BAR 61] BARLOW H., « Possible principles underlying the transformations of sensory messages », *Sensory communication*, p. 425-443, MIT press, 1961.
- [BAR 89] BARLOW H., « Unsupervised learning », *Neural Computation*, vol. 1, 1989, p. 295-311.
- [CAR 98] CARDOSO J.-F., « Blind signal separation : statistical principles », *Proceedings of the IEEE*, vol. 9, n° 10, 1998, p. 2009-2025.
- [DEN 02] DENQUIVE N., TARROUX P., « Catégorisation de scènes visuelles », *Technique et Science Informatique*, vol. X, n° X, 2002, p. 1-18.
- [FIE 94] FIELD D., « What is the goal of sensory coding? », *Neural Computation*, vol. 6, 1994, p. 559-601.
- [HYV 01] HYVARINEN, KARHUNEN, OJA, *Independent component analysis (Adaptive and learning systems for signal processing, communications and control series)*, John Wiley and Sons, 2001.
- [THO 96] THORPE S., FIZE D., MARLOT C., « Speed of processing in the human visual system », *Nature*, vol. 381, n° 6, 1996, p. 520-522.