Sélection adaptative de caractéristiques pertinentes et classification hiérarchique des images dans les bases hétérogènes

Rostom Kachouri, Khalifa Djemal, Hichem Maaref

Laboratoire d'Informatique, Biologie Intégrative et Systèmes Complexes
Université d'Évry Val-d'Essonne
40, rue du Pelvoux, F-91020 Évry

Rostom.Kachouri@ibisc.univ-evry.fr, Khalifa.Djemal@iup.univ-evry.fr
Hichem.Maaref@univ-evry.fr

RÉSUMÉ. *Dans les bases hétérogènes, les images appartiennent souvent à différentes classes thématiques et nécessitent une large description permettant leur reconnaissance. Cependant, les caractéristiques utilisées ne sont pas toujours adaptées au contenu de la base d'images considérée. Nous proposons dans cet article une nouvelle approche se basant sur deux originalités, à savoir la sélection adaptative de caractéristiques et la classification multi-modèle intitulée MC-MM. La sélection adaptative permet de ne considérer que les caractéristiques les mieux adaptées au contenu de la base d'images utilisée. La méthode MC-MM assure la reconnaissance des images en se servant hiérarchiquement des caractéristiques sélectionnées. Les résultats expérimentaux obtenus confirment l'efficacité et la robustesse de notre approche.*

ABSTRACT. *In heterogeneous databases, images often provided from different sources and belong to different topics, hence there is a need for a large description to ensure efficient representation of their content. However, extracted features are not always adapted to the considered image database. In this paper we propose a new image recognition approach based on two innovations, namely adaptive feature selection and Multi-Model Classification Method (MC-MM). The adaptive selection considers only the most adapted features with the used image database content. The MC-MM method ensures image recognition using hierarchically selected features. Experimental results confirm the effectiveness and the robustness of our proposed approach.*

MOTS-CLÉS : *extraction d'attributs, sélection adaptative des caractéristiques pertinentes, classification multi-modèle, reconnaissance d'images, bases hétérogènes.*

KEYWORDS: *feature extraction, adaptive relevant feature selection, multi-model classification, image recognition, heterogeneous image database.*

DOI:10.3166/TS.28.547-574 © 2011 Lavoisier

Extended abstract

In this paper, we are interested in Content Based Image Recognition (*CBIR*) in heterogeneous databases. Unlike the text-based approaches, *CBIR* systems allow image access according to their visual characteristics. The process of describing the image content with informations that can be derived from the image itself such as color, texture and shape is called feature extraction. Images in heterogeneous databases often belong to different topics and then a large description is generally required. In this work, average color vectors, histogram and correlogram are used as color features. First order statistics, co-occurence matrix coefficients and Gradient norm vectors are used as texture features. The GIST descriptor is also employed as feature covering both color and texture and finally the invariant moments of Hu are used as shape features. However, the encountered problem so far is the choice of relevant features depending on the considered image database. Indeed, extracted features are not always adapted to the content of images. Consequently, relevant feature selection is strongly needed.

Several feature selection techniques are available in the literature. Mainly, we distinguish two known selection method, named *wrappers* and *filters* approches. As they rely only on theoretical considerations, *Filter* methods are very fast, but not always efficient. Contrariwise, *Wrapper* methods use the classifier in the selection process, so they perform high recognition rates but still less fast especially for a large number of features. New methods that combine the two selection techniques are recently proposed. In this context, we propose in this paper a new adaptive feature selection. We use Support Vector Machine classifier (*SVM*) to evaluate the extracted features. Actually, we carry out multiple *SVM* learning using each feature separately. Subsequently, we apply Fisher Linear Discriminant (*FLD*) to select the most relevant features based on the performed *SVM* evaluation. In fact, we compute an *FLD* threshold that ensures better separation between relevant and irrelevant features depending on the obtained training performance. Hence, the proposed method achieves automatically relevant feature selection according to the observed image database content. However, selected features have not the same relevance.

Considering the negative effect of the least efficient ones, a simple concatenation of selected features does not lead to optimal recognition results. For this, we propose in this paper to recognize images by means of a hierarchical classification technique, that we call *MC-MM* (méthode de classification multi-modèle). It derives directly from the above described adaptive feature selection. The *SVM* classifier is used and the employed hierarchical order relies essentially on the selected feature training rate. Images are initially classified according to the feature model having the lowest training performance among the selected ones. Afterwards, image classification is progressively refined through different hierarchical levels. In fact, at each upcoming level in the classification hierarchy, images are classified according to the subsequent model until reaching the most relevant one at the last level. Furthermore, the own classification of each level is usually compared to that obtained within the previous level. In the case of dissimilar classification, the Nearest Cluster Center (*NCC*) classifier is employed. The

NCC classifier consists of a simple process in which the considered image is assigned, among two evaluated clusters, to the closest one in a given feature space.

Given selection and classification results in this paper are obtained from experiments on different COREL database subsets. For each COREL subset, we use the "$3/4$ - $1/4$" proportion to learning and testing respectively. Thus from the 100 images of each cluster in the COREL database, 75 images are randomly sampled for learning. The remaining 25 images are used for test. To assess the individual relevance of the extracted features, we carried out *SVM* evaluations of the corresponding models. The performed evaluations within the different used COREL subsets prove that the training rate efficiency of a given model relies basically on the employed image content. In this context, the selected features through the proposed adaptive selection vary always depending on the processed image subsets. The proposed generalization procedure is also assessed by comparison with two other generalization methods. The first proceeds in an opposite manner, ie from the most relevant models to the less ones. The second assigns images to the considered cluster by the majority of selected models. For the different COREL subsets results prove that the proposed method provides always better classification rates.

Furthermore, the proposed hierarchical way to combine features is assessed. For comparison, we carried out the *SVM* based classical classification method. In addition, we compare the obtained recognition results of the proposed method with those of methods present in the literature, such as *k*-means-SVM, DD-SVM, MILES, MI-SVM and the SIFT based Bag of Features. This comparative analysis demonstrates that the proposed method outperforms the different evaluated methods. Indeed, it provides 83.7% as average classification accuracy while the best obtained performance from all the other discussed methods does not exceed 82.6%.

1. Introduction

De nombreuses applications des systèmes de vision par ordinateur (Moeslund, Granum, 2001) comme les interactions homme-machine, la détection des événements, etc., ont fait apparaître un grand besoin de systèmes de recherche et de reconnaissance d'images dans les bases de données. La recherche d'images a été basée initialement sur l'annotation textuelle des images (Ismail *et al.*, 2008). Le coût de génération d'annotations textuelles fiables dans de grandes bases d'images, ainsi que le « fossé linguistique [1] » associé à ces annotations expliquent pourquoi des techniques de recherche d'images basées directement sur leur contenu visuel sont aujourd'hui d'un grand intérêt (Gevers, Smeulders, 2004). Contrairement aux méthodes d'annotation textuelle, ces nouvelles techniques permettent l'accès aux images à partir de leurs propriétés visuelles, c'est-à-dire à partir des informations qui peuvent être extraites

1. Le fossé linguistique reflète les difficultés rencontrées lors de la description du contenu des images par des mots-clés textuelles. En particulier, la grande variance des différentes interprétations sémantiques possibles pour un même contenu image.

des images elles-mêmes, telles que la couleur, la texture, la forme, etc. Depuis une dizaine d'années, différents systèmes de recherche et de reconnaissance d'images par le contenu, souvent désignés par l'acronyme CBIR (*Content Based Image Retrieval*), sont mis en place. L'architecture type de tels systèmes est composée principalement de deux étapes importantes, à savoir l'extraction d'attributs visuels pour la description des images et la classification qui mène à la reconnaissance de ces images (figure 1).

Dans la littérature, il est d'usage de distinguer deux types d'attributs connus par deux désignations : les attributs locaux et les attributs globaux (Aly *et al.*, 2009). En effet, les attributs visuels sont extraits à partir de l'image selon trois niveaux d'abstraction : le niveau points d'intérêt (Lowe, 2004 ; Mikolajczyk, Schmid, 2004 ; Mikolajczyk *et al.*, 2005), le niveau régions (Moghaddam *et al.*, 1999 ; Malki *et al.*, 1999 ; Del Bimbo, Vicario, 1998 ; Carson *et al.*, 1999 ; Ma, Manjunath, 1999) et le niveau global (Pentland *et al.*, 1994 ; Flickner *et al.*, 1995 ; Bach *et al.*, 1996 ; Sclaroff *et al.*, 1997 ; Gevers, Smeulders, 1999 ; Brunelli, Mich, 2000 ; Boujemaa *et al.*, 2001). Les deux premiers niveaux reflètent une approche locale de description d'images, tandis que le dernier permet d'assurer une description globale. L'approche locale se base essentiellement sur la description de parties spécifiques dans l'image, alors que l'approche globale permet la reconnaissance des images en se basant sur des similarités visuelles mesurées sur les images entières. On distingue également des procédures de classification supervisée (Bi *et al.*, 2003 ; Zhu *et al.*, 2003 ; Vapnik, 1999 ; Datta *et al.*, 2006) et non supervisée (Portnoy *et al.*, 2002 ; Shamir, Sharan, 2001 ; Berry, 2003) selon que l'on dispose ou non d'une connaissance *a priori* sur les images. La classification supervisée se base sur une technique d'apprentissage où l'on cherche à produire automatiquement des règles de décision à partir d'une base d'apprentissage contenant des images déjà connues (ou classifiées). Par contre, la classification non supervisée se ramène au regroupement des images de la façon la plus naturelle possible sans aucune connaissance *a priori* de la base. Dans cet article, nous nous intéressons à la reconnaissance d'images par le contenu dans les bases hétérogènes. Selon le contexte de la constitution de ce type de base, les images appartenant sémantiquement à un thème donné, représenté par une même classe d'images, ne présentent pas forcément une majorité de sous-régions ou de points d'intérêt communs. De plus, certaines images appartenant à des classes différentes peuvent se révéler d'un aspect très proche, alors que des images de la même classe peuvent avoir des contenus dissemblables. De ce fait, nous retenons dans ce travail l'approche globale pour une description fiable des images et nous adoptons le principe de classification supervisée pour assurer une séparation adéquate des différentes classes d'images dans la base.

Toutefois, une caractérisation robuste et discriminante des images dans les bases hétérogènes reste un grand défi. En effet, face à un très large éventail de caractéristiques disponibles dans la littérature (Androutsos *et al.*, 2005 ; Manjunath *et al.*, 1998), le problème rencontré est souvent le choix des caractéristiques pertinentes en fonction de la base d'images considérée. Plusieurs méthodes de sélection sont présentées dans la littérature (Weston *et al.*, 2000 ; Tieu, Viola, 2004 ; Muller *et al.*, 2004). Cependant, les caractéristiques sélectionnées ne présentent pas de pouvoirs discriminants identiques. Considérant l'effet négatif des caractéristiques les moins performantes, un

unique vecteur constitué par une simple concaténation des attributs associés à l'ensemble des caractéristiques sélectionnées et représenté dans un espace vectoriel de même dimension, ne mène pas toujours à des résultats de classification optimaux. Pour ce faire, des méthodes de classification hiérarchique basées sur l'ordre de pertinence des caractéristiques retenues ont été proposées (Wenyin *et al.*, 2000 ; Landré, Truchetet, 2007). Par contre, une intervention d'un expert pour assurer l'organisation hors ligne des caractéristiques utilisées est souvent requise pour chaque base d'images. Ainsi de nouveaux systèmes CBIR sont recherchés. Ces systèmes doivent permettre l'amélioration des résultats de classification dans les bases d'images hétérogènes à travers une adaptation automatique des caractéristiques sélectionnées au contenu des bases considérées. C'est dans ce cadre que nous proposons dans cet article d'abord une méthode de sélection adaptative des caractéristiques pertinentes ; ensuite, en nous basant sur les résultats de cette sélection, nous présentons une nouvelle Méthode de Classification hiérarchique Multi-Modèle (MC-MM). Les résultats expérimentaux obtenus prouvent la performance de cette nouvelle méthode de classification. Une évaluation de l'approche proposée par rapport à un ensemble de travaux récents (Chen *et al.*, 2006 ; Chen, Wang, 2004 ; Andrews *et al.*, 2003 ; Csurka *et al.*, 2004) est réalisée. Une comparaison avec l'approche de sac des descripteurs (Csurka *et al.*, 2004 ; Willamowski *et al.*, 2004 ; Zhang *et al.*, 2007) est aussi effectuée.

L'organisation de l'article est la suivante. Dans la section 2, la méthode de sélection adaptative est présentée après une description des caractéristiques globales utilisées dans ce travail. La nouvelle méthode de classification MC-MM est proposée dans la section 3. Dans la section 4, nous présentons les résultats expérimentaux obtenus. Une étude comparative ainsi qu'une discussion justifiant la pertinence de l'approche proposée par rapport aux méthodes existantes sont aussi fournies dans cette section. Une conclusion ainsi que des perspectives sont présentées dans la section 5.

2. Extraction d'attributs et sélection adaptative

L'architecture standard d'un système de reconnaissance d'images basé sur une classification supervisée est illustrée par la figure 1. Les deux principales fonctions de ce système sont l'extraction d'attributs et la classification. Des règles de décision sont apprises hors ligne sur des images de la base d'apprentissage. Ensuite, le classifieur utilisé doit bien généraliser ces règles en ligne sur de nouvelles images. Par ailleurs, les performances de classification obtenues dépendent essentiellement de la description des images. Nous présentons dans ce qui suit les caractéristiques utilisées dans ce travail tout en justifiant les choix effectués. Ensuite, nous décrivons la méthode de sélection adaptative proposée.

2.1. Extraction d'attributs

L'objectif principal de l'extraction d'attributs est de déterminer pour chaque image une représentation (signature) qui soit d'une part compacte pour être rapidement accessible et d'autre part suffisamment complète pour bien caractériser l'image. Une

littérature abondante constitue l'état de l'art de l'approche de description globale (Pentland *et al.*, 1994 ; Flickner *et al.*, 1995 ; Bach *et al.*, 1996 ; Sclaroff *et al.*, 1997 ; Gevers, Smeulders, 1999 ; Brunelli, Mich, 2000 ; Boujemaa *et al.*, 2001). Plusieurs contributions récentes sont aussi proposées (Datta *et al.*, 2008 ; He, Cercone, 2009).

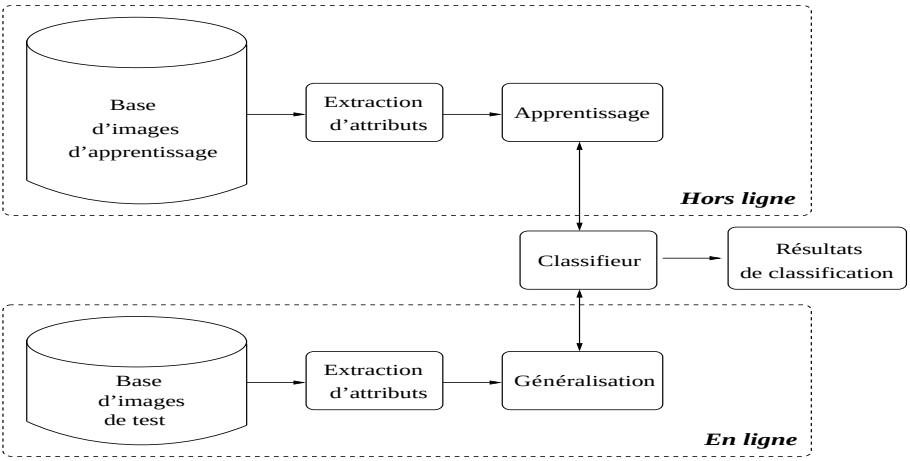
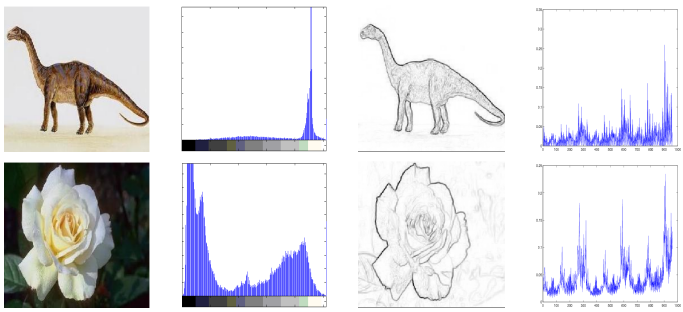


Figure 1. Architecture standard d'un système de reconnaissance d'images par le contenu basé sur une technique de classification supervisée

Les caractéristiques de couleurs sont très utilisées pour leur simplicité d'extraction, leur richesse de description et leur efficacité de reconnaissance. Nous employons pour ce type de description les vecteurs couleurs moyens (Faloutsos *et al.*, 1994) calculés dans quatre espaces couleurs différents, à savoir les espaces RVB, HSV, HMMD et YCrCb. L'espace RVB est généralement l'espace couleur de base dans lequel sont capturées les images. L'espace HSV est très proches de la perception humaine, tandis que les espaces HMMD et YCrCb représentent des transformations directes de l'espace RVB permettant d'améliorer la représentativité colorimétrique des images. Pour plus de détails concernant les espaces couleurs, nous recommandons aux lecteurs intéressés de consulter la référence (Danny, 2003). La caractéristique histogramme couleur (Hafner *et al.*, 1995) ainsi que la caractéristique corrélogramme (Huang *et al.*, 1997) sont également utilisées pour la description couleur. L'histogramme permet de représenter le nombre de pixels pour chaque niveau de couleur dans l'image, alors que le corrélogramme prend en compte la distribution spatiale des différentes intensités de couleur dans l'image. Les caractéristiques de textures permettent d'améliorer les performances de description des images, en particulier quand les distributions des couleurs sont très proches. Dans ce travail, nous employons les statistiques de premier ordre (Press *et al.*, 1987) calculées à partir de l'histogramme des niveaux de gris des images. Les attributs extraits de la matrice de co-occurrence (Haralick *et al.*, 1973), appelée souvent par l'acronyme anglais SGLD (*Spatial Gray Level Dependency*), sont aussi utilisés. Nous calculons en plus l'image gradient contenant les normes du gradient (Delingette, Montagnat, 2000) de tous les pixels dans l'image originelle. Elle présente le degré de différence d'intensité de chaque pixel par rapport à ses voisins.

Le descripteur GIST (Oliva, Torralba, 2001) est aussi utilisé comme une caractéristique représentant la couleur et la texture. Il permet d'identifier les principales régions homogènes de couleurs et de textures dans l'image. Les caractéristiques de couleur et de texture s'intéressent au contenu général de l'image. Par contre, les caractéristiques de forme sont capables de décrire les différents objets contenus dans l'image. Souvent, elles indiquent l'aspect général d'un objet comme son contour. Les moments de Hu (Hu, 1962) sont employés dans cet article. Ces moments sont invariants aux changements d'échelle, aux translations et aux rotations. La figure 2 présente quelques échantillons de caractéristiques utilisées pour représenter les images Dinosaure et Rose appartenant à deux classes différentes.



a) b) c)

Figure 2. Échantillons de caractéristiques représentant respectivement les images Dinosaure et Rose : a) histogramme, b) négatif des images gradient et c) GIST

Pour réduire la dimension de certaines caractéristiques tout en préservant leur pertinence de description, voire en l'améliorant, quelques traitements sont réalisés dans ce travail (Kachouri *et al.*, 2008) :

– nous utilisons l'auto-corrélogramme (Huang *et al.*, 1998) à la place du corrélogramme. C'est une sous-caractéristique qui capture seulement la corrélation spatiale entre les intensités de couleurs identiques dans l'image. Son équation est fournie en annexe A ;

– nous effectuons une projection selon les deux axes x et y de l'image gradient, ce qui permet de réduire sa dimension. La figure 3 montre la capacité de discrimination pertinente de cette nouvelle caractéristique ;

– enfin, les moments de Hu sont calculés à partir de l'image gradient. Cela permet de considérer le contour plus que toutes autres informations dans l'image, ce qui enrichit davantage la description de type forme de cette caractéristique. Les équations des sept moments de Hu sont données en annexe B.

Dans le reste de l'article, nous notons les différents vecteurs couleurs moyens par RVB, HSV, HMMD et YCrCb, l'histogramme couleur par Hist, l'auto-corrélogramme par A-Cor, la caractéristique contenant les statistiques de premiers ordres par FOS, la caractéristique contenant les attributs extraits de la matrice de co-occurence par

SGLD, les deux vecteurs d'attributs contenant respectivement les projections selon x et y de l'image gradient par G-norm, le descripteur GIST par GIST et la caractéristique contenant les moments invariants par Hu.

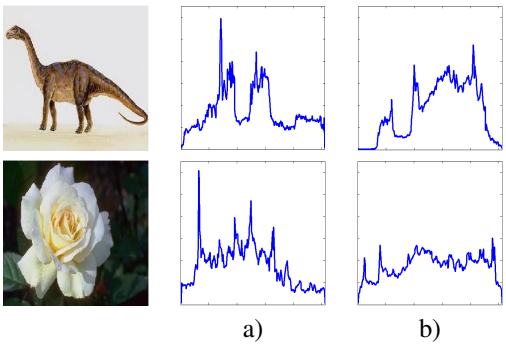


Figure 3. Nouvelle forme de la caractéristique image gradient des images Dinosaure et Rose : a) la projection selon x et b) la projection selon y

Face à une telle description et vu l'hétérogénéité des base considérées, quelques limitations apparaissent. En effet, plusieurs caractéristiques peuvent être redondantes ou non pertinentes parce qu'elles sont relativement similaires ou bien non discriminantes pour la classification d'une base d'images donnée. Ainsi les performances de généralisation risquent de souffrir sensiblement ; il est donc nécessaire de proposer des méthodes adaptées pour la sélection des caractéristiques les plus pertinentes. Dans ce contexte, nous proposons une méthode de sélection adaptative.

2.2. Sélection adaptative des caractéristiques pertinentes

La sélection d'un sous-ensemble de caractéristiques jugé comme le plus pertinent est un domaine sérieusement étudié et les algorithmes proposés sur ce sujet font l'objet, depuis plusieurs années, d'une littérature foisonnante (Guyon, Elisseeff, 2003). En effet, les méthodes de sélection proposées s'avèrent assez diverses et variées, mais il est souvent d'usage de distinguer deux branches principales, à savoir les méthodes filtres et les méthodes enveloppantes. Les méthodes filtres sont très rapides. Elles reposent sur des considérations théoriques, qui étudient les relations de dépendance entre les variables (Yu, Liu, 2004). Mais comme elles ne prennent pas en compte le classifieur utilisé en phase de généralisation (Jolliffe, 1986 ; Fukunaga, 1990), les méthodes filtres restent souvent incapables de garantir de bons taux de reconnaissance (Fenton *et al.*, 2001). Bien que conceptuellement plus simples que les filtres, les méthodes enveloppantes de sélection ont été introduites plus récemment (John *et al.*, 1994). Ce type de sélection fait appel à un algorithme de classification. Ainsi le sous-ensemble sélectionné de caractéristiques est bien adapté à l'algorithme de classification utilisé et les taux de reconnaissance obtenus sont généralement élevés (Bi *et al.*, 2003 ; Zhu *et al.*, 2003 ; Chen *et al.*, 2006). L'inconvénient majeur de cette technique

de sélection est son temps de calcul qui est assez coûteux, surtout quand le nombre d'images croît.

De nouvelles méthodes qui combinent ces deux types de sélection ont été récemment proposées (Guyon, Elisseeff, 2003). Elles permettent de profiter à la fois des points forts des deux techniques : de la rapidité des filtres et de l'adaptation caractéristiques-classifieur des méthodes enveloppantes. Dans ce même contexte, nous présentons dans cette section une méthode de sélection adaptative des caractéristiques pertinentes. Pour la partie enveloppante de la méthode, nous employons les machines à vecteurs supports (Vapnik, 1999), connues par l'acronyme SVM (*Support Vectors Machines*), pour l'évaluation de l'ensemble des caractéristiques utilisées en phase d'apprentissage, ce qui assure une adaptation du pouvoir discriminant de chaque caractéristique évaluée avec le contenu de la base d'images considérée. Le discriminant linéaire de Fisher, noté souvent par l'acronyme anglais FLD (Fisher, 1936), est ensuite utilisé pour la partie filtre de la méthode. FLD assure la sélection des caractéristiques les plus pertinentes en se basant sur les évaluations SVM. Dans ce qui suit, nous détaillons le principe de la méthode de sélection proposée.

Soit $\mathcal{B}_{App}$ une base d'apprentissage. On note par N_{Tot} le nombre total des images dans cette base et par $\mathcal{F}$ l'espace vectoriel d'attributs utilisé pour la représentation des n caractéristiques employées. Soit $\mathcal{C} = \{C_1, C_2, ..., C_N\}$ l'ensemble des N classes thématiques contenues dans $\mathcal{B}_{App}$. Pour évaluer l'ensemble des caractéristiques en fonction du contenu de $\mathcal{B}_{App}$, on entraîne le classifieur SVM-$multiclasses$ en considérant chacune des caractéristiques candidates. C'est-à-dire, pour les n caractéristiques utilisées, on effectue n apprentissages différents où l'on utilise à chaque fois une caractéristique unique. On note par $\mathcal{M}$ l'espace des modèles obtenu. Les n modèles appartenant à $\mathcal{M}$ sont les différentes hyper-surfaces qui correspondent aux n caractéristiques utilisées. Ces hyper-surfaces représentent les n frontières de séparation entre les différentes classes $\mathcal{C}$ de la base $\mathcal{B}_{App}$. Ainsi, pour chaque modèle, les N_{Tot} images de la base d'apprentissage sont divisées en deux groupes : groupe des images mal classifiées et groupe des images correctement classifiées. On note dans ce qui suit le nombre d'images dans ces deux groupes respectivement par $N_{Err}(i)$ et $N_{Corr}(i)$, $i = [1, ..., n]$. Pour chaque modèle dans $\mathcal{M}$, on évalue le taux d'erreur et le taux de classification en phase d'apprentissage (ou la performance d'apprentissage), qu'on note respectivement $\forall i \in \{1, ..., n\}$, $E(i)$ et $Per(i)$:

$$E(i) = \frac{N_{Err}(i)}{N_{Tot}},\ Per(i) = \frac{N_{Corr}(i)}{N_{Tot}},\ i = [1, ..., n] \tag{1}$$

Comme chaque caractéristique est évaluée en fonction des performances d'apprentissage du modèle associé, nous proposons de réaliser la sélection des caractéristiques pertinentes à travers la sélection des modèles les plus performants dans $\mathcal{M}$. L'avantage de cette sélection est qu'elle assure une réduction du nombre de modèles avec une faible augmentation de la la probabilité d'erreur de classification. En effet, elle permet de déterminer un sous-ensemble de modèles pertinents et d'exclure le reste des modèles qui sont considérés comme non discriminants pour la classification de la

base d'images utilisée. Ceci permet non seulement d'alléger le traitement de reconnaissance en ligne, mais aussi d'éviter l'effet négatif des modèles non pertinents sur le résultat final de reconnaissance.

Soit $\mathcal{T} = \{Per(1), Per(2), ..., Per(n)\}$ la liste organisées des performances d'apprentissage des différents modèles dans $\mathcal{M}$ ($Per(1) \geq Per(2) \geq ... \geq Per(n)$). Cette organisation par ordre de pertinence décroissant est très importante puisqu'elle est bien adaptée à la fois aux caractéristiques employées et au contenu même de la base d'images. En effet, elle change dès que l'ensemble de caractéristiques ou le contenu de la base change. À partir de la liste $\mathcal{T}$, nous proposons de déterminer selon le critère discriminant de Fisher un indice k qui représente le seuil adaptatif de sélection des modèles pertinents (k sépare la catégorie des performances d'apprentissage correspondant aux modèles pertinents de celles correspondant aux modèles non pertinents dans la liste $\mathcal{T}$). Pour cela, on commence par le calcul de $m_1(i)$ et $m_2(i)$, $\forall i \in [1, ..., n]$. Ces deux valeurs représentent les deux moyennes des performances d'apprentissage qui sont situées respectivement, de part et d'autre de la performance d'apprentissage $Per(i)$ dans la liste $\mathcal{T}$:

$$m_1(i) = \frac{1}{i}\sum_{j=1}^{i} Per(j),\ m_2(i) = \frac{1}{n-i}\sum_{j=i+1}^{n} Per(j),\ i = [1, ..., n] \tag{2}$$

En fonction de ces deux valeurs moyennes $m_1(i)$ et $m_2(i)$, on calcule les deux variances respectives $v_1^2(i)$ et $v_2^2(i)$:

$$v_1^2(i) = \sum_{j=1}^{i} |Per(j) - m_1(i)|^2,\ v_2^2(i) = \sum_{j=i+1}^{n} |Per(j) - m_2(i)|^2,\ i = [1, ..., n] \tag{3}$$

Maximiser la marge de séparation entre les modèles pertinents et ceux non pertinents revient essentiellement à maximiser l'entité $P(i)$:

$$P(i) = \frac{|m_1(i) - m_2(i)|}{v_1^2(i) + v_2^2(i)}\ , i = [1, ..., n] \tag{4}$$

Ainsi le seuil k est déterminé comme suit :

$$1 \leq k \leq n = \arg\max_i P(i) \tag{5}$$

k représente l'indice de la performance d'apprentissage à partir de laquelle le reste des performances dans la liste $\mathcal{T}$ correspondent à des modèles considérés comme non pertinents. Soit $\mathcal{T}_s = \{Per(1), Per(2), ..., Per(k)\}$ la liste des performances sélectionnées, où $Per(i)$ est la performance d'apprentissage du modèle sélectionné M_{si}. En

phase de généralisation de la classification MC-MM (section 3), seul le sous-ensemble des modèles sélectionnés $\mathcal{M}_s = \{M_{s1}, M_{s2}, ..., M_{sk}\}$ est employé.

La classification hiérarchique multi-modèle basée sur les résultats de cette sélection adaptative est présentée dans la section 3.

3. Classification hiérarchique multi-modèle

Nous présentons dans cette section une nouvelle méthode de classification qu'on appelle (MC-MM) destinée aux systèmes de reconnaissance d'images par le contenu dans les bases hétérogènes. MC-MM combine hiérarchiquement les résultats de classification des différents modèles sélectionnés. L'ordre hiérarchique employé est basé sur les degrés de pertinence des modèles sélectionnés, c'est-à-dire celui de leurs performances d'apprentissage croissantes. Ceci permet d'assurer une amélioration progressive de la décision de classification et par la suite une qualité de reconnaissance ascendante. Une évaluation de cette procédure de généralisation est fournie dans la section 4. La classification MC-MM est basée essentiellement sur une bivalidation hiérarchique des décisions SVM. Toutefois, une image requête peut être assignée différemment selon deux niveaux hiérarchiques consécutifs. Afin de résoudre un tel conflit de décision, le classifieur du plus proche centre de classe que nous appellons NCC (pour *Nearest Cluster Center*) est employé.

3.1. Résolution de conflit de décision par NCC

NCC est une procédure simple où les images sont assignées à la classe la plus proche dans un espace donné d'attributs.

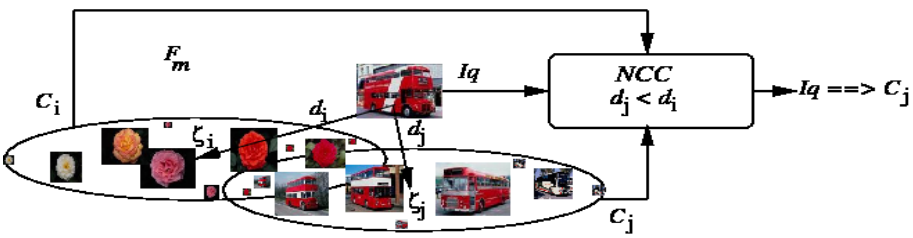


Figure 4. Règle de décision du classifieur NCC: classification de l'image requête I_q à travers les deux classes évaluées C_i et C_j dans l'espace d'attributs $\mathcal{F}_m$

Soit $1 \leq i \neq j \leq N$, I_q une image requête attribuée aux deux classes C_i et C_j selon deux niveaux hiérarchiques consécutifs L_l et L_m. Supposant que L_m est le niveau correspondant au modèle le plus performant parmi les deux modèles considérés et $\mathcal{F}_m$ est son espace d'attributs correspondant. NCC calcule initialement dans $\mathcal{F}_m$ les centres ζ_i et ζ_j respectivement des deux classes C_i et C_j. Ces centres sont fictifs et peuvent ne correspondre à aucune image dans la classe considérée. Ils sont représentés par la caractéristique moyenne des images d'apprentissage appartenant aux classes

correspondantes. On note la projection de l'image I_q dans l'espace $\mathcal{F}_m$ par $\mathcal{F}_m(I_q)$. Tel que présentée par la figure 4, la règle de décision du classifieur NCC est basée sur les mesures de similarité entre ζ_i, ζ_j et $\mathcal{F}_m(I_q)$. Pour cela, nous employons dans cet article la distance euclidienne (Wang *et al.*, 2005). Cette distance est la plus utilisée pour les mesures de similarité dans les espaces d'attributs. Toutefois, il est toujours possible d'utiliser d'autres métriques de l'état de l'art (Dengsheng, Lu, 2003).

Bien que basée sur les deux classifieurs SVM-$multiclasses$ et NCC, la classification MC-MM est principalement discriminante *via* la classification des SVM. En effet, l'apport du classifieur NCC n'est requis qu'en cas de conflit entre les décisions des différents niveaux hiérarchiques consécutifs. Il est employé pour situer les images candidates exclusivement par rapport au deux classes proposées par les modèles SVM considérés. Effectivement, les SVM, agissant par séparation, exploitent les relations de chaque classe avec les autres. Par contre, le classifieur NCC modélise chaque classe séparément. Dans ce qui suit, nous détaillons le principe de la méthode de classification MC-MM.

3.2. Classification MC-MM

Pour une base d'images de test $\mathcal{B}_{Test}$, on commence par l'extraction des seules caractéristiques correspondantes aux modèles appartenant à $\mathcal{M}_s$. Nous rappelons que l'ensemble des modèles pertinents $\mathcal{M}_s$ est déterminée par la sélection adaptative proposée dans la section 2.2. La première classification considérée pour une image requête I_q dans $\mathcal{B}_{Test}$ est celle obtenue avec le modèle M_{sk} (celui ayant la plus faible performance d'apprentissage dans $\mathcal{M}_s$). Notons par $C^{M_{sk}}(I_q) \in \mathcal{C}$ le résultat de classification de I_q avec M_{sk}. Cette classification est affinée progressivement à travers les différents niveaux hiérarchiques utilisés. En fait, à chaque niveau ultérieur dans l'architecture MC-MM (figure 5) I_q est classifiée avec le modèle subséquent dans $\mathcal{M}_s$ jusqu'à atteindre le modèle le plus performant (M_{s1}) au dernier niveau. Ainsi la classification de I_q selon chaque niveau est toujours comparée à celle obtenue avec le niveau précédent dans la hiérarchie MC-MM. Cette classification est principalement validée si les deux niveaux comparés présentent une classification similaire de I_q. Lorsqu'il y a conflit entre les décisions prises à ces deux niveaux consécutifs, l'ambiguïté est levée grâce à l'utilisation du classifieur NCC (section 3.1).

Au premier niveau hiérarchique, le résultat de classification de l'image I_q avec le modèle M_{sk-1}, noté par $C^{M_{sk-1}}(I_q) \in \mathcal{C}$, est comparé avec $C^{M_{sk}}(I_q)$. Si I_q est similairement classifiée avec M_{sk} et M_{sk-1}, la classification de I_q dans ce premier niveau est directement validée ($C^{M_{sk},M_{sk-1}}(I_q) = C^{M_{sk}}(I_q) = C^{M_{sk-1}}(I_q)$). Dans le cas contraire ($C^{M_{sk}}(I_q) \neq C^{M_{sk-1}}(I_q)$), I_q est assignée à la classe validée par le classifieur NCC. La classe $C^{M_{sk},M_{sk-1}}(I_q)$ est donc celle ayant le centre le plus proche de $\mathcal{F}_m(I_q)$ parmi les deux classes $C^{M_{sk}}(I_q)$ et $C^{M_{sk-1}}(I_q)$. De la même façon, les classes $C^{M_{sk},M_{sk-1}}(I_q)$ et $C^{M_{sk-2}}(I_q)$ sont comparées dans le deuxième niveau. Ainsi de suite, la classification de I_q est affinée progressivement avec la mé-

thode MC-MM jusqu'à atteindre le dernier niveau hiérarchique. La figure 5 illustre l'architecture de la classification multi-modèle MC-MM.

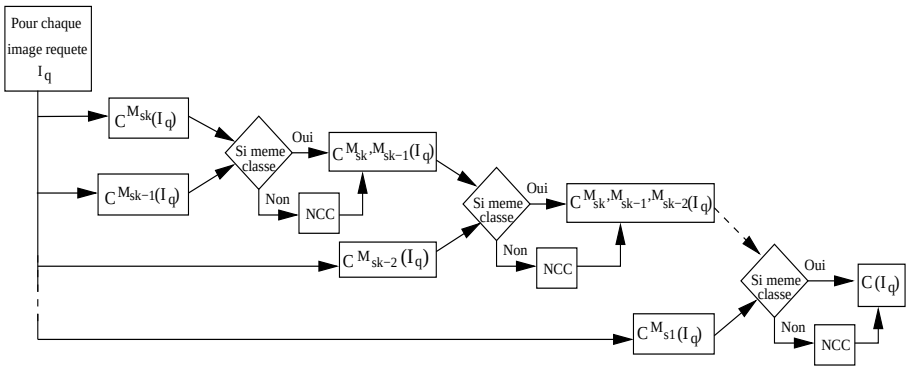


Figure 5. Diagramme illustrant l'architecture de la méthode de classification hiérarchique multi-modèle MC-MM

Notons par $C(I_q) = C^{M_{sk},M_{sk-1},\ldots,M_{s1}}(I_q)$ la classification de I_q avec MC-MM. On peut formuler cette classification comme suit :

$$C(I_q) = [\,[\,[C^{M_{sk}}(I_q) \otimes C^{M_{sk-1}}(I_q)]\ldots \otimes C^{M_{s2}}(I_q)] \otimes C^{M_{s1}}(I_q)], \qquad (6)$$

avec

$$C^{M_{si}}(I_q) \otimes C^{M_{sj}}(I_q) = \begin{cases} C^{M_{sj}}(I_q) \text{ si } C^{M_{si}}(I_q) = C^{M_{sj}}(I_q) \\ NCC(I_q, C^{M_{si}}(I_q), C^{M_{sj}}(I_q)) \text{ sinon} \end{cases}$$

La classification MC-MM est une classification automatique qui ne nécessite aucune rétroaction de l'utilisateur ou un bouclage de pertinence. De plus, elle assure de bonnes performances de classification grâce à l'utilisation hiérarchique de modèles bien adaptés au contenu de la base d'images. Nous présentons et nous discutons les performances de classification MC-MM dans la section 4.

4. Résultats expérimentaux

Nous présentons dans cette section l'évaluation de la sélection adaptative et de la classification MC-MM à travers une base d'images hétérogène de référence. La section 4.1 est réservée à la description de la base d'images utilisée. L'évaluation des modèles correspondant aux différentes caractéristiques employées ainsi que les résultats de sélection adaptative est présentée dans la section 4.2. La section 4.3 expose les résultats de classification MC-MM. Une discussion des performances obtenues ainsi qu'une comparaison avec les résultats de méthodes connues tirées de la littérature est aussi fournie dans cette section.

4.1. Base d'images

La difficulté est accrue pour un système de reconnaissance d'images par le contenu dans les bases hétérogènes. Généralement, il requiert l'identification préalable de thèmes spécifiques pour la création d'une base vérité-terrain. Dans cet article, nous utilisons la base COREL [2]. Cette base comprend 1000 images catégorisées par le concepteur en une vérité-terrain de 10 classes thématiques (personnes, animaux, fleurs, icebergs, etc.) contenant 100 images chacune. La base COREL est considérée comme une référence pour les spécialistes du domaine. En effet, la plupart des publications considérant la reconnaissance d'images dans les bases hétérogènes utilise cette base de données. Quelques échantillons des différentes classes d'images contenues dans la base utilisée sont représentés dans la figure 6. Les noms des classes ainsi que les identifiants correspondants sont consignés dans le tableau 1 [3].

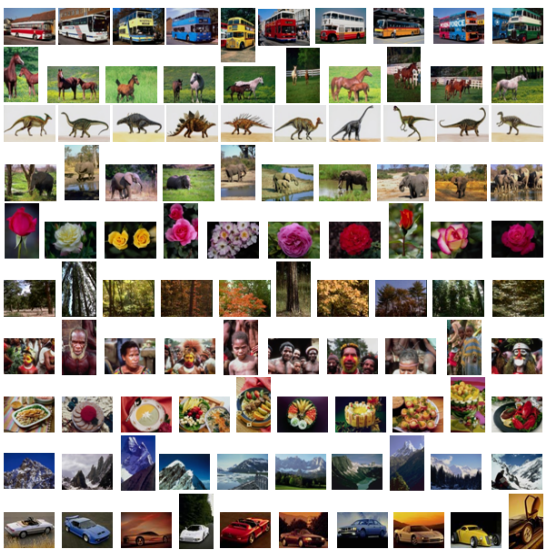


Figure 6. Échantillons des 10 classes de la base d'images hétérogène COREL utilisée

Les expérimentations sont réalisées en utilisant quatre sous-bases d'images, contenant respectivement 3, 5, 7 et 10 classes. Les identifiants des classes constituant ces différentes sous-bases sont les suivants : (2, 4 et 5), (2, 4, 5, 8 et 9), (0, 2, 4, 5, 6, 8 et 9) et (0, 1, 2, 3, 4, 5, 6, 7, 8 et 9). Pour ces quatre sous-bases, la proportion « $3/4$ - $1/4$ » est utilisée par convention pour l'apprentissage et le test. Ainsi, à partir des 100 images de chaque classe, 75 images sont tirées aléatoirement pour l'apprentissage. Les 25 images restantes sont utilisées pour le test.

2. La base COREL utilisée peut être téléchargée du site web de James Z. Wang's Research Group, disponible en ligne à l'adresse suivante : http://wang.ist.psu.edu/docs/home.shtml.

3. L'appellation anglaise est maintenue dans ce tableau.

Tableau 1. Noms et identifiants des 10 classes de la base d'images COREL utilisée

Identifiant	Nom de classe
0	Buses
1	Horses
2	Dinosaurs
3	Elephants
4	Flowers
5	Forests
6	African peoples and villages
7	Food
8	Mountains and glaciers
9	Cars

4.2. Évaluation des modèles

La connaissance de la vérité-terrain permet de réaliser une évaluation automatique et fournir des performances fiables. Les critères de qualité retenus pour l'évaluation sont les mesures de performance standard (Del Bimbo, 1999), c'est-à-dire la précision et le rappel. Soit pour une image requête donnée, TP le taux des vrais positifs, FP le taux des faux positifs et P le nombre total des images pertinentes [4] dans la base. Notons que les taux TP et FP indiquent respectivement le nombre d'images pertinentes et non pertinentes retrouvées. La précision et le rappel sont définis comme suit :

$$\text{Précision} = \frac{TP}{TP + FP}, \ \text{Rappel} = \frac{TP}{P} \tag{7}$$

Tableau 2. Évaluation de l'efficacité de classification des modèles individuels associés aux caractéristiques de couleurs pour les différentes sous-bases COREL utilisées

Sous-bases COREL		RVB	HSV	HMMD	YCrCb	Hist	A-Cor
3 classes	P_{moy}	0,8320	0,9071	0,9338	**0,8129**	**0,9643**	**0,9643**
	R_{moy}	0,8000	0,9067	0,9333	**0,7733**	**0,9600**	**0,9600**
5 classes	P_{moy}	**0,6735**	0,6889	0,7384	0,6897	0,8928	**0,9338**
	R_{moy}	**0,6640**	0,6880	0,7280	**0,6640**	0,8800	**0,9280**
7 classes	P_{moy}	0,6236	0,5684	**0,5657**	0,6444	**0,8593**	0,8549
	R_{moy}	0,6057	**0,5429**	0,5543	0,6114	**0,8400**	0,8343
10 classes	P_{moy}	0,5655	**0,4585**	0,4830	0,5750	**0,8055**	0,7996
	R_{moy}	0,5560	**0,4480**	0,4760	0,5720	0,7800	**0,7840**

4. Les images pertinentes sont les images présentant une bonne similarité avec une image requête donnée selon le principe de forme ou de contexte. Souvent, les images pertinentes d'une base sont celles appartenant à la même classe thématique de l'image requête dans cette base.

Tableau 3. Évaluation de l'efficacité de classification des modèles individuels associés aux caractéristiques de textures et de formes pour les différentes sous-bases COREL utilisées

Sous-bases COREL		FOS	SGLD	G-norm	GIST	Hu
3 classes	P_{moy}	**0,6801**	**0,8981**	0,8648	0,8919	0,7351
	R_{moy}	**0,6800**	**0,8800**	0,8267	0,8400	0,7333
5 classes	P_{moy}	**0,5858**	0,7439	0,7016	**0,8457**	0,6451
	R_{moy}	**0,5680**	**0,7200**	0,5840	0,5440	0,6480
7 classes	P_{moy}	**0,5087**	0,6844	0,7288	**0,8231**	0,5147
	R_{moy}	0,5143	**0,6743**	**0,4857**	0,5314	0,5200
10 classes	P_{moy}	**0,4228**	0,6387	0,6776	**0,7089**	0,4336
	R_{moy}	**0,4520**	**0,6160**	0,4760	0,4760	**0,4520**

Tableau 4. Nombres et listes des modèles sélectionnés pour les différentes sous-bases COREL utilisées

Sous-bases COREL	k	Listes $\mathcal{M}_s$ des k modèles sélectionnés (modèles consignés dans l'ordre de performances décroissant)
3 classes	9	{Hist, A-Cor, G-norm, GIST, HMMD, HSV, SGLD, RVB, YCrCb}
5 classes	9	{Hist, A-Cor, G-norm, GIST, RVB, HSV, HMMD, YCrCb, SGLD}
7 classes	9	{Hist, G-norm, GIST, A-Cor, RVB, YCrCb, SGLD, HSV, HMMD}
10 classes	4	{Hist, G-norm, GIST, A-Cor}

Pour tester la pertinence individuelle de chacune des caractéristiques employées dans ce travail (se référer à la section 2.1), une évaluation des modèles SVM correspondants est effectuée *via* les quatre sous-bases COREL. Les tableaux 2 et 3 présentent les valeurs moyennes de la précision et du rappel des modèles évalués. Les meilleures et les pires performances en précision et rappel dans ces tableaux sont affichées en gras pour les différentes sous-bases d'images utilisées. En plus de la variation évidente de ces performances en fonction de l'ensemble des caractéristiques employées, les résultats obtenus montrent clairement la variation de l'efficacité de classification d'un modèle donné par rapport aux différentes sous-bases d'images. Ceci prouve que le pouvoir discriminant des différentes caractéristiques n'est pas absolu, mais varie sensiblement en fonction du contenu de la base d'images utilisée. Le tableau 4 présente les résultats de sélection adaptative pour les différentes sous-bases d'images COREL. D'après ce tableau, nous constatons que selon les performances d'apprentissage des 11 modèles utilisés, 9 modèles sont sélectionnés pour les trois sous-bases contenant 3, 5 et 7 classes. Par contre pour la base contenant 10 classes, uniquement 4 modèles sont sélectionnés. Ceci résulte de l'accroissement de la complexité du contenu des images en fonction du nombre de classes d'images dans la base. En effet, l'augmentation du nombre de classes dans la base restreint de manière évidente le nombre de modèles pertinents qui permettent d'assurer des performances d'apprentissage conséquentes dans cette base. De plus, nous signalons que l'ordre de pertinence des modèles dans

l'ensemble $\mathcal{M}_s$ diffère d'une sous-base à l'autre tel que le cas des trois premières sous-bases où le nombre de modèles sélectionnés reste inchangé (tableau 4).

4.3. Résultats de la classification MC-MM

En se basant sur les résultats de la sélection adaptative (section 4.2), nous discutons dans cette section des résultats de classification MC-MM. Comme il est détaillé dans la section 3, la généralisation MC-MM procède hiérarchiquement à partir des modèles les moins performants vers les plus performants. Nous commençons par une évaluation de la procédure de généralisation proposée. Dans ce cadre, nous comparons les performances de la généralisation MC-MM avec deux autres procédures de généralisation possibles. La première adopte le même principe proposé sauf qu'elle procède de façon opposée, c'est-à-dire à partir des modèles les plus performants vers les moins performants. La deuxième assigne les images requêtes à la classe considérée par la majorité des modèles sélectionnés. Dans ce qui suit, ces deux généralisations sont indiquées respectivement par généralisation décroissante et généralisation max_modèles. Le tableau 5 présente l'étude comparative des performances de ces deux généralisations avec la généralisation MC-MM pour les différentes sous-bases COREL utilisées.

Tableau 5. Évaluation de la généralisation MC-MM. Étude comparative avec les deux procédures de généralisation décroissante et max_modèles pour les différentes sous-bases COREL utilisées

Sous-bases COREL	Précisions moyennes de classification des trois procédures de généralisation		
	décroissante	max_modèles	**MC-MM**
3 classes	0,8248	0,9803	**0,9879**
5 classes	0,7555	0,8425	**0,9750**
7 classes	0,7512	0,7914	**0,9123**
10 classes	0,7096	0,7638	**0,8371**

D'après le tableau 5, nous remarquons que la généralisation MC-MM est la plus performante parmi les trois procédures comparées. Pour les différentes sous-bases COREL, MC-MM assure des taux de classifications meilleurs que ceux obtenus avec les deux autres procédures de généralisation. La généralisation décroissante, qui procède de façon opposée à la généralisation MC-MM, est la moins performante. En fait, l'ordre d'utilisation des modèles dans les deux procédures de généralisation MC-MM et de généralisation décroissante constitue une sorte de pondération de ces modèles. Les premiers modèles utilisés sont les moins influents vu que la classification des images avec ces modèles risque d'être modifiée par les modèles qui suivent. Par conséquent, il est évident que les modèles les moins performants doivent être employés en premier lieu dans de telles architectures. Ce qui explique la supériorité des taux de classifications MC-MM par rapport à ceux de la généralisation décroissante. La généralisation max_modèles assure des résultats meilleurs que ceux de la généralisation

décroissante, mais ces résultats restent moins performants que ceux de la généralisation MC-MM. Ceci est justifié par une amélioration progressive de la classification des images à travers les différents niveaux hiérarchiques de la généralisation MC-MM. En plus, du fait qu'elle ne prenne pas en compte les degrés de pertinence des modèles sélectionnés, la généralisation max_modèles risque de classifier les images de test selon des modèles moins performants.

Tableau 6. Matrice de confusion de la classification MC-MM pour la base COREL 10 classes utilisée

	Cl. 0	Cl. 1	Cl. 2	Cl. 3	Cl. 4	Cl. 5	Cl. 6	Cl. 7	Cl. 8	Cl. 9
Cl. 0	**84**	0	0	0	0	0	8	8	0	0
Cl. 1	0	**96**	0	0	0	0	0	0	0	4
Cl. 2	0	0	**100**	0	0	0	0	0	0	0
Cl. 3	0	0	0	**73**	0	0	0	0	27	0
Cl. 4	9	12	0	0	**71**	4	0	4	0	0
Cl. 5	8	0	0	0	0	**92**	0	0	0	0
Cl. 6	4	0	0	12	0	0	**72**	0	12	0
Cl. 7	8	0	0	8	0	0	0	**76**	8	0
Cl. 8	6	0	0	17	0	0	4	0	**73**	0
Cl. 9	0	0	0	0	0	0	0	0	0	**100**

Nous présentons dans le tableau 6 la matrice de confusion obtenue à la suite de la classification MC-MM de la base COREL 10 classes. Les colonnes de cette matrice représentent les classes estimées, tandis que les lignes représentent les classes réelles (ou de référence). Ainsi le pourcentage des images correctement classifiées est présenté par la diagonale de cette matrice. En effet, un système de classification est d'autant meilleur que sa matrice de confusion se rapproche d'une matrice diagonale. D'après le tableau 6 il est à signaler que la classification est réussie à 100 % pour les classes 2 et 9. Pour le reste des classes le taux de classification varie entre 71 % et 96 %, ce qui prouve l'efficacité de la classification MC-MM face à la complexité et à la diversité du contenu de la base d'images hétérogène COREL. Nous illustrons à titre indicatif dans la figure 7 un exemple d'images-requêtes de la classe 6, les images-candidates correctement classifiées et celles mal classifiées. La figure 7 montre que 18 images-requêtes parmi 25 sont correctement classifiées et seulement 7 sont mal classifiées. Nous tenons à signaler que les caractéristiques sélectionnées pour la base COREL 10 classes s'avèrent pertinentes pour la classification de 72 % des images-requêtes de la classe 6. Cependant, nous précisons que la faiblesse de description des images mal classifiées est due principalement au contenu même de ces images. En effet d'après la figure 7, nous pouvons bien remarquer qu'au contraire des images correctement classifiées, la plupart des images mal classifiées contiennent une majorité de régions communes avec d'autres classes. Ces régions (d'herbes ou de ciel) semblent être spécifiques pour certaines classes autres que la classe 6 dans la base. Ceci explique bien les inévitables erreurs de reconnaissance présentées par la classification erronée des 7 images de la figure 7 dans les classes 3 et 8 par exemple (tableau 6).

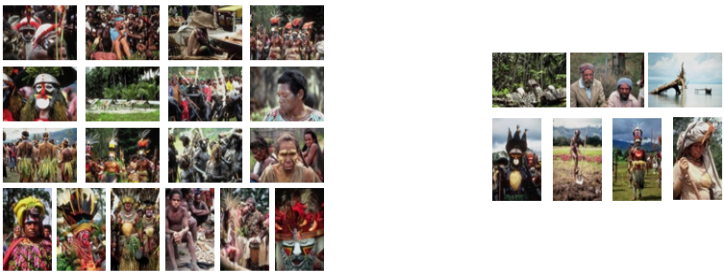


Figure 7. Exemple des résultats de classification MC-MM pour la base d'images COREL 10 classes utilisée : les images correctement classifiées (à gauche) et celles mal classifiées (à droite) de la classe « African peoples and villages »

Le tableau 7 présente les performances de classification de la méthode proposée pour les différentes sous-bases COREL utilisées. Nous fournissons également dans ce tableau les performances de classification de la méthode classique, dite aussi méthode simultanée (Kachouri *et al.*, 2008). La méthode de classification classique est basée sur un modèle SVM unique où l'ensemble des caractéristiques utilisées est employé simultanément. D'après les résultats présentés, il est clair que la classification MC-MM est beaucoup plus performante. Cette évaluation montre bien l'apport de la sélection adaptative des modèles pertinents, ainsi que l'emploi hiérarchique des modèles sélectionnés à travers la méthode MC-MM. Réellement, la prise en compte hiérarchique des caractéristiques sélectionnées permet de manipuler ces données sans difficulté en dépit de leurs énormes disparités de dimensions. En effet, en examinant le tableau 4, nous constatons de grandes disparités de taille entre les espaces de représentation des différentes caractéristiques sélectionnées pour chaque sous-base d'images. Par exemple, les vecteurs couleurs moyens (RVB et HSV) ont seulement 3 composantes, les attributs extraits des matrices de co-occurrence (SGLD) sont de l'ordre de la dizaine alors que les projections monodimensionnelles des images gradient (G-norm) ont pour dimensions la taille des images, c'est-à-dire de l'ordre de quelques centaines de valeurs.

Tableau 7. Évaluation des deux méthodes de classification MC-MM et classique pour les différentes sous-bases d'images COREL utilisées

Sous-bases COREL		Précisions de classification	
		Classique	MC-MM
3 classes	P_{moy}	0,7222	**0,9879**
	R_{moy}	0,4933	**0,9333**
5 classes	P_{moy}	0,7760	**0,9750**
	R_{moy}	0,3920	**0,9120**
7 classes	P_{moy}	0,7711	**0,9123**
	R_{moy}	0,3771	**0,8229**
10 classes	P_{moy}	0,7065	**0,8371**
	R_{moy}	0,3960	**0,8160**

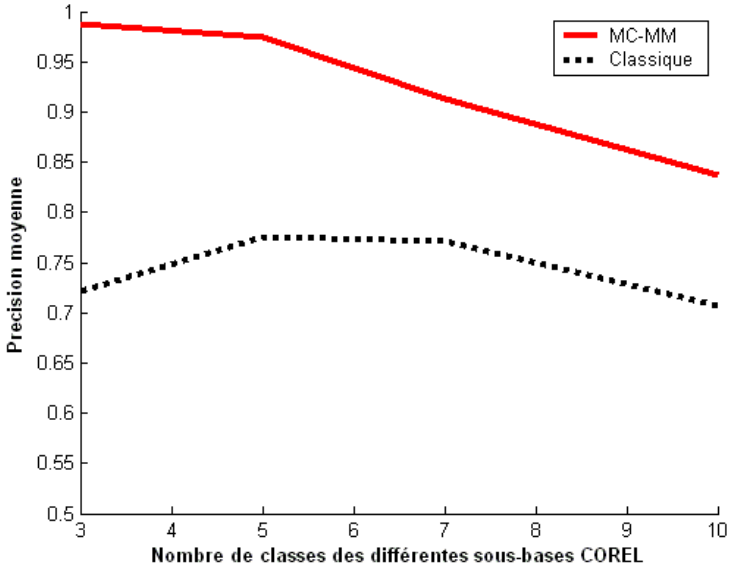


Figure 8. Comparaison des précisions moyennes de MC-MM et de la méthode de classification classique pour les différentes sous-bases d'images COREL utilisées

Les précisions moyennes des deux méthodes de classification pour les différentes sous-bases COREL sont présentées dans la figure 8. La performance de reconnaissance MC-MM décroît quand le nombre de classes augmente. En effet, nous obtenons des précisions moyennes de $0,9879$ pour 3 classes et $0,8371$ pour 10 classes. Ceci est dû à l'accroissement de la complexité du contenu des bases utilisées en fonction du nombre de classes. Par ailleurs, nous remarquons que l'utilisation hiérarchique des modèles SVM ainsi que leur complémentarité avec les décisions du classifieur NCC mènent toujours à des résultats de classification meilleurs que ceux obtenus avec les modèles sélectionnés séparément (tableaux 7, 2 et 3). Par conséquent, l'utilisation hiérarchique des modèles sélectionnés par le classifieur MC-MM permet toujours d'améliorer la caractérisation des images et d'assurer par la suite une meilleure reconnaissance.

Nous comparons les performances de classification de la méthode MC-MM proposée avec celles de différentes méthodes présentes dans la littérature, par exemple les méthodes *k*-means-SVM (Csurka *et al.*, 2004), DD-SVM [5] (Chen, Wang, 2004), MILES [6] (Chen *et al.*, 2006), MI-SVM (Andrews *et al.*, 2003) et la méthode sac des

5. Une implantation MatLab de DD-SVM est disponible en ligne à l'adresse suivante : http://www.cs.uno.edu/∼yixin/ddsvm.html.

6. Une implantation MatLab de MILES est disponible à l'adresse suivante : http://www.cs.olemiss.edu/∼ychen/MILES.html.

descripteurs [7] (Csurka *et al.*, 2004 ; Willamowski *et al.*, 2004 ; Zhang *et al.*, 2007) que l'on désigne souvent par l'acronyme BoF (*Bag of Features*). La méthode *k*-means-SVM construit des groupes de régions d'images en utilisant l'algorithme *k*-means. Dans cette méthode, les SVM sont construits en utilisant comme attributs le nombre de régions désignant chaque classe. Par contre, la méthode DD-SVM entraîne le classifieur SVM 1-norme dans l'espace des attributs construit à partir d'une projection définie par les maximums et les minimums locaux de la fonction de densité diverse (DD). Notons que la fonction DD mesure les co-occurrences des régions similaires dans les différentes images d'une même classe (Maron, Lozano-Pérez, 1998). La méthode MILES a été proposée comme une amélioration de DD-SVM par l'emploi d'une sélection enveloppante des attributs performants. En conséquence, le classifieur SVM 1-norme est appliqué simultanément, dans cette méthode, pour la sélection et pour la construction du classifieur. La méthode MI-SVM est basée quant à elle sur la sélection des régions considérées comme positives pour la classification des images. Elle utilise le même ensemble d'attributs régions que DD-SVM. Enfin, la méthode BoF consiste principalement à évaluer un jeu de descripteurs visuels locaux sur un ensemble de points d'intérêt extraits des images. Pour l'implémentation BoF utilisée dans cet article, la détection et la description des points d'intérêt dans l'ensemble des images de la base sont assurées par le descripteur SIFT (Lowe, 2004). La méthode BoF permet ensuite de caractériser les statistiques de ces descripteurs locaux par le biais de leur histogramme de valeurs quantifiées. Ceci permet de caractériser globalement le contenu des images. Nous fournissons les performances de classification de ces méthodes pour la base COREL 10 classes dans le tableau 8 [8].

Tableau 8. Comparaison de la précision moyenne de classification de la méthode MC-MM avec celles d'autres méthodes de l'état de l'art pour la base COREL 10 classes utilisée

Méthodes comparées	Précisions de classification
MC-MM	**83,7**
MILES (Chen *et al.*, 2006)	82,6
DD-SVM (Chen, Wang, 2004)	81,5
MI-SVM (Andrews *et al.*, 2003)	74,7
BoF (Csurka *et al.*, 2004)	74,0
k-means-SVM (Csurka *et al.*, 2004)	69,8

Cette analyse comparative met en évidence la supériorité de la performance de notre méthode par rapport à celles des différentes méthodes évaluées. En effet, la méthode MC-MM fournit une précision moyenne de classification de l'ordre de $83,7\ \%$, tandis que la meilleure performance obtenue parmi toutes les autres méthodes examinées est celle de la méthode MILES qui atteint seulement $82,6\ \%$.

7. Une implantation MatLab de la méthode sac des descripteurs (Bag of Features) est disponible en ligne à l'adresse suivante : http://www.vlfeat.org/ vedaldi/code/bag/bag.html.

8. Les valeurs illustrées par ce tableau sont les précisions moyennes de classification (en pourcentage) des différentes méthodes comparées.

5. Conclusion et perspectives

Dans ce travail, nous avons présenté une nouvelle approche de reconnaissance d'images par le contenu dans les bases hétérogènes. Deux originalités ont été proposées dans cette approche, à savoir une méthode de sélection adaptative et une classification hiérarchique multi-modèle, intitulée MC-MM. À la suite d'un apprentissage multi-modèle des classifieurs SVM, la méthode de sélection proposée détermine un sous-ensemble des caractéristiques pertinentes. Les caractéristiques sont sélectionnées en se basant sur une discrimination (Fisher) de leurs taux d'apprentissage individuels. Cette sélection est totalement adaptée au contenu de la base utilisée et aux caractéristiques employées. Durant la classification MC-MM, uniquement les modèles des caractéristiques sélectionnées sont considérés. Ils sont employés hiérarchiquement selon l'ordre de leurs performances d'apprentissage d'une façon automatique et sans aucune rétroaction de l'utilisateur ou un bouclage de pertinence. Pour toute image requête le résultat de classification est affiné progressivement tout au long des niveaux hiérarchiques de la méthode MC-MM. Des résultats de classification similaires apportent une confirmation de décision des niveaux consécutifs. Par contre en cas de conflit, l'ambiguïté est levée grâce à l'utilisation du classifieur NCC.

Nous avons évalué la méthode de sélection adaptative ainsi que la classification MC-MM grâce aux expérimentations effectuées pour la base d'images de référence COREL. En comparaison avec des méthodes existantes dans la littérature, MC-MM présente de meilleures performances de classification et génère des résultats très compétitifs pour la classification de 10 classes thématiques. Ainsi l'approche proposée est bien robuste face aux deux phénomènes les plus perturbateurs dans un système de reconnaissance d'images par le contenu, à savoir les images aberrantes contenues dans les bases hétérogènes et les caractéristiques non pertinentes qui peuvent être utilisées pour la description des images. Notons que MC-MM ne peut pas être employée avec des systèmes mono-modèle basés sur la description d'une caractéristique unique. Ceci est dû essentiellement à l'hypothèse principale de la méthode MC-MM (classification multi-modèle). En effet, MC-MM est destinée aux systèmes basés sur une description composée par un ensemble varié de caractéristiques.

Comme nous l'avons évoqué précédemment, le contenu complexe des images dans les bases hétérogènes engendre souvent des structures dissemblables pour les images appartenant à la même classe thématique. Nous pensons qu'on peut améliorer les performances du système de reconnaissance en utilisant, avec l'architecture multi-modèle proposée, des caractéristiques de type sémantique. Ce type de description est soutenu par l'utilisateur, essentiellement pour les images requêtes, à travers un mécanisme de bouclage de pertinence (rétroaction). Les degrés de pertinence que le système accorde initialement aux différentes caractéristiques employées (par la sélection adaptative) sont ainsi modifiés à travers les différentes boucles.

L'intégration de la sélection adaptative dans la méthode de classification MC-MM peut être envisagée. En effet, on peut s'orienter vers une pondération dynamique des différentes caractéristiques. Cette pondération sera déterminée itérativement au cours

de la phase de généralisation tout en s'appuyant sur le retour de pertinence de l'utilisateur et les taux d'apprentissage initialement calculés.

Bibliographie

Aly M., Welinder P., Munich M., Perona P. (2009). Automatic Discovery of Image Families: Global vs. Local Features. In *Ieee international conference on image processing.*

Andrews S., Tsochantaridis I., Hofmann T. (2003). Support Vector Machines for Multiple-Instance Learning. In *Advances in neural information processing systems 15*, p. 561-568. Cambridge, MA, MIT Press.

Androutsos P., Kushki A., Plataniotis K. N., Venetsanopoulos A. N. (2005). Aggregation of Color and Shape Features for Hybrid Query Generation in Content Based Visual Information Retrieval. *Signal Processing*, vol. 85, p. 385-393.

Bach J. R., Fuller C., Gupta A., Hampapur A., Horowitz B., Humphrey R., Jain R., Shu C.-F. (1996). The Virage Image Search Engine: An Open Framework for Image Management. In *Spie conference on storage and retrieval for image and video databases*, vol. 2670, p. 76-87.

Berry M. (2003). *Survey of Text Mining : Clustering, Classification, and Retrieval*. Springer. Hardcover.

Bi J., Bennett K., Embrechts M., Breneman C., Song M. (2003). Dimensionality Reduction Via Sparse Support Vector Machines. *Journal of Machine Learning Research*, vol. 3, p. 1229-1243.

Boujemaa N., Fauqueur J., Ferecatu M., Fleuret F., Gouet V., Saux B. L., Sahbi H. (2001). IKONA: Interactive Generic and Specific Image Retrieval. In *International workshop on multimedia content-based indexing and retrieval (MMCBIR'2001).*

Brunelli R., Mich O. (2000). Compass: An Image Retrieval System for Distributed Databases. In *IEEE International Conference on Multimedia and Expo (I)*, p. 145-148.

Carson C., Thomas M., Belongie S., Hellerstein J., Malik J. (1999). *Blobworld: a System for Region-based Image Indexing and Retrieval*. Rapport technique. Berkeley, CA, USA.

Chen Y., Bi J., Wang J. Z. (2006). MILES: Multiple-Instance Learning via Embedded Instance Selection. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 28, n° 12, p. 1931-1947.

Chen Y., Wang J. Z. (2004). Image Categorization by Learning and Reasoning with Regions. *J. Mach. Learn. Res.*, vol. 5, p. 913-939.

Csurka G., Dance C. R., Fan L., Willamowski J., Bray C. (2004). Visual Categorization with Bags of Keypoints. In *Workshop on statistical learning in computer vision, eccv*, p. 59-74.

Danny P. (2003). *A Review of RGB Color Spaces...From xyY to R'G'B'*. Rapport technique. Montreal (Quebec) Canada, The BabelColor company.

Datta R., Joshi D., Li J., Wang J. Z. (2006). Studying Aesthetics in Photographic Images Using a Computational Approach. In *Proceedings of the european conference on computer vision*, p. 7-13.

Datta R., Joshi D., Li J., Wang J. Z. (2008). Image Retrieval: Ideas, Influences, and Trends of the New Age. *ACM Computing Surveys*, vol. 40, n° 2, p. 1-60.

Del Bimbo A. (1999). *Visual Information Retrieval*. San Francisco, CA, USA, Morgan Kaufmann Publishers Inc.

Del Bimbo A., Vicario E. (1998). Using Weighted Spatial Relationships in Retrieval by Visual Contents. In *Content-based access of image and video libraries, proceedings. ieee workshop on*, p. 35-39.

Delingette H., Montagnat J. (2000). Shape and Topology Constraints on Parametric Active Contours. *Computer Vision and Image Understanding*, vol. 83, n° 2, p. 140-171.

Dengsheng Z., Lu G. (2003). Evaluation of Similarity Measurement for Image Retrieval. In *International conference on neural networks and signal processing*, vol. 2, p. 928-931.

Faloutsos C., Equitz W., Flickner M., Niblack W., Petkovic D., Barber R. (1994). Efficient and Effective Querying by Image Content. *Journal of Intelligent Information Systems*, vol. 3, p. 231-262.

Fenton W. G., McGinnity T. M., Maguire L. P. (2001). Fault Diagnosis of Electronic Systems Using Intelligent Techniques: a Review. *IEEE Transactions on Systems, Man, and Cybernetics, Part C*, vol. 31, n° 3, p. 269-281.

Fisher R. A. (1936). The Use of Multiple Measurements in Taxonomic Problems. *Annals Eugenics*, vol. 7, p. 179-188.

Flickner M., Sawhney H., Niblack W., Ashley J., Huang Q., Dom B., Gorkani M., Hafner J., Lee D., Petkovic D., Steele D., Yanker P. (1995). Query by Image and Video Content: The QBIC System. *Computer*, vol. 28, p. 23-32.

Fukunaga K. (1990). *Introduction to statistical pattern recognition (2nd ed.)*. San Diego, CA, USA, Academic Press Professional, Inc.

Gevers T., Smeulders A. W. M. (1999). The PicToSeek WWW Image Search System. In *International conference on mathematics and computer science*, vol. 1, p. 264-269.

Gevers T., Smeulders A. W. M. (2004). Emerging Topics in Computer Vision. In M. et S. B. Kang (Ed.),, chap. Content-based image retrieval: An overvie. Prentice Hall.

Guyon I., Elisseeff A. (2003). An Introduction to Variable and Feature Selection. *Journal of Machine Learning Research*, vol. 3, p. 1157-1182.

Hafner J., Sawhney H. S., Equitz W., Flickner M., Niblack W. (1995). Efficient Color Histogram Indexing for Quadratic Form Distance Functions. *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 17, p. 729-736.

Haralick R. M., Shanmugam K., Dinstein I. (1973). Textural Features for Image Classification. *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 3, n° 6, p. 610-621.

He D., Cercone N. (2009). Local Triplet Pattern for Content-Based Image Retrieval. In *Proceedings of the 6th international conference on image analysis and recognition*, p. 229-238. Berlin, Heidelberg, Springer-Verlag.

Hu M. K. (1962). Visual Pattern Recognition by Moment Invariants. *Information Theory, IRE Transactions on*, vol. 8, n° 2, p. 179-187.

Huang J., Kumar R., Mitra M., Zhu J. (1998). Spatial Color Indexing and Applications. In *Proceedings of the 1998 conference on computer vision*, vol. 35, p. 245-268.

Huang J., Kumar S. R., Mitra M., Zhu W.-J., Zabih R. (1997). Image Indexing Using Color Correlograms. In *Proceedings of the 1997 conference on computer vision and pattern recognition (cvpr '97)*, p. 762-768. Washington, DC, USA, IEEE Computer Society.

Ismail M. B., Frigui H., Caudill J. (2008). Empirical Comparison of Automatic Image Annotation Systems. In *First international workshop on image processing theory, tools and applications (ipta)*, p. 114-121. Sousse, TUNISIA.

John G. H., Kohavi R., Pfleger K. (1994). Irrelevant Features and the Subset Selection Problem. In *International conference on machine learning*, p. 121-129.

Jolliffe I. (1986). *Principal Component Analysis*. New-York, Springer Verlag.

Kachouri R., Djemal K., Maaref H., Masmoudi D. S., Derbel N. (2008). Heterogeneous Image Retrieval System Based on Features Extraction and SVM Classifier. In J. Filipe, J. Andrade-Cetto, J.-L. Ferrier (Eds.), *Icinco-spsmc*, p. 137-142. INSTICC Press.

Landré J., Truchetet F. (2007). Image Retrieval with Binary Hamming Distance. In A. Ranchordas, H. Araújo, J. Vitrià (Eds.), *Visapp (2)*, p. 237-240. Barcelona-Espania, INSTICC - Institute for Systems and Technologies of Information, Control and Communication.

Lowe D. G. (2004). Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision*, vol. 60, p. 91-110.

Ma W. ying, Manjunath B. (1999). NeTra: A toolbox for navigating large image databases. In *Multimedia systems*, vol. 7, p. 568-571.

Malki J., Boujemaa N., Nastar C., Winter R. (1999). Region Queries without Segmentation for Image Retrieval by Content. In *Third international conference on visual information systems (visual'99)*, p. 115-122.

Manjunath B. S., Ohm J. rainer, Vasudevan V. V., Yamada A. (1998). Color and Texture Descriptors. *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 11, p. 703-715.

Maron O., Lozano-Pérez T. (1998). A Framework for Multiple-Instance Learning. In *Advances in neural information processing systems*, vol. 10, p. 570-576. Cambridge, MA: MIT Press.

Mikolajczyk K., Schmid C. (2004). Scale and Affine Invariant Interest Point Detectors. *International Journal of Computer Vision*, vol. 60, n° 1, p. 63-86.

Mikolajczyk K., Tuytelaars T., Schmid C., Zisserman A., Matas J., Schaffalitzky F., Kadir T., Gool L. V. (2005). A Comparison of Affine Region Detectors. *International Journal of Computer Vision*, vol. 65, n° 1/2, p. 43-72.

Moeslund T. B., Granum E. (2001). A Survey of Computer Vision-Based Human Motion Capture. *Computer Vision and Image Understanding*, vol. 81, n° 3, p. 231-268.

Moghaddam B., Biermann H., Margaritis D. (1999). Defining Image Content with Multiple Regions-of-Interest. In *8993, proceedings of ieee workshop on content-based access of image and video libraries*, p. 89-93.

Muller H., Michoux N., Bandon D., Geissbuhler A. (2004). A Review of Content-Based Image Retrieval Systems in Medical Applications - Clinical Benefits and Future Directions. *International Journal of Medical Informatics*, vol. 73, n° 1, p. 1-23.

Oliva A., Torralba A. (2001). Modeling the Shape of the Scene: A Holistic Representation of the Spatial Envelope. *International Journal of Computer Vision*, vol. 42, n° 3, p. 145-175.

Pentland A., Picard R., Sclaroff S. (1994). *Photobook: Content-based Manipulation of Image Databases* (vol. 2) n° 2185.

Portnoy D., Dr A., Bellaachia A., Bellaachia A., Chen Y., Elkahloun A. G. (2002). E-CAST: A Data Mining Algorithm for Gene Expression Data. In *Workshop on data mining in bioinformatics*, p. 49-54.

Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. P. (1987). *Numerical Recipes: The Art of Scientific Computing*. Cambridge University Press.

Sclaroff S., Taycher L., La Cascia M. (1997). ImageRover: a Content-Based Image Browser for the World Wide Web. In *Content-based access of image and video libraries, 1997. proceedings. ieee workshop on*, p. 2-9.

Shamir R., Sharan R. (2001). Algorithmic Approaches to Clustering Gene Expression Data. In *Current topics in computational biology*, p. 269-300. MIT Press.

Teague M. R. (1980). Image Analysis via the General Theory of Moments. *Journal of the Optical Society of America (1917-1983)*, vol. 70, p. 920-930.

Tieu K., Viola P. (2004). Boosting Image Retrieval. *International Journal of Computer Vision*, vol. 56, p. 17-36.

Vapnik V. N. (1999). An Overview of Statistical Learning Theory. *Neural Networks, IEEE Transactions on*, vol. 10, n° 5, p. 988-999.

Wang L., Zhang Y., Feng J. (2005). On the Euclidean Distance of Images. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 27, n° 8, p. 1334-1339.

Wenyin L., Wang T., Zhang H. (2000). *A Hierarchical Characterization Scheme for Image Retrieval* (vol. 3).

Weston J., Mukherjee S., Chapelle O., Pontil M., Poggio T., Vapnik V. (2000). Feature Selection for SVMs. In *Advances in neural information processing systems 13*, p. 668-674. MIT Press.

Willamowski J., Arregui D., Csurka G., Dance C. R., Fan L. (2004). Categorizing Nine Visual Classes Using Local Appearance Descriptors. In *Icpr workshop on learning for adaptable visual systems.*

Yu L., Liu H. (2004). Efficient Feature Selection via Analysis of Relevance and Redundancy. *Journal of Machine Learning Research*, vol. 5, p. 1205-1224.

Zhang J., Lazebnik S., Schmid C. (2007). Local Features and Kernels for Classification of Texture and Object Categories: a Comprehensive Study. *International Journal of Computer Vision*, vol. 73.

Zhu J., Rosset S., Hastie T., Tibshirani R. (2003). 1-Norm Support Vector Machines. In *Neural information processing systems*, vol. 16, p. 49-56.

Annexe A. Auto-corrélogramme

Soit $\mathcal{I}$ une image couleur quantifiée sur cl couleurs $\{c_1, c_2, ..., c_{cl}\}$. Chaque pixel p_i, $i = [1, 2, ..., N_p]$, de coordonnés (x, y) est doté d'une couleur $\mathcal{I}(p_i)$: $p_i \in \mathcal{I}_{c_j} \Leftrightarrow \mathcal{I}(p_i) = c_j, \forall j \in \{1, 2, ..., cl\}$.

La caractéristique corrélogramme (Huang *et al.*, 1997) présente les corrélations spatiales entre couleurs en fonction d'une distance inter-pixel, prise au sens de la norme L_∞ :

$$|p_1 - p_2| = max\{|x_1 - x_2|, |y_1 - y_2|\} \tag{8}$$

Pour une distance $Dist \in \{1, 2, ..., N\}$ fixée *a priori*, le corrélogramme de la couleur c_i par rapport à la couleur c_j est défini par :

$$Cor_{c_i,c_j}^{Dist} = \underset{p_1 \in \mathcal{I}_{c_i}}{prob}\ [p_2 \in \mathcal{I}_{c_j} \quad /|p_1 - p_2| = Dist] \tag{9}$$

Pour tout pixel p_1 de couleur c_i dans l'image $\mathcal{I}$, Cor_{c_i,c_j}^{Dist} retourne la probabilité [9] de trouver un pixel p_2 de couleur c_j à une distance $Dist$ de p_1.

L'auto-corrélogramme (Huang *et al.*, 1998) capture seulement la corrélation spatiale entre les intensités de couleurs identiques dans l'image :

$$A\text{-}Cor_{c_i}^{Dist} = \underset{p_1 \in \mathcal{I}_{c_i}}{prob}\ [p_2 \in \mathcal{I}_{c_i} \quad /|p_1 - p_2| = Dist] \tag{10}$$

Annexe B. Hu

Les « moments » sont des attributs de forme très puissants. Un moment μ_{ij} est une somme pondérée de tous les pixels en fonction de leurs positions dans l'image (Teague, 1980). En 1962, les sept moments de Hu ont été proposés (Hu, 1962) :

$$\begin{cases} \phi_1 = \mu_{20} + \mu_{02} \\ \phi_2 = (\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2 \\ \phi_3 = (\mu_{30} - 3\mu_{12})^2 + (3\mu_{21} - \mu_{03})^2 \\ \phi_4 = (\mu_{30} + \mu_{12})^2 + (\mu_{21} + \mu_{03})^2 \\ \phi_5 = (\mu_{30} - 3\mu_{12})(\mu_{30} + \mu_{12})[(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2] \\ \qquad +(3\mu_{21} - \mu_{03})(\mu_{21} + \mu_{03})[3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2] \\ \phi_6 = (\mu_{20} - \mu_{02})[(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2] \\ \qquad +4\mu_{11}(\mu_{30} + \mu_{12})(\mu_{21} + \mu_{03}) \\ \phi_7 = (3\mu_{21} - \mu_{03})(\mu_{3,0} + \mu_{12})[(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{0,3})^2] \\ \qquad -(\mu_{3,0} - 3\mu_{12})(\mu_{21} + \mu_{03})[3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2] \end{cases} \tag{11}$$

Article reçu le 29/11/2010
Accepté le 06/07/2011

9. Cette probabilité correspond au nombre d'occurrences des pixels p_2 de couleur c_j dans le voisinage du pixel p_1 ($p_1 \in \mathcal{I}_{c_i}$). Le voisinage considéré est fixé par la distance $Dist$.

Rostom Kachouri *est docteur en traitement des images et du signal de l'université d'Évry Val-d'Essonne, France. Depuis 2010, il occupe un poste de recherche post-doctoral à l'ESIEE Engineering, Paris. Ses travaux de recherche portent essentiellement sur le traitement des images et du signal et l'adéquation algorithme-architecture.*

Khalifa Djemal *est maître de conférences au département génie électrique et informatique industrielle à l'institut universitaire de technologie de l'université d'Évry Val-d'Essonne, France. Il travaille au sein de l'équipe TADIB du laboratoire IBISC. Ses recherches s'intéressent aux domaines du traitement des images et données (restauration, segmentation, classification et CBIR). Il a présidé la conférence internationale IPTA et le workshop international MIAD. Il est rapporteur d'un certain nombre de conférences et revues internationales.*

Hichem Maaref *est professeur des universités au sein de l'université d'Évry Val-d'Essonne. Depuis 2001, il est le responsable de l'équipe TADIB (traitement et analyse de données et d'image - biométrie) du laboratoire IBISC. Ses recherches s'intéressent principalement aux méthodes de reconnaissance de forme pour des applications d'analyse d'image, de fusion sensorielle et de navigation autonome de robots.*