



**HAL**  
open science

## Reconnaissance de symboles graphiques à partir d'une combinaison de descripteurs en intégrant leur comportement sur une base d'apprentissage

Jean-Pierre Salmon, Laurent Wendling, Salvatore Tabbone

### ► To cite this version:

Jean-Pierre Salmon, Laurent Wendling, Salvatore Tabbone. Reconnaissance de symboles graphiques à partir d'une combinaison de descripteurs en intégrant leur comportement sur une base d'apprentissage. Sep 2006, pp.97-102. hal-00113764

**HAL Id: hal-00113764**

**<https://hal.science/hal-00113764>**

Submitted on 14 Nov 2006

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Reconnaissance de symboles graphiques à partir d'une combinaison de descripteurs en intégrant leur comportement sur une base d'apprentissage

J.-P. Salmon – L. Wendling – S. Tabbone

LORIA,  
Campus scientifique, BP 239  
54 506 Vandœuvre-les-Nancy Cedex France  
{salmon,wendling,tabbone}@loria.fr

**Résumé :** *dans cet article nous proposons une nouvelle approche pour combiner des descripteurs de formes fondée sur l'étude comportementale de ceux-ci vis-à-vis d'un corpus d'apprentissage. Chaque descripteur est calculé sur plusieurs classes d'objets ou de symboles. Pour chaque échantillon et pour tous les descripteurs nous déterminons un profil type défini à partir d'un jeu d'apprentissage en prenant en compte les conflits pouvant exister entre descripteurs. Enfin, nous montrons que l'utilisation de la combinaison de descripteurs permet d'améliorer la classification sur des cas d'études réels.*

**Mots-clés :** descripteurs de formes, étude comportementale, combinaison de descripteurs.

## 1 Introduction

La diversité des images, et les différents types de bruits pouvant les altérer, ont entraîné l'apparition de nouveaux descripteurs destinés à reconnaître correctement ces images. Suivant le type d'image, il revient alors à la méthode de combinaison de prendre en compte principalement les descripteurs adéquats.

Notre travail se situe dans le cadre d'un problème de classification d'échantillon d'images diverses, avec plusieurs sources d'information à notre disposition pour prendre notre décision. Dans ce domaine de nombreuses solutions ont été proposées parmi les plus répandues on citera, sans être exhaustif : les SVM [21], les réseaux de neurones [22], les algorithmes génétiques [24, 27, 26], et les techniques de fusion classique [25, 23]... Certaines approches comme la fusion d'informations [20] ou la sélection de caractéristiques ont été spécialement créées pour répondre à ce problème de classification. Parmi ces approches dédiées : les différentes techniques de fusion (conjonctive, disjonctive, Dempster-Schafer [20]) permettent de gérer les conflits ou les accords entre plusieurs sources d'information. Tandis que pour la seconde approche, le choix des caractéristiques est un problème de sélection d'un sous ensemble de caractéristiques  $d$  parmi un ensemble  $D$  basé sur un critère d'optimisation. Les caractéristiques retenues sont les plus discriminantes et les moins redondantes. Notre approche s'inspire de ces deux techniques. Nous utilisons les principes de fusions pour définir des profils comportementaux de chaque source d'information et nous reprenons le concept de sélection pour favoriser les sources d'information les plus pertinentes. Dans notre cas la sélection s'effectue par le biais de pondération

des descripteurs. Notre critère d'optimisation est le taux de reconnaissance directement lié au taux d'ordonnement qui reflète la séparabilité des classes. Etudier le comportement de chaque source d'information et se restreindre à l'utilisation des informations les plus adaptées contribue à rendre les classes plus compactes et plus facilement dissociables.

L'importante variabilité des symboles à reconnaître requiert l'utilisation de descripteurs invariants à des transformations géométriques classiques. De nombreuses approches ont été proposées au cours des années passées. Habituellement les descripteurs de formes sont regroupés en deux catégories [19] : les approches de type contour et celles de type région avec des avantages et inconvénients bien connus. Certaines sont fondées sur les descripteurs de Fourier [10, 12, 14], d'autres sur le calcul des moments [3, 17], sur la transformée de Hough [2, 9] ou encore sur des descriptions angulaires [5, 16]. Dans cette article, nous limitons notre étude à des méthodes globales à la forme considérée mais permettant de conserver ou de retrouver des propriétés géométriques fondamentales comme la translation, la rotation et l'homothétie.

Un autre problème rencontré en reconnaissance des formes est le type de distance la mieux adaptée à l'appariement. Pour pallier ce problème le choix de la métrique est souvent relatif au problème considéré. Bhattacharya [4], distance Euclidienne ou de Mahalanobis, EMD [13] sont fréquemment utilisées. Dans le cadre de cet article, nous utiliserons la distance Euclidienne pour comparer des valeurs scalaires et essentiellement l'index de Tanimoto (min sur max) pour des descriptions d'objets sous la forme de vecteur. Ceci a permis d'obtenir des résultats intéressants, pour notre étude expérimentale, tout en nécessitant peu de temps de calcul pour la comparaison.

Nous commencerons par présenter les différentes notations employées dans cet article au §2. Les mesures de classification calculées pour comparer les différentes approches sont rappelées dans le §3. Puis nous présentons (§4) la méthode permettant de définir les profils de chaque descripteur ainsi que la méthode de combinaison. Enfin, une étude expérimentale sur des bases de test réelles est proposée §5.

## 2 Notations utilisées

Tout au long de cet article, nous considérons  $K$  classes de même taille  $V$  pour faciliter les explications mais ce n'est en aucun cas une nécessité. Soit  $I = \{i_u^q\}_{u=1,V \wedge q=1,K}$  l'en-

semble des  $V * K$  images et  $C = \{c_q\}_{q=1,K}$  l'ensemble des  $K$  classes tel que :  $c_q = \{i_u^q\}_{u=1,V}$ . Chaque classe  $c_q$  est divisée en deux : une partie de  $T$  images destinée à l'apprentissage, notée  $C_{learn}$ , et une autre de  $V - T$  images pour les tests, notée  $C_{test}$  :

$$\begin{aligned} C &= C_{learn} \cup C_{test} \\ C_{learn} &= \{C_{learn_q}\}_{q=1,K} \quad \text{avec} \quad C_{learn_q} = \{i_u^q\}_{u=1,T} \\ C_{test} &= \{C_{test_q}\}_{q=1,K} \quad \text{avec} \quad C_{test_q} = \{i_u^q\}_{u=T+1,V} \end{aligned}$$

Soit  $D = \{d_m\}_{m=1,N}$  l'ensemble des  $N$  descripteurs où  $d_m(i_u^q)$  correspond à la valeur du descripteur calculée sur l'image  $i_u^q$  et  $s_m(i_u^q, i_{u'}^q) \in [0, 1]$  le degré de similarité entre deux images et  $dist_m(d_m(i_u^q), d_m(i_{u'}^q))$  la distance telle que :  $s_m(i_u^q, i_{u'}^q) = 1 - dist_m(d_m(i_u^q), d_m(i_{u'}^q))$

Afin de déterminer la classe d'appartenance d'une image test, durant la phase expérimentale, on considère parmi l'ensemble des images de la base d'apprentissage celle qui possède le plus grand degré de similarité puis on affecte l'image test à cette classe telle que :  $i_q^u \in C_{test}$  et  $i_{q'}^{u'} \in C_{learn}$  deux images, si  $s_m(i_q^u, i_{q'}^{u'}) = \max\left(\left\{s_m(i_q^u, i_{q''}^{u''})\right\}_{i_{q''}^{u''} \in C_{learn}}\right)$  alors  $i_q^u \in C_{q'}$ .

En utilisant un seul descripteur  $d_m$  le taux de reconnaissance chute lorsque la taille de la base d'image grandit. De plus un descripteur  $d_m$  n'est pas assez pertinent pour toutes les classes d'images  $c_q$ . L'objectif de notre méthode est donc de pallier ces problèmes en améliorant le pouvoir de discrimination de chacun des descripteurs et en les combinant.

### 3 Mesures de classification

#### 3.1 Matrice de confusion

Une matrice de confusion permet d'évaluer l'aspect discriminant des descripteurs  $D$  par rapport à une classe d'image donnée  $C_q$ . Habituellement, une tête de classe  $t_q^u$  (symbole le plus représentatif) - ou un modèle type (graphe structurel, vecteur de caractéristiques...) - est déterminée à partir de la base d'apprentissage  $t_q^u \in C_{learn_q}$ . Lors de la phase d'apprentissage, les images  $i_u^q \in C_{learn_q}$   $\{t_q^u\}$  sont utilisées comme des requêtes. Les têtes de classe sont alors triées suivant la distance  $d_m$  qui les sépare de la requête. La tête de classe la plus proches est alors comptabilisée dans la matrice, à la ligne correspondant à classe de l'image requête.

Dans notre cas on considère chaque image de la base comme une tête de classe. Par conséquent, on comptabilise les  $V$  têtes de classes plus proches pour définir nos matrices de confusion.

#### 3.2 Mesure d'ordonnement

Cette mesure est fréquemment utilisée pour évaluer des méthodes de reconnaissance par le contenu [7]. Elle a été adaptée ici pour évaluer la robustesse des différentes approches. Soit une classe  $c_q$ , sa mesure d'ordonnement est définie par :

$$\begin{aligned} r(c_q) &= \frac{1}{V-T} \sum_{i_u^q \in C_{test_q}} \frac{1}{T} \left( \sum_{a=1}^T [\delta(cl(i_u^q, a), q)] \right. \\ &\quad \left. + \sum_{a=T}^{T*K} \left[ \delta(cl(i_u^q, a), q) \times \left(1 - \frac{a-T}{T*K-T}\right) \right] \right) \end{aligned}$$

$$\delta(cl(i_u^q, a), q) = \begin{cases} 1 & \text{si } cl(i_u^q, a) = q \\ 0 & \text{sinon} \end{cases}$$

avec  $K$  le nombre de classes,  $T$  et  $(V - T)$ , respectivement, le nombre d'images d'apprentissages et de tests par classe,  $cl(i_u^q, a)$  représente la classe de l'image réponse située au rang  $a$ , classée par ordre décroissant des degrés de similarité obtenus pour l'image requête  $i_u^q$ . Dans notre étude,  $\delta$  renvoie 1 lorsque l'image au rang  $a$  correspond à la classe  $c_q$  et 0 dans le cas contraire. La valeur d'ordonnement permet de mesurer la répartition des erreurs après une requête, ce qui n'apparaît pas à travers le taux de reconnaissance. De plus, quand la taille de la base de données augmente, l'évolution de la mesure fournit une estimation intéressante de la stabilité du descripteur de formes.

## 4 Fonctions de pondération

Un descripteur est généralement considéré comme discriminant lorsque le taux de reconnaissance associé est grand. La matrice de confusion est utilisée ici pour évaluer les taux de reconnaissance des descripteurs. Par définition si une image est correctement reconnue, il devrait avoir uniquement des images de la classe correspondante. A l'opposé, une mauvaise reconnaissance sera caractérisée par la présence d'images issues d'autres classes parmi les  $V$  premières images qui ont obtenu les degrés de similarité les plus élevés.

Le tableau 1 représente une matrice de confusion obtenue sur une base d'images comportant 9 classes et 11 objets par classe (base de Sharvit - cf. partie expérimentale). Le descripteur utilisé ici est la compacité. L'objectif n'est pas d'étudier la compacité mais de mettre en évidence les étapes effectuées pour un descripteur et une matrice donnée.

Cluster	1	2	3	4	5	6	7	8	9
1	<b>28</b>	30	0	7	19	0	22	0	15
2	29	<b>51</b>	0	7	5	0	20	0	9
3	0	0	<b>44</b>	4	0	30	0	42	1
4	7	6	4	<b>47</b>	19	0	7	1	30
5	20	6	0	17	<b>34</b>	0	15	0	29
6	0	0	33	0	0	<b>49</b>	0	39	0
7	20	30	0	7	16	0	<b>43</b>	0	5
8	0	0	40	0	0	35	0	<b>46</b>	0
9	16	11	0	25	29	0	3	0	<b>37</b>

TAB. 1 – Matrice de confusion de la compacité.

Par exemple, si on considère la 6<sup>ème</sup> ligne de la matrice, les valeurs correspondent à la répartition des  $V$  premières images qui ont obtenu les plus haut degrés de similarité pour chaque image de la base d'apprentissage. On constate que les

images sont regroupées principalement entre 3 classes (la 3, 6 et 8). Le taux de reconnaissance pour cette classe est de 40% (nombre d'objets correctement reconnus/nombre d'objets testés=49/121).

Maintenant, si on considère une classe d'images moins bien reconnue, par exemple celles de la classe 5 (5<sup>ème</sup> ligne), les images issues d'autres classes sont beaucoup plus nombreuses, parmi les  $V$  images réponses. Le descripteur est donc moins discriminant que dans l'exemple précédent. Ainsi on peut en déduire que la compacité n'est pas adaptée pour déterminer quels objets appartiennent à la classe 5. Ceci est mis en évidence à travers la représentation graphique donnée figure 1. Celle-ci est obtenue en calculant l'histogramme des distances entre les objets d'une classe et tous les autres objets de la base.

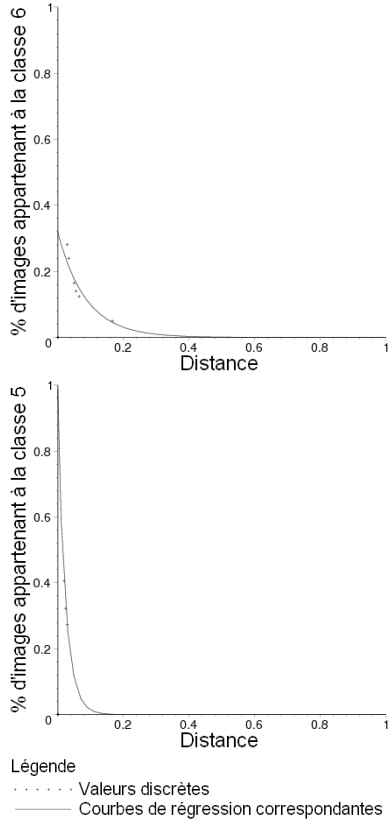


FIG. 1 – Différence entre un descripteur discriminant (bas) et non discriminant (haut). Une interpolation des données a été réalisée pour une meilleure visualisation.

En conclusion, on peut retenir qu'il est important de prendre en compte non seulement le taux de reconnaissance mais également la distribution des erreurs afin d'améliorer les performances du processus.

L'étude des erreurs relative à un descripteur, pour une classe donnée, nous a amené à constater que les images qui possèdent le plus haut degré de similarité (à contrario la plus faible distance) par rapport à une image test ne correspondent pas toujours à la même classe. Ainsi, il faut parfois considérer les images ayant un degré de similarité plus faible pour retrouver les images correspondant à la classe de la requête.

L'objectif est donc d'intégrer le comportement de chaque descripteur vis-à-vis d'une classe donnée en déterminant des

intervalles de confiance selon les distances. Une confiance élevée reflète une forte présence d'images de la classe considérée par rapport au nombre total d'images trouvées à cette distance.

Pour cela nous avons établi un profil  $W_{m,q}$  qui alloue une valeur de pondération proportionnelle à la confiance que l'on a dans le descripteur  $m$  en fonction de la distance, pour chaque classe  $q$  de la base d'apprentissage. Ce type de profil est fortement dépendant de la précision des distances.

L'algorithme fig. 2 résume la technique employée pour définir le profil.

**Pour Chaque**  $d_m \in D$   
**Pour Chaque**  $C_{learn_q} \in C_{learn}$   
**Pour Chaque**  $i_u^q \in C_{learn_q}$   
**Pour Chaque**  $i_{u'}^{q'} \in C_{learn}$   
 $distance = dist_m(d_m(i_u^q), d_m(i_{u'}^{q'}))$   
**Si**  $i_{u'}^{q'} \in C_{learn_q}$  **Alors**  
 Incréments ( $nbImageDeLaClasse_{m,q}(distance)$ )  
 Incréments ( $nb\_image\_total_{m,q}(distance)$ )  
 $W_{m,q}(distance) = \frac{nb\_image\_de\_la\_classe_{m,q}(distance)}{nb\_image\_total_{m,q}(distance)}$

FIG. 2 – Définition des poids.

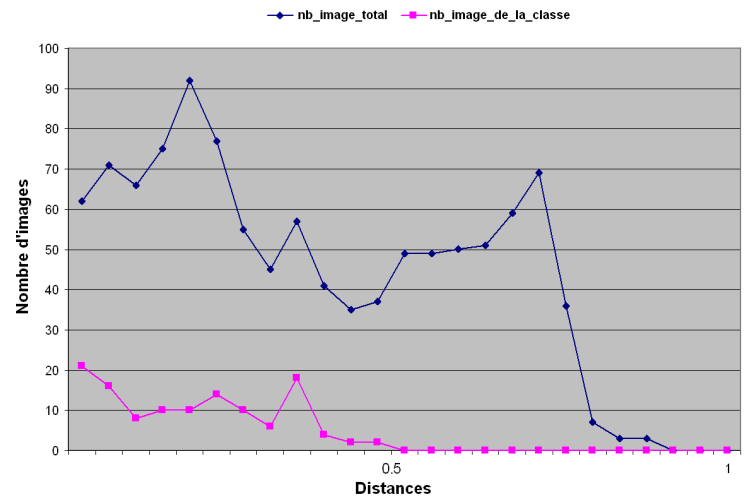


FIG. 3 – Distribution du nombre d'images par rapport à la  $distance_{Compacité,classe7}$ .

Par exemple, le profil donné fig. 4 (haut), correspondant à la classe 7 et au descripteur de compacité, est calculé à partir des courbes de la fig. 3.

La figure 4 (bas) décrit l'influence du profil  $w_{Compacité,classe7}(distance)$  sur la distance entre les images de la classe 7 (avant/après).

Le score est alors défini ainsi :

$$s(i_u^q, i_{u'}^{q'}) = \sum_{m=1}^N \left( 1 - dist_m(d_m(i_u^q), d_m(i_{u'}^{q'})) \right) \times W_{m,q}(dist_m(d_m(i_u^q), d_m(i_{u'}^{q'})))$$

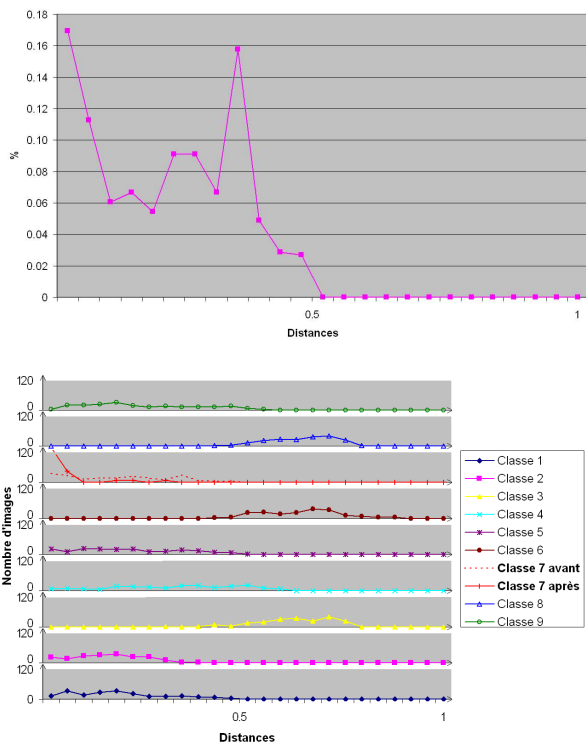


FIG. 4 – Profil des poids  $w_{Compacité,classe7}(distance)$  (haut). Influence de la fonction de pondération  $W_{m,q}$  sur la répartition du nombre d’images (bas) de la classe 7 (avant/après).

## 5 Etude expérimentale

### 5.1 Ensemble de descripteurs utilisés

Nous nous intéressons ici à des descripteurs simples à mettre en œuvre et rapides pour la plupart : compacité  $C$ , degré d’ellipticité [17]  $E$ , Signature angulaire [5]  $S_A$ , descripteur de Fourier générique [18]  $GFD$ ,  $\mathcal{R}$ -signature [16], Fourier-Mellin analytique [6, 8]  $FM$  et moments de Zernike [17, 11]  $MZ$ .

### 5.2 Résultats obtenus sur les bases d’images

Les images utilisées pour nos tests représentent des formes pleines, compatibles avec l’ensemble de nos descripteurs. Elles sont réparties dans 5 bases de données de 9 classes comprenant 11 échantillons (5 échantillons sont utilisés lors de la phase d’apprentissage et le reste pour les tests).

Ci-dessous, nous présentons les résultats obtenus sur deux bases contenant des images de différentes natures : pour la première il s’agit d’objets quelconques et de lettres pour la seconde. En plus de leur nature, les images de ces bases se différencient par les perturbations qu’elles comportent, mettant en avant la nécessité d’utiliser des descripteurs spécialisés.

#### 5.2.1 Base de données de formes

Pour montrer l’intérêt de notre approche, nous avons dans un premier temps, utilisé une base de données test proposée par D. Sharvit [15] sur son site web<sup>1</sup>. Des formes présentent

<sup>1</sup> <http://www.lems.brown.edu/vision/researchAreas/SIID/>

des occultations, d’autres sont déformées. Enfin certains objets sont de tailles différentes ou tournés.

La figure 5 donne les taux de reconnaissance moyens obtenus pour chacune des classes avant et après l’intégration des profils. Nous pouvons remarquer que la méthode que nous proposons permet d’améliorer de manière notable les taux de reconnaissance de la plupart des descripteurs.

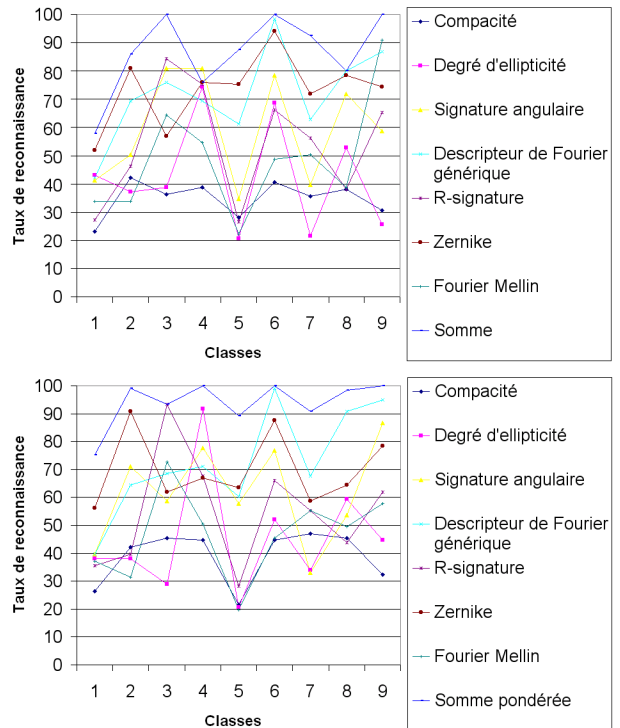


FIG. 5 – Taux de reconnaissance par descripteurs (avant/après).

Le tableau 2 donne les taux de reconnaissance moyen obtenus par les différentes méthodes. Nous intégrons ici  $W_S$  fondée sur une somme pondérée qui associe les profils et le taux global de reconnaissance obtenu par descripteur. Nous pouvons remarquer que ceci améliore de manière notable la reconnaissance.

Desc.	$C$	$E$	$S_A$	$G_{FD}$	$T_{Rf}$	$FM$	$MZ$	$W_S$
Av.	35	43	60	73	54	73	49	85
Ap.	39	45	62	73	55	76	50	94

TAB. 2 – Taux de reconnaissance obtenus sur la base Sharvit (avant/après).

Les taux d’ordonnement moyens donnés dans le tableau 3 soulignent aussi le bon comportement de la combinaison de descripteurs.

Nous pouvons remarquer également que les résultats obtenus sont en cohérence avec les méthodes employées. Les scores les plus bas sont atteints par les descripteurs dits « scalaires ». Alors que les méthodes fondées sur une signature unidimensionnelle ont des résultats moyens relativement proches. Les descripteurs de Fourier, qui peuvent être assimilés à une représentation bidimensionnelle, donnent de très

bons taux de reconnaissance car la signature est plus consistante. Enfin la combinaison qui intègre les performances de chacun des descripteurs permet d'améliorer encore la classification.

Desc.	<i>C</i>	<i>E</i>	<i>S<sub>A</sub></i>	<i>G<sub>FD</sub></i>	<i>T<sub>Rf</sub></i>	<i>FM</i>	<i>MZ</i>	<i>W<sub>S</sub></i>
Av.	87	85	88	96	89	98	87	94
Ap.	91	91	96	98	92	98	90	100

TAB. 3 – Taux d'ordonnement moyens obtenus sur la base Sharvit (avant/après).

### 5.2.2 Lettres graphiques extraites de Lettrines

Dans le cadre de l'ACI Madonne<sup>2</sup> axée sur la reconnaissance de documents anciens et dont le porteur est le L3i (La Rochelle), nous avons évalué notre approche sur une base composée de lettrines<sup>3</sup>.

Afin d'utiliser correctement nos descripteurs sur la base des lettrines, nous avons défini une séquence d'opérations permettant d'extraire les lettres graphiques contenues dans la plupart des documents qui nous étaient fournis. Pour cela nous avons réalisé une binarisation fondée sur un critère d'entropie floue, puis effectué une succession d'ouverture et fermeture pour enlever le bruit et rejoindre les parties des lettres éclatées. Nous nous sommes ensuite intéressés aux principales composantes connexes. Enfin nous avons extrait les lettres en considérant divers critères (compacité, taille, localisation spatiale...).

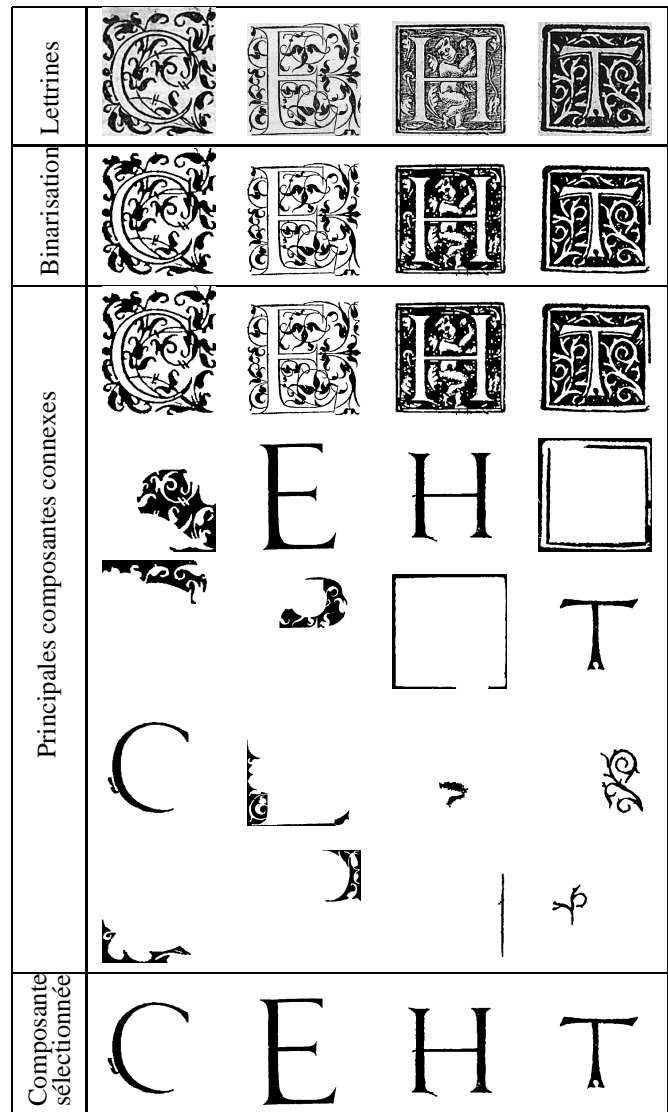
Ces traitements ne font pas l'objet de cet article mais nous permettent de re-situer le contexte de notre étude. Le tableau 4 présente un aperçu de la chaîne de traitements réalisée sur plusieurs lettrines.

A partir d'un ensemble de lettres extraites, nous avons défini une base d'apprentissage sur laquelle nous avons appliqué directement notre approche. Les tests ont ensuite été réalisés sur le reste de la base de données (50% de données d'apprentissage et 50% de données tests) pour évaluer de manière cohérente l'efficacité de notre approche.

Nous pouvons remarquer, sur cette autre base, que notre approche fournit de meilleurs résultats dans la majorité des cas (cf. tableaux 5 et 6).

Nous avons augmenté progressivement les tailles des bases (jusqu'à 45 classes de symboles) afin d'évaluer le comportement de notre méthode lorsque les corpus d'apprentissage et de reconnaissance grandissent.

Nous avons comparé notre approche à celle des ANN [1] implémentée par David D. Mount et S. Arya (le package est disponible sur internet<sup>4</sup>). L'apprentissage est effectué sur chaque classe en considérant la matrice des distances entre les symboles. Les taux de reconnaissance sont clairement plus bas que ceux obtenus en considérant l'ensemble de la base mais restent relativement élevés en utilisant notre approche (environ 88% de taux de reconnaissance et 99% de taux d'ordonnement).



TAB. 4 – Exemple d'extraction de lettres.

Descripteurs	<i>C</i>	<i>E</i>	<i>S<sub>A</sub></i>	<i>G<sub>FD</sub></i>	<i>T<sub>Rf</sub></i>	<i>FM</i>	<i>MZ</i>	<i>W<sub>S</sub></i>
Avant	49	41	70	59	50	77	40	60
Après	52	39	75	69	50	81	42	84

TAB. 5 – Taux de reconnaissance obtenus sur la base des lettres (avant/après).

Descripteurs	<i>C</i>	<i>E</i>	<i>S<sub>A</sub></i>	<i>G<sub>FD</sub></i>	<i>T<sub>Rf</sub></i>	<i>FM</i>	<i>MZ</i>	<i>W<sub>S</sub></i>
Avant	85	78	90	89	80	92	77	97
Après	93	81	97	94	87	98	79	100

TAB. 6 – Taux d'ordonnement moyens obtenus sur la base des lettres (avant/après).

<sup>2</sup><http://l3iexp.univ-lr.fr/madonne/>

<sup>3</sup>Nous remercions le Centre d'études Supérieures de la Renaissance pour la permission d'utiliser leurs documents d'archive.

<sup>4</sup>[www.cs.umd.edu/~mount/ANN](http://www.cs.umd.edu/~mount/ANN)

Les résultats donnés dans les tableaux 7 et 8 montrent une bonne robustesse de notre approche à ce facteur.

Taille de la base	T	2×T	3×T	4×T	5×T
Avant ( $W_S$ )	87	79	76	70	67
Après ( $W_S$ )	88	91	90	89	88
Ann	89	90	86	88	86

TAB. 7 – Taux de reconnaissance obtenus en considérant des bases de données de tailles (T) croissantes.

Taille de la base	T	2×T	3×T	4×T	5×T
Avant ( $W_S$ )	98	98	98	98	98
Après ( $W_S$ )	99	99	99	99	99

TAB. 8 – Taux d'ordonnement moyens obtenus en considérant des bases de données de tailles (T) croissantes.

## 6 Conclusion

Dans cet article nous avons montré que l'étude du comportement des descripteurs permet d'améliorer la reconnaissance et ceci sur des applications distinctes. Nous avons choisi d'utiliser comme référence pour nos profils la distance entre les images tests et les images d'apprentissage car nous manipulons des valeurs scalaires normées entre 0 et 1. Cependant, notre approche devrait rester valable avec d'autres références comme : une mesure de bruits par exemple. Actuellement nous cherchons à déterminer comment intégrer ces fonctions de pondération dans d'autres mécanismes de combinaison de données (règle de Dempster-Shafer, fusion possibiliste quantifiée ou non...). Par ailleurs nous envisageons d'étudier la robustesse de notre approche lors du passage à l'échelle en considérant des bases de données de plus en plus importantes. Enfin, nous recherchons aussi à spécialiser notre ensemble de descripteurs en fonction des classes de symboles recherchées en étudiant l'importance de ceux-ci lors de la reconnaissance et leurs interactions avec les autres classes.

## Références

[1] S. Arya, D.M. Mount, N.S. Netanyahu, R. Silverman and A. Wu. An optimal algorithm for approximate nearest neighbor searching. *J. ACM*, 45 :891–923, 1998.

[2] D.H. Ballard. Generalizing the Hough Transform to Detect Arbitrary Shapes. *Pattern Recognition*, 13(2) :111–122, 1981.

[3] R. Bamieh and R. de Figueiredo. A General Moments Invariants/Attributed Graph Method for the Three Dimensional Object Recognition from a Single Image, *IEEE Journal of Robotics and Automation*, 2 :240–242, 1986.

[4] A. Bhattacharya. On a measures of divergence between two statistical populations defined by their probability distributions, *Bull. Calculus Math Soc.*, 99–109, 1943.

[5] T. Bernier and J-A. Landry. A new method for representing and matching shapes of natural objects. *Pattern Recognition*, 36(8) :1711–1723, 2003.

[6] S. Derrode and F. Ghorbel. Robust and efficient Fourier-Mellin transform approximations for invariant grey-level image description and reconstruction, *Computer Vision and Image Understanding*, 84(1) :25–39, 2001.

[7] T. Gevers and A.W.M. Smeulders. Context-based image retrieval by viewpoint-invariant color indexing, *Image and Vision Computing*, 17 :475–488, 1999.

[8] F. Ghorbel. A complete invariant description for gray-level images by the harmonic analysis approach, *Pattern Recognition Letters*, 15 :1043-1051, 1994.

[9] A.A. Kassim, T. Tan and K.H. Tan. A Comparative Study of Efficient Generalised Hough Transform Techniques. *Image and Vision Computing*, 17 :737–748, 1999.

[10] H. Kauppinen, T. Seppänen and M. Pietikäinen. An Experimental Comparison of Autoregressive and Fourier-Based Descriptors in 2D Shape Classification. *IEEE Transactions on PAMI*, 17(2) :201–207, 1995.

[11] A. Khotanzad and J. Lu. Classification of invariant image representation using a neural network, *IEEE Transactions on Acoustic, Speech and Signal Processing*, 028–1190, 1990.

[12] E. Persoon and K. Fu. Shape Discrimination using Fourier Descriptors. *IEEE Transactions on Systems, Man, and Cybernetics*, 7(3) :170–179, 1977.

[13] Y. Rubner and S. Tomasi. *Perceptual metrics for image database navigation*, kluwer, Boston, 2000.

[14] Y. Rui, A. She and T.S. Huang. A Modified Fourier Descriptor for Shape Matching in MARS, *Image Databases and Multimedia Search*, ed. S. K. Chang, 8 :165–180, 1998.

[15] D. Sharvit, J. Chan, H. Tek and B. Kimia. Symmetry-based Indexing of Image Databases. *Journal of Visual Communication and Image Representation*, 12, 1998.

[16] S. Tabbone and L. Wendling. Binary shape normalization using the Radon transform, *DGCI, LNCS 2886*, Naples, Italy, 2886 :184–193, 2003.

[17] R. Teag. Image Analysis via the General Theory of Moments. *Journal of the Optical Society of America*, 70(8) :920–921, 1980.

[18] D. Zhang and G. Lu. Shape-based image retrieval using generic Fourier descriptor, *Signal Processing : Image Communication*, 17 :825–848, 2002.

[19] D. Zhang and G. Lu. Review of shape representation and description techniques, *Pattern Recognition*, 37(1) :1–19, 2004.

[20] Shafer. *A Mathematical Theory of Evidence*, Princeton University Press, Princeton, New Jersey, USA 1979.

[21] E. Osuna, R. Freund and F. Girosi. Support Vector Machines : Training and Applications, *AIM-1602*, 38, 1997.

[22] Y. Amit and D. Geman. Shape Quantization And Recognition With Randomized Trees, *Neural Computation*, 9 :1545-1588, 1997.

[23] L. Roux and J. Desachy. Information fusion for supervised classification in a satellite image, 4th IEEE Int. Conf. on Fuzzy Systems, Yokohama, Japan., 9 :1119–1124, 1995.

[24] I.S. Oh, J.S. Lee and B.R. Moon. Hybrid genetic algorithms for feature selection, *IEEE Trans Pattern Anal Mach Intell.*, 26(11) :1424-37, 2004.

[25] M. Arif, S.Hafid, T. Brouard and N. Vincent. AWFO (un opérateur d'agrégation) pour la reconnaissance des formes, 12e Conférence Reconnaissance de Forme et Intelligence Artificielle, 2006.

[26] J. Yang and V. Honavar. Feature Subset Selection Using a Genetic Algorithm, *IEEE Intelligent Systems*, 13 :44-49, 1998.

[27] M.L. Raymer, W.F. Punch, E.D. Goodman, L.A. Kuhn and A.K Jain. Dimensionality Reduction Using Genetic Algorithms, *IEEE Transactions on Evolutionary Computation*, A Publication of the IEEE Neural Networks Council, 4, 2000.