

Classification de formes par apprentissage des géodésiques de formes

Abdesslam BENZINOUE & Kamal NASREDDINE

Laboratoire des Sciences et Techniques de l'Information, de la Communication et de la Connaissance, UMR CNRS 6285 Lab-STICC
Ecole Nationale d'Ingénieurs de Brest (ENIB), 29238 BREST cedex, France
benzinou@enib.fr, nasreddine@enib.fr

Résumé – Dans un travail précédent, nous avons proposé une méthode de reconnaissance de formes (classification et recherche) basée sur le recalage et les géodésiques de formes. Dans cet article, nous y introduisons une étape supplémentaire d'estimation de pose pour l'accélération du temps de traitement. Nous montrons ensuite que les géodésiques de formes peuvent également être utilisées par des techniques de classification de type SVM ou forêts aléatoires. Les expérimentations menées sur la base de formes MPEG-7 montrent (1) une réduction du temps de calcul des attributs caractéristiques d'un facteur dix en moyenne, et (2) une amélioration du taux de bonne classification par rapport au classifieur plus proche voisin : En analyse mono-échelle : Plus proche voisin (95,05%), Forêt aléatoire (96,10%), SVM (97,66%). En analyse multi-échelle : Plus proche voisin (98,86%), Forêt aléatoire (99,02%), SVM (99,44%).

Abstract – In a previous work, we proposed a method of pattern recognition (classification and retrieval) based on signal registration and shape geodesics. In this paper, we introduce a preliminary step of pose estimation for accelerating the processing time. We then show that shape geodesics may also be used by classification techniques such as SVM and random forest algorithms. Experiments conducted on the shapes of the benchmark database MPEG-7 show (1) a reduction in computation time of the attributes of a factor of ten in average, and (2) an improvement in the classification rate compared to the nearest neighbor classifier: In single-scale analysis: Nearest Neighbor (95.05%), Random Forest (96.10%), SVM (97.66%). Multi-scale analysis: Nearest Neighbor (98.86%), random forest (99.02%), SVM (99.44%).

1 Introduction

Cet article présente une méthode de classification des formes par apprentissage des géodésiques de formes. Classiquement, le schéma d'un processus de classification supervisée de formes comporte deux étapes essentielles : la première consiste à apprendre la description de formes à partir d'une base d'exemples (apprentissage), la seconde va reconnaître une forme à partir de sa description extraite de l'image. Plusieurs travaux ont été consacrés à la définition de descripteurs de formes invariants par un certain groupe de transformations géométriques (translation, rotation, facteur d'échelle). Dans un travail précédent [1], nous avons proposé une approche de reconnaissance de formes établie sur un processus de recalage robuste entre les formes. Fondée sur les géodésiques dans l'espace des formes, cette approche exploite le résidu du recalage pour définir une métrique permettant la comparaison des formes. Nous présentons dans cet article une extension de cette méthode par l'utilisation des Séparateurs à Vaste Marge [2] ou des forêts aléatoires [3]. Nous montrons en particulier comment utiliser les distances géodésiques avec ces méthodes de classification. La deuxième extension abordée dans ce travail concerne l'accélération du processus de recalage par l'utilisation d'une technique d'estimation de pose et du point initial entre les formes à aligner.

L'approche proposée est testée sur la base d'images MPEG-7, partie B [4]. Cette base d'images est composée d'un grand nombre de formes de différents types : 70 classes de formes

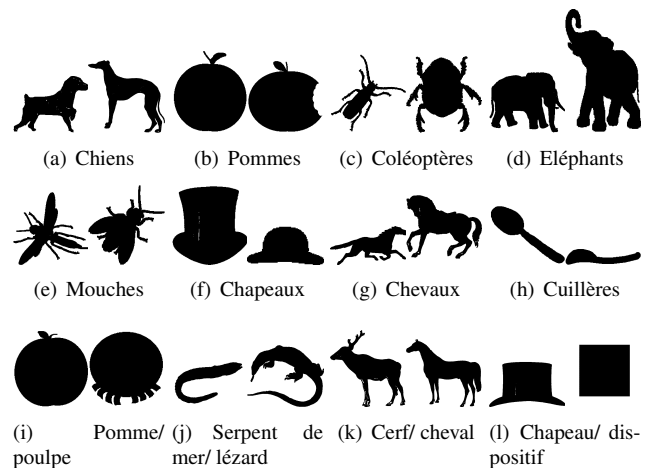


FIGURE 1 – Illustrations des difficultés de reconnaissance sur la base de test MPEG-7. (a)-(h) : Exemples de formes qui ne sont pas visuellement similaires aux autres échantillons de leur propre classe. (i)-(l) : Exemples de formes issues de classes différentes mais qui sont très similaires.

avec 20 exemples de chaque classe, pour un total de 1400 formes. Les classes comprennent des objets naturels et artificiels. La reconnaissance de formes sur cette base d'images n'est pas une affaire simple car certains échantillons de la base sont visuellement différents des autres membres de leur propre classe, et certaines formes sont très similaires à des exemples d'autres

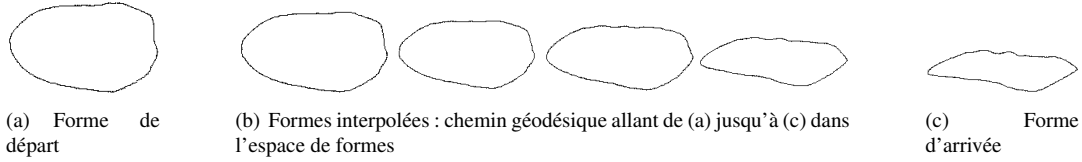


FIGURE 2 – Chemin géodésique dans un espace de formes. Le chemin de déformation représenté est celui allant de (a) à (c).

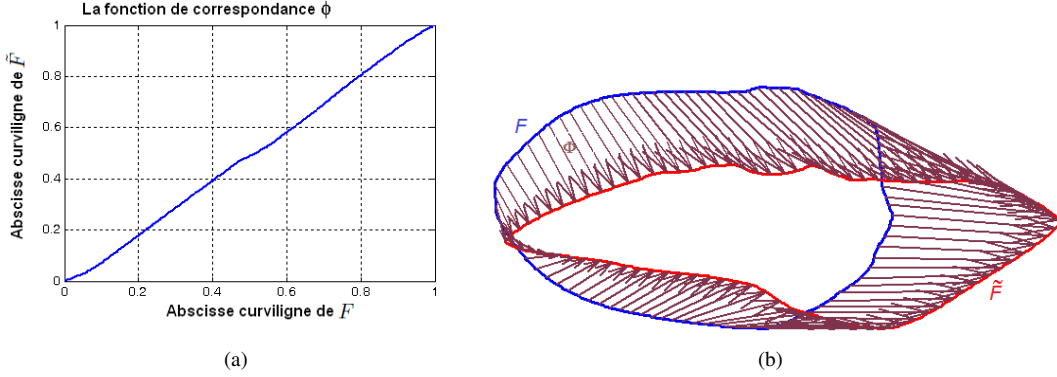


FIGURE 3 – La fonction de correspondance ϕ des deux contours F et \tilde{F} de la figure 2. A gauche, la fonction de correspondance est une fonction monotone qui fait correspondre une abscisse curviligne entre 0 et 1 sur le premier contour à l’abscisse curviligne du deuxième contour ; à droite, visualisation de la fonction de correspondance comme recalage de contours $2D$.

classes (Figure 1). Il est montré ici que cette approche surpasse toutes les approches de l’état d’art avec un taux de classification qui dépasse les 99%.

2 Métrique basée recalage et géodésiques de formes

Soient F et \tilde{F} deux formes paramétrées par les fonctions d’angle θ et $\tilde{\theta}$ respectivement. La fonction angle est définie comme étant l’angle que fait la tangente au point de la courbe avec l’axe horizontal. Notre approche de recalage robuste entre les deux formes F et \tilde{F} consiste à minimiser la fonctionnelle $E^{F\tilde{F}}(\phi)$ donnée par :

$$E^{F\tilde{F}}(\phi) = (1 - \alpha) E_D^{F\tilde{F}}(\phi) + \alpha E_R(\phi) \quad (1)$$

$$E_D^{F\tilde{F}}(\phi) = \arccos \int_{s \in [0, 1]} \sqrt{\phi_s(s)} \left| \cos \frac{\rho(r(s))}{2} \right| ds \quad (2)$$

$$E_R(\phi) = \alpha \int_{s \in [0, 1]} |\phi_s(s)|^2 ds \quad (3)$$

où ϕ désigne la fonction de correspondance, s est l’abscisse curviligne, $\phi_s = \frac{d\phi}{ds}$ et $r(s) = \theta(s) - \tilde{\theta}(\phi(s))$. La mesure de similarité $E_D^{F\tilde{F}}(\phi)$ considérée ici permet de rechercher le trajet géodésique entre les deux formes $2D$, F et \tilde{F} (Figure 2). La norme robuste $\rho(r)$ est empruntée au domaine des statistiques robustes [5] et permet d’améliorer la robustesse du recalage aux données aberrantes. α est un paramètre qui contrôle la régularité de la solution ϕ^* . Pour assurer l’invariance par rotation,

la minimisation de $E^{F\tilde{F}}(\phi)$ pour tous les choix d’origines des courbes est considérée. La figure 3 illustre le principe du recalage à partir de la fonction de correspondance ϕ .

La métrique proposée exploite le résidu du recalage de la façon suivante.

Distance mono-échelle :

$$d(F, \tilde{F}) = E_D^{F\tilde{F}}(\phi^*) \text{ où } \phi^* = \underset{\phi \in T}{\operatorname{argmin}} E^{F\tilde{F}}(\phi) \quad (4)$$

avec T l’espace des transformations admissibles (croissantes et régulières).

Distance multi-échelle :

$$d(F, \tilde{F}) = \frac{1}{N} \sum_{k=1}^N d_k(F, \tilde{F}) \quad (5)$$

où d_k est la distance définie dans l’équation (4) entre les mêmes formes à la $k^{\text{ième}}$ échelle et N le nombre d’échelles considérées. Cette distance est donc une combinaison de N distances mesurées à N échelles différentes.

3 Estimation de pose et du point initial

La minimisation de $E^{F\tilde{F}}(\phi)$ utilise un schéma itératif incrémental détaillé dans [1]. Dans cet algorithme, l’invariance par rotation permet d’aligner des formes indépendamment des points de départ. Cependant, le nombre d’itérations nécessaires à sa convergence augmente avec l’éloignement des positions initiales. Afin de réduire ce temps d’alignement, nous proposons d’estimer au préalable la pose et le point initial par une

technique de recalage rigide entre les deux formes à aligner. Contrairement au recalage, l'estimation du point initial reste un problème peu étudié dans la littérature. Cette tâche est souvent réalisée par des procédures d'optimisation complexes, suivies dans certains cas par une correction manuelle. Notre technique d'estimation de pose et de point initial utilise un algorithme décrit dans [6] pour l'alignement rigide d'une paire de formes 2D. Nous avons expérimenté cette méthode sur la base de test MPEG-7, partie B, dont les objets peuvent avoir des déformations assez importantes. En moyenne, le temps d'alignement est amélioré de 90%, soit une réduction d'un facteur 10 par rapport au schéma initial sans estimation préalable de pose et de point initial.

4 Apprentissage des distances géodésiques

On suppose avoir un ensemble de formes classées, (F_i, C_i) , où F_i est la forme du $i^{\text{ème}}$ échantillon de la base et C_i la classe correspondante. La classification d'une nouvelle forme F est issue du classifieur employé en fonction de la distance d considérée. En plus de la méthode des k-plus proches voisins (k-PPV), nous voulons montrer que les géodésiques peuvent également être utilisées par des techniques de classification de type SVM [2] ou forêts aléatoires [3]. Nous proposons ici d'utiliser les distances géodésiques calculées entre les formes deux à deux comme attributs d'apprentissage dans l'algorithme SVM, ou comme caractéristiques d'éclatement des feuilles dans l'algorithme de la forêt aléatoire. Dans un premier temps, la matrice des distances est directement utilisée comme matrice de caractéristiques dans les trois classifieurs k-PPV, SVM et forêt aléatoire. Nous avons utilisé la technique d'ajustement par validation croisée (*deux tiers-un tiers*) pour la sélection du modèle et l'optimisation des hyper-paramètres du classifieur.

Le procédé de reconnaissance doit être évalué sur des images non incluses dans la base d'apprentissage. Un test de validation croisée *un-sorti* (aussi appelée *leave-one-out*) est effectué sur la base des images disponibles. Les images sont extraites une à une de la base d'images afin d'être utilisées à tour de rôle comme image à classifier.

5 Résultats et discussion

Afin de comparer l'approche proposée à l'état de l'art, des méthodes de classification de formes en vision par ordinateur, nous procédons à une évaluation des performances à partir de la base de test MPEG-7, partie B [4]. La représentation de forme est faite par échantillonnage de points également distants le long du contour. L'échelle est considérée comme liée à la résolution de l'échantillonnage de la forme. Les échantillonnages de forme à différentes échelles avec 32, 48, 64 et 192 points sont considérés.

Dans les tests de classification par forêt aléatoire, le taux de réussite atteint 96,1% en analyse mono-échelle et 99,02% en analyse multi-échelle. En figure 4 nous avons donné le taux

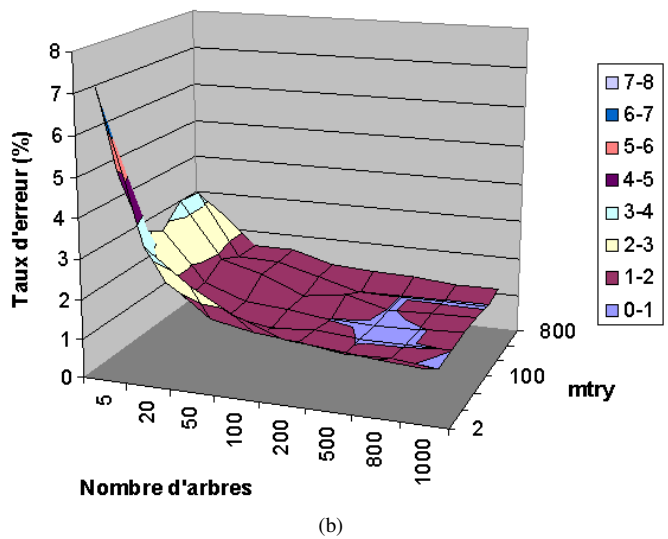
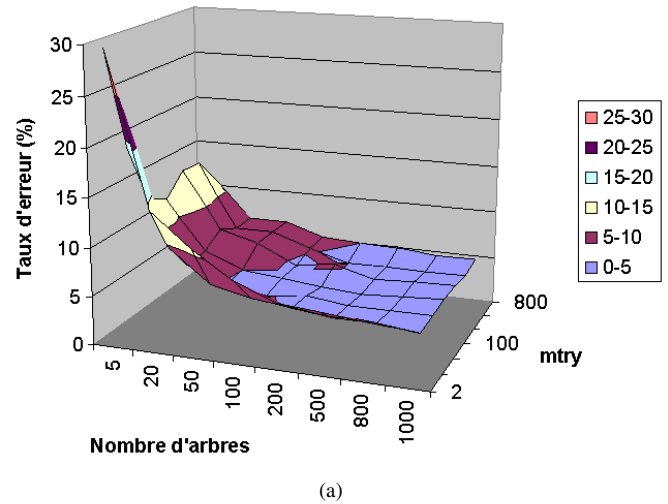


FIGURE 4 – La variation du taux d'erreur de classification par forêt aléatoire. (a) : En analyse mono-échelle. (b) : En analyse multi-échelle.

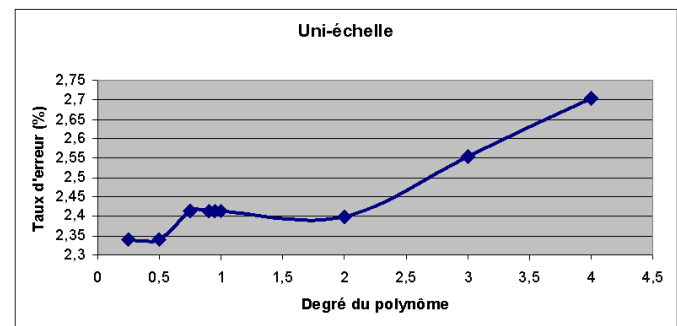


FIGURE 5 – La variation du taux d'erreur de classification en analyse mono-échelle par SVM à noyau polynômial en fonction du degré d du polynôme.

Tableau 1 – Les différents taux de réussite obtenus selon la distance et le type de classification utilisés sur la base MPEG-7.

	Distances géodésiques			Descripteurs de Fourier
	Plus proche voisin	Forêt aléatoire	SVM	Plus proche voisin
Distance mono-échelle	95,05%	96,10%	97,66%	82% [7]
Distance multi-échelle	98,86%	99,02%	99,44%	95,5% [8]

d’erreur de classement en fonction des deux paramètres, N le nombre d’arbres, et $mtry$ le nombre de variables d’essai choisies par hasard à chaque nœud. En analyse mono-échelle (figure 4(a)), le taux d’erreur minimal correspond à un nombre d’arbres $N = 500$ et à $mtry = 150$. Ce taux descend à 3,9%, ce qui correspond à un gain de 1% par rapport à la classification par l’algorithme du plus proche voisin développée dans [1]. En analyse multi-échelle (figure 4(b)), le taux d’erreur minimal descend à 0,98%, ce qui correspond à un gain de 0,16% par rapport à la classification par l’algorithme du plus proche voisin.

Dans les tests de classification par SVM, le taux de réussite est maximal avec un noyau polynomial d’ordre 0,5. Ce taux atteint 97,66% en analyse mono-échelle et 99,44% en analyse multi-échelle. En figure 5, nous avons donné le taux d’erreur en fonction du degré du polynôme. Le taux d’erreur minimal correspond à un ordre égal à 0,5. Ce taux descend à 2,34%, ce qui correspond à un gain de 2,61% par rapport à la classification par l’algorithme du plus proche voisin et 1,56% par rapport à la classification par forêt aléatoire. En analyse multi-échelle, le taux d’erreur de classement avec $d = 0,5$ est égal à 0,56%, ce qui correspond à un gain de 0,58% par rapport à la classification par l’algorithme du plus proche voisin et 0,42% par rapport à la classification par forêt aléatoire.

Nous avons récapitulé dans le tableau 1 les différents résultats obtenus, selon la distance et le type de classification utilisés.

6 Conclusion et perspectives

L’approche proposée surpasse toutes les approches reportées dans la littérature avec un taux de bonne classification de 99,44% en multi-échelle et 97,66% en mono-échelle. De notre point de vue, la supériorité de la méthode proposée par rapport aux autres méthodes de l’état d’art est due à la pertinence de la mesure de similarité utilisée, fondée sur le recalage et les géodésiques de formes. La métrique qui en résulte prend en compte la correspondance des points de caractéristiques similaires.

Comme perspective, nous allons utiliser la distance géodésique proposée à la place de la distance euclidienne dans un noyau SVM [9].

Références

[1] K. Nasreddine, A. Benzinou, and R. Fablet, “Variational shape matching for shape classification and retrieval,” *Pat-*

tern Recognition Letters, vol. 31, pp. 1650–1657, 2010.

[2] V. Vapnik, *The nature of statistical learning theory*. New York, USA : Springer-Verlag, 1995.

[3] L. Breiman, “Random forests,” *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

[4] S. Jeannin and M. Bober, *Description of Core Experiments for MPEG-7 Motion/Shape*. MPEG7, ISO/IEC JTC1/SC29/WG11 N2690, document N2690, Seoul, 1999.

[5] P. Huber, *Robust statistics*. New York : John Wiley and Sons, 1981.

[6] J. S. Marques and A. J. Abrantes, “Shape alignment : optimal initial point and pose estimation,” *Pattern Recognition Letters*, vol. 18, pp. 49–53, January 1997.

[7] M. Nixon and A. Aguado, *Feature extraction and image processing*. London, UK : Academic Press, 2007.

[8] C. Direkoglu and M. Nixon, “Shape classification using multiscale Fourier-based description in 2-D shape,” in *9th IEEE International Conference on Signal Processing*, vol. 1, 2008, pp. 820–823.

[9] B. Haasdonk and C. Bahlmann, “Learning with distance substitution kernels,” in *Pattern Recognition*, ser. Lecture Notes in Computer Science, C. Rasmussen, H. Bülthoff, B. Schölkopf, and M. Giese, Eds. Springer Berlin Heidelberg, 2004, vol. 3175, pp. 220–227.