

## UN DESCRIPTEUR DE FORMES INVARIANT PAR ROTATION ET CHANGEMENT D'ECHELLES

C. OLIVIER, K. ROMEO-PAKKER, A. RABIHA

La3i, U.F.R. des Sciences, Université de Rouen, 76821 Mont Saint Aignan cedex, France.

### RESUME

Il existe plusieurs caractéristiques invariantes par rotation, changement d'échelle, contribuant à la reconnaissance de formes telles que des caractères imprimés ou manuscrits. Ces descripteurs comme les moments invariants ou les descripteurs de Fourier, ne sont, en général, que des éléments d'un vecteur de caractéristiques et il convient donc que ceux-ci ne soient pas trop pénalisants en place mémoire. Nous proposons dans ce papier une méthode basée sur un suivi de contour que nous allons appliquer aux caractères dits "bâtons" ou "droits". Le descripteur présenté est obtenu par construction d'un graphe de directions cumulées qui est normalisé et réduit à un nombre de classes minimal par un critère d'information.

### 1- INTRODUCTION

Le domaine de la reconnaissance de formes utilise différentes méthodes invariantes par rotation, changements d'échelle, symétrie. On peut citer les moments invariants de Hu ou de Zernicke [1], les descripteurs de Fourier ou mieux de Mellin-Fourier [2], [3], les descripteurs topologiques [4], les descripteurs circulaires concentriques [5]. Les techniques de suivi de contours sont aussi utilisées mais leur principal défaut est qu'ils sont très sensibles au bruit dû à la numérisation de l'image d'où des contours parfois très différents pour des formes très voisines. Une description par contours obtenue par un codage à partir des directions de Freeman, s'il permet un lissage des formes, est en général difficile à analyser à cause de la trop grande variabilité de la longueur de la chaîne codée à comparer.

Dans ce papier, nous proposons une méthode utilisant un algorithme de suivis cumulés de contours externes à partir d'un codage de Freeman périodique ce qui autorise des décalages et une normalisation des graphes de suivis. Ces graphes sont invariants par rotation et changement d'échelles. Ils ne tiennent donc pas compte de l'intensité interne de la forme comme pour les moments ou Mellin-Fourier. La description et l'obtention de ces graphes seront présentées en partie II. Pour réduire la taille du vecteur induit de chaque graphe, nous utilisons un critère d'information de type Akaike[6], permettant de résumer celui-ci à un histogramme au nombre de classes et à la forme de ces classes optimaux au sens d'un coût de type distance de Hellinger entre distributions [7]. C'est l'objet de la partie III. Nous validons notre méthode sur des caractères "bâtons" de

### ABSTRACT

There are many rotation and scale change invariant features for the description of handwritten or printed characters. Generally, these features, like the moment invariant or Fourier descriptors, are elements of a description vector and as such cannot occupy too much memory space. We propose in this paper a new method based on the contour following algorithm which will be applied on handwritten segmented letters written with constraints. The features are presented with a graph of a cumulative distribution of Freeman codes on the external contourline of analyzed characters. This graph is normalized and the number of classes on the x-axis is optimally reduced by an information criterion.

tailles et d'orientations différentes dans un tableau présenté dans une dernière partie.

### 2- PRINCIPE DU SUIVI DE CONTOURS

Nous allons expliquer dans ce paragraphe, les différentes étapes qui nous ont permis d'élaborer une nouvelle méthode de reconnaissance de formes invariantes à la rotation et au changement d'échelles. La représentation présentée, basée initialement sur un codage de Freeman, donne des informations (concavités et convexités) pertinentes sur la forme à reconnaître.

#### 2-1- Prétraitement

Les images que nous traitons sont des caractères scannés; il est donc inutile de passer par les différentes phases de prétraitements comme la binarisation puisqu'elles sont en noir et blanc (caractères noirs sur fond blanc). Toutefois, ces images sont généralement bruitées: présence de nombreuses imperfections dues aux artefacts de l'acquisition et souvent à la qualité du document. Il apparaît donc nécessaire d'effectuer un lissage sur l'image afin d'améliorer sa qualité et d'obtenir un contour plus régulier. Ainsi, nous lissons les contours préalablement par masques 3x3 permettant un ébarbulage du tracé.

#### 2-2- Suivi de contour



Il s'agit d'une méthode incrémentale sur le suivi de la frontière qui sépare un tracé (objet connexe) des autres qui l'entourent (dans notre cas, le fond de l'image).

Le suivi de contour est basé sur l'approximation angulaire la plus simple: le code de Freeman. Elle consiste à coder chaque vecteur entre deux points successifs par un chiffre allant de 0 à 7, représentant les huit directions de Freeman. Ce codage est une représentation exacte de chaque pixel du contour relativement à ses voisins. Le codage global du contour d'un caractère est donc la succession de tous les codes correspondants aux directions prises pour suivre le contour. On effectue un seul suivi de contour: celui du contour externe. L'origine du suivi est le pixel le plus haut et le plus à gauche du caractère, et l'image est balayée de haut en bas et de gauche à droite. Cette méthode, étroitement liée à un repère, est donc très sensible à l'orientation de l'objet codé. C'est pourquoi nous devons traiter ces codages globaux afin de les rendre invariants à l'origine du suivi et à l'orientation.

### 2-3- Contour cumulé

En chaque point du contour, une variable entière (de 0 à 7) est mémorisée. Nous considérons alors les variations relatives du code de Freeman sur le contour. Ainsi, on incrémentera une variable entière du nombre de directions dont on a tourné dans le sens trigonométrique positif et on décrémentera dans le sens négatif. Un ajustement s'avère évidemment nécessaire pour certaines variations dues au passage 7 à 0 par exemple. On construit ainsi un graphe, où à chaque pixel du contour (en abscisse) est associé un entier relatif (en ordonnée). La figure 1 montre quatre représentations de "U" issues d'une base de toponymes cartographiques du laboratoire La3i, ainsi que les graphes de contours correspondants.

Ces graphes reflètent bien les concavités, mais trop dépendants du nombre de points et de l'orientation du caractère, leurs différences n'autorisent pas encore une possible classification.

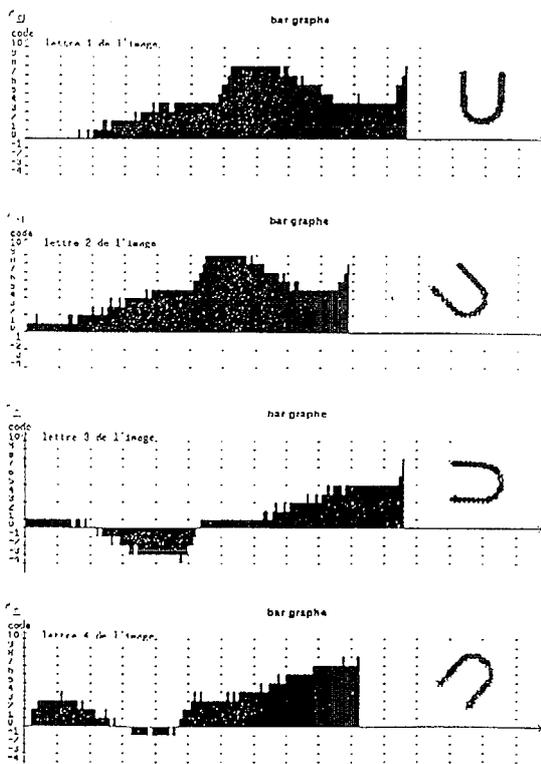


Figure 1: Exemple de graphes de contours.

### 2-4- Graphe cumulé invariant

L'invariance par rotation est obtenue en deux étapes:

- Recherche de maxima locaux: un maximum local est difficilement détectable sans lissage (voir figure 1). Celui-ci est effectué par moyennage sur 3 points successifs. Sur notre base d'apprentissage, on a ensuite pu empiriquement décider qu'un maximum local sera un point de valeur supérieure à toutes les autres sur un intervalle  $[-1/5 m, +1/5 m]$  où  $m$  est le nombre de points du graphe.

- Recalage des graphes: pour un graphe, la plus grande distance entre deux maxima locaux est notée. Pour une forme donnée représentée par différents graphes cumulés, nous repérons le maximum correspondant à la plus courte de ces distances et nous recalons tous les graphes à partir de ce point. Cette manipulation, correspondant à un changement d'origine des graphes, est l'application d'une simple rotation sur les différentes représentations de la forme à apprendre.

L'invariance au changement d'échelle est obtenue par normalisation des nouveaux graphes.

La figure 2 montre les graphes cumulés invariants des quatre "U" de la figure 1. Le nombre de codes (les points) est normalisé à  $m = 100$ . Nous avons ainsi en abscisse  $m$  points  $x_i$  d'ordonnées  $m$  variables cumulées  $y_i$ .

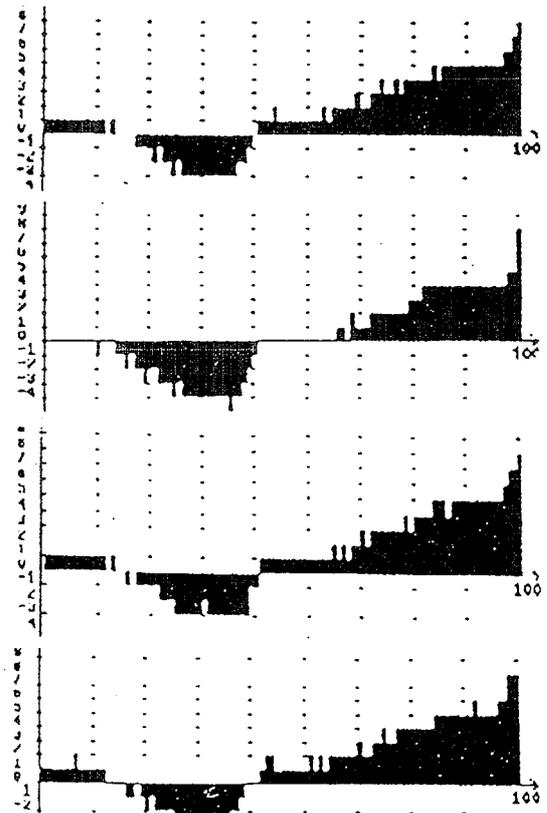


Figure 2: Graphes cumulés invariants non réduits de l'exemple précédent.

### 3- REDUCTION DE DONNEES

Nous avons construit pour chaque classe de caractères une caractéristique formée d'un vecteur de  $m$  coordonnées. Comme on le voit sur la figure 2, l'invariance à la rotation est globalement vérifiée mais un lissage s'impose, avant comparaison des graphes, pour éliminer quelques pics non significatifs. Un premier lissage par moyennage de 4 valeurs successives est proposé en figure 3 sur trois "2" de tailles et

orientations différentes. Cet exemple confirme les invariances par rotation et changement d'échelle.

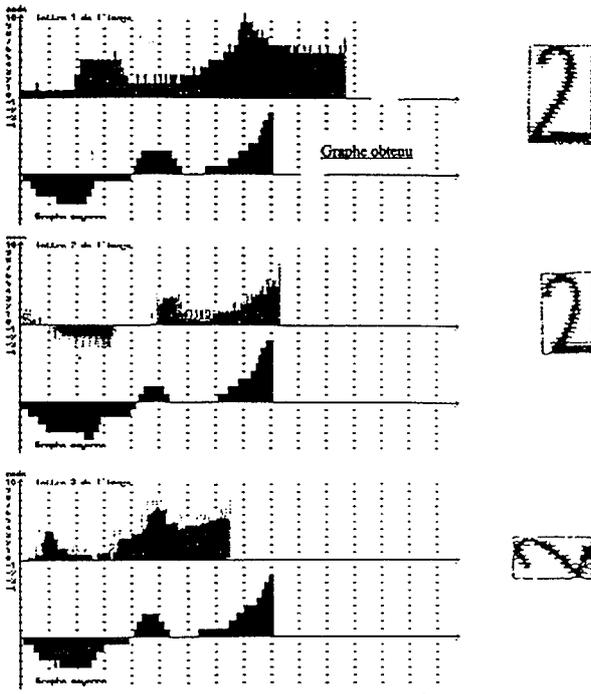


Figure 3: Exemple de graphes cumulés lissés d'une même forme.

Le lissage précédent conserve le même nombre de classes ( $m = 100$ ). C'est pourquoi nous nous proposons d'utiliser une méthode élaborée initialement dans le cadre de l'approximation de distributions par des histogrammes [7] pour réduire ce nombre de classes de manière optimale suivant un coût.

Le critère utilisé est proche par sa forme du critère d'information AIC d'Akaike [8] initialement élaboré pour la recherche de l'ordre de modèles linéaires AR. Il est issu d'une première approximation au sens du maximum de vraisemblance d'une loi par un histogramme puis de l'approximation asymptotique d'un coût défini à partir de la distance de Hellinger entre deux lois de probabilité, et non d'un contraste de Kullback comme pour l'ordre de modèles. Le terme d'entropie est inchangé, bien qu'appliqué à des histogrammes, et la pénalisation, de la forme du critère de type BIC (Bayesian Information Criterion) utilisé pour les modèles ARMA. Le critère proposé est aussi voisin du MDL (Minimum Description Length) de J. Rissanen que celui-ci a déterminé à partir de la notion de complexité stochastique [9].

Nous ne rentrerons pas dans les détails théoriques:

Le graphe initial est normalisé, c'est à dire que le point  $x_i$  a pour ordonnée  $y_i/N$ , où  $N$  est la surface limitée par le graphe entre  $x_1$  et  $x_m$  et l'axe des abscisses.

A partir de la partition initiale  $M$  (le graphe à  $m$  points des figures 2), nous cherchons le nombre de classes  $c$ ,  $c < m$ , et la forme de ces  $c$  classes qui minimisent le critère:

$$AIC^*(c) = c \frac{\text{Log}N}{N} - 2 \sum_{B \in C} \theta_c(B) \text{Log} \frac{\theta_c(B)}{\mu(B)}$$

où  $C$  est la partition de l'histogramme cherché à  $c$  classes,  $\theta_c(B)$  est la mesure normalisée d'une classe  $B$  de  $C$ ,  $\mu$  une loi

*a priori* (initialement choisie uniforme:  $\forall B \in M: \mu(B) = 1/m$ ).

A partir de  $m$ , la partition à  $(m-1)$  classes minimisant  $AIC^*(m-1)$  est recherchée par fusion de classes adjacentes puis une partition à  $(m-2)$  classes, etc ..., puis une partition à  $(k-1)$  classes tant que  $AIC^*(k) < AIC^*(k+1)$ , jusqu'à l'obtention de l'entier  $c$  et de la partition  $C$  telle que:

$$c = \arg \min_{k=1 \dots m} (AIC^*(k)).$$

La figure 4 montre le comportement de  $AIC^*$  sur le graphe cumulé du premier "U" de la figure 2, à partir d'un algorithme de fusion des classes donné en [7]. Sur notre exemple,  $c = 6$ .

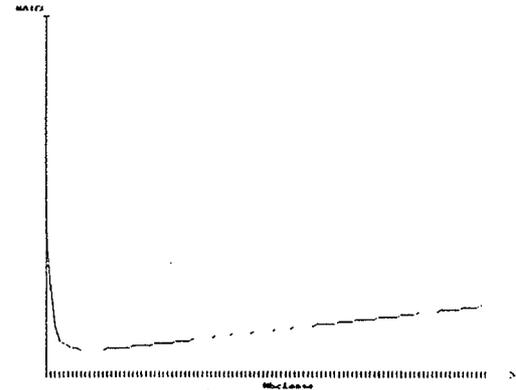


Figure 4: Exemple de comportement de  $AIC^*$

Enfin, nous montrons en figure 5 à partir de l'exemple des "U", les quatre histogrammes définitifs obtenus représentant ces quatre formes. Le nombre  $c$  retenu varie de 6 à 10 soit une réduction de données du vecteur graphe cumulé de 6 à 12, sans perte notable d'information.

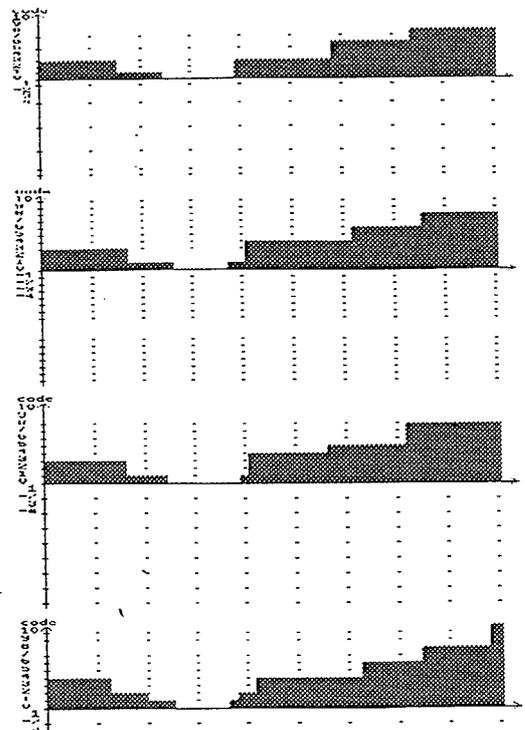


Figure 5: Graphes cumulés réduits par  $AIC^*$  des figures 1 et 2.



#### 4- UN OUTIL POUR LA CLASSIFICATION

Pour montrer l'intérêt de la méthode, nous avons comparé deux à deux l'ensemble des graphes réduits obtenus à partir de l'ensemble des caractères de notre base d'apprentissage. La distance choisie entre histogrammes est une mesure de Kolmogorov et un indice I classique de dissemblance est calculé. Un seuil est fixé empiriquement à  $I_0 = 0.18$ , les valeurs moyennes intra-classes variant de 0.05 à 0.12, sauf pour la classe "I" (0.17) mais celle-ci (tableau 1) est fort dissemblable des autres classes. Les résultats présentés au tableau 1 sont les valeurs moyennes inter-classes des indices I calculés entre graphes cumulés réduits par AIC\* sur notre base d'apprentissage. Ils permettent de conclure à la validité de la méthode et aux limites de celle-ci (inhérente à l'analyse externe des formes): confusion des formes: "C", "J", "U", "A" et "Y", "B", "O", "Q" et "D", "E" et "W"...

En phase de reconnaissance, la démarche peut être la suivante: un caractère à reconnaître, représenté par son graphe réduit par AIC\*, ne sera pas dans une classe si l'indice de vraisemblance avec celle-ci est supérieur à  $I_0$ , à chaque classe étant associé un graphe cumulé réduit moyen caractéristique. Sinon, recherche de la distance de Kolmogorov minimale et étiquetage.

#### Conclusion

La méthode présentée n'est qu'une contribution au problème de la reconnaissance de caractères. La caractéristique proposée, définie à partir de contours cumulés réduits par une fusion de données adaptée (critère d'information), peu coûteuse en place mémoire, et robuste pour des caractères particuliers (caractères "bâtons"), ne peut évidemment suffire, seule, à la reconnaissance de formes mais sera intégrée dans un vecteur composé d'autres primitives invariantes aux similitudes, comme le nombre d'occlusions, les moments, etc...

Nous avons pu réduire de plus par le critère présenté un vecteur caractéristique d'un facteur variant de 5 à 10. Ce critère pourrait aussi bien permettre la discrétisation de caractéristiques continues ou la réduction de données sur des primitives de formes telles que les projections de pixels, par exemple.

#### Eléments bibliographiques

- [1]: El Khali F., Sid-Ahmed M.A.: *Machine recognition of optically captured machine printed arabic text*. Pattern Recognition, **23**, n°11, 1990, pp 199-217.
- [2]: Chang C.C., Hwang S.M., Buehrer D.J.: *A shape recognition scheme based on relative distances of feature points from the centroid*. Pattern Recognition, **24**, n°11, 1991, pp 1053-1063.
- [3]: Fonga-Tchinkante H.: *Analyse harmonique sur les groupes et reconnaissance des formes*. Thèse de doctorat, INPG, 1992.
- [4]: Lin Y., Dou J., Wang H.: *Contour shape description based on an arch height function*. Pattern Recognition, **25**, n°1, 1992, pp 17-23.
- [5]: Kita N.: *Object locating based on concentric circular description*, 11th IAPR, **1**, 1992, pp 660-663.
- [6]: Colot O., Olivier C., Courtellemont P.: *Information criteria and abrupt changes of probability laws*. EUSIPCO-Signal Processing VII, n°3, 1994, pp 1855-1858.
- [7]: Olivier C., Courtellemont P., Colot D., De Bruçq D.: *Comparison of histogram: a tool for detection*. Eur. J. of Diagnosis & Safety in Auto., **4**, n°3, 19904, pp 335-355.
- [8]: Akaike H.: *A new look at the statistical model identification*. IEEE Trans. Auto. Contr., **19**, 1974, pp 716-723.
- [9]: Rissanen J., Speed T.P., Yu B.: *Density estimation by stochastic complexity*. Trans. Info. Theory, **38**, n°2, 1992, pp 315-323.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	
0,09	0,41	0,29	0,44	0,31	0,20	0,39	0,19	0,64	0,23	0,21	0,37	0,24	0,23	0,19	0,28	0,16	0,18	0,30	0,18	0,30	0,21	0,32	0,27	0,07	0,30	A
	0,07	0,59	0,10	0,64	0,51	0,52	0,47	0,27	0,52	0,51	0,24	0,45	0,43	0,29	0,18	0,31	0,36	0,62	0,35	0,47	0,44	0,67	0,26	0,39	0,31	B
		0,10	0,63	0,38	0,25	0,18	0,31	0,75	0,11	0,32	0,48	0,28	0,36	0,37	0,46	0,36	0,37	0,31	0,35	0,24	0,24	0,36	0,44	0,31	0,46	C
			0,13	0,67	0,52	0,75	0,49	0,30	0,56	0,53	0,28	0,47	0,45	0,32	0,20	0,34	0,39	0,64	0,37	0,50	0,46	0,70	0,32	0,40	0,33	D
				0,06	0,22	0,36	0,30	0,85	0,39	0,23	0,62	0,37	0,33	0,43	0,51	0,40	0,35	0,26	0,44	0,51	0,46	0,17	0,42	0,35	0,44	E
					0,07	0,31	0,14	0,71	0,25	0,18	0,44	0,24	0,19	0,31	0,37	0,27	0,31	0,20	0,30	0,34	0,29	0,27	0,33	0,22	0,29	F
						0,09	0,39	0,88	0,26	0,39	0,63	0,37	0,43	0,49	0,58	0,47	0,40	0,33	0,46	0,40	0,37	0,36	0,57	0,42	0,55	G
							0,08	0,67	0,35	0,17	0,41	0,17	0,10	0,30	0,38	0,26	0,32	0,19	0,28	0,32	0,29	0,30	0,30	0,23	0,22	H
								0,17	0,67	0,73	0,30	0,62	0,64	0,52	0,41	0,54	0,59	0,82	0,52	0,54	0,56	0,88	0,49	0,60	0,48	I
									0,09	0,26	0,40	0,23	0,29	0,29	0,39	0,28	0,29	0,36	0,28	0,17	0,15	0,37	0,37	0,25	0,39	J
										0,10	0,46	0,29	0,18	0,27	0,39	0,25	0,27	0,22	0,32	0,35	0,30	0,25	0,28	0,20	0,29	K
											0,12	0,37	0,37	0,33	0,20	0,31	0,40	0,57	0,25	0,31	0,28	0,65	0,29	0,33	0,24	L
												0,09	0,19	0,28	0,35	0,25	0,31	0,30	0,21	0,27	0,25	0,35	0,35	0,27	0,27	M
													0,08	0,29	0,35	0,25	0,33	0,21	0,31	0,32	0,29	0,33	0,27	0,23	0,19	N
														0,06	0,23	0,06	0,14	0,42	0,19	0,33	0,27	0,43	0,18	0,18	0,28	O
															0,07	0,21	0,27	0,51	0,19	0,34	0,31	0,55	0,26	0,25	0,26	P
																0,12	0,14	0,38	0,16	0,31	0,25	0,41	0,17	0,15	0,26	Q
																0,10	0,41	0,21	0,41	0,32	0,35	0,27	0,20	0,34	R	
																	0,09	0,42	0,44	0,41	0,32	0,43	0,35	0,39	S	
																		0,11	0,29	0,25	0,45	0,31	0,18	0,31	T	
																			0,08	0,15	0,52	0,38	0,29	0,39	U	
																				0,07	0,46	0,34	0,21	0,33	V	
																					0,10	0,45	0,36	0,46	W	
																						0,06	0,26	0,23	X	
																							0,06	0,29	Y	
																								0,10	Z	

Tableau 1: Comparaison inter-classe des graphes cumulés réduits.