CONSTRUCTION D'UN GRAPHE STRUCTUREL REPRESENTATIF D'UNE FORME

Patrice DARGENTON, Nicole VINCENT, Hubert EMPTOZ
INSA LYON, LISPI-RF&D[1] (FRANCE)

I CONSTRUCTION DU GRAPHE STRUCTUREL 1
1°) SQUELETTISATION DE LA FORME 1
2°) DÉTECTION DES ARCS DU SQUELETTE 2
3°) TRAITEMENT DES ARCS 2
II EXPLOITATION DU GRAPHE STRUCTUREL 2
BIBLIOGRAPHIE 4

Dans le domaine de la reconnaissance omniscripteur des mots manuscrits cursifs (hors ligne), on distingue deux approches principales :

La première approche concerne les méthodes qui consistent, à partir d'une hypothèse lexicale, obtenue par une reconnaissance globale du mot, à vérifier la segmentation et la reconnaissance des lettres (méthodes descendantes) ;

La seconde approche concerne les méthodes analytiques (méthodes ascendantes) qui se basent sur des hypothèses de reconnaissance de lettres pour aboutir à la reconnaissance du mot.

La critique essentielle que l'on peut adresser à la première approche est que la génération d'hypothèse sur le mot n'exploite pas suffisamment la notion de lettre (l'extraction de primitives est globale) alors que dans le cas des documents imprimés, la seule reconnaissance des lettres suffit pour aboutir à la reconnaissance complète du mot.

La deuxième approche, quant à elle, conduit souvent à une indécision car, en présence des fortes variations propres aux lettres manuscrites, naturellement ambiguës, on emploie une description trop approximative des lettres. En outre, il est difficile de segmenter les lettres avant de les avoir reconnues.

Partant de ce constat, nous nous efforçons de réduire le plus possible l'incertitude au niveau de la lettre, grâce à une reconnaissance de type structurel. Dans cet article, nous proposons une méthode robuste de construction d'un graphe structurel représentatif d'une forme. Notre objectif est de réduire l'information contenue dans l'image au minimum nécessaire pour décrire précisément la lettre manuscrite.

I CONSTRUCTION DU GRAPHE STRUCTUREL

La méthode comporte trois étapes :

1°) Squelettisation de la forme

Trois étapes lui sont consacrées :

a) le calcul de la distance minimale de chaque pixel de la forme à son contour
b) la détermination du noyau de la forme à partir de cette distance au bord
c) l'amincissement du noyau en squelette par suivi de contour.

L'épaisseur du trait de la forme peut être paire ou impaire : si elle est exclusivement impaire, le noyau a une épaisseur de un pixel et est alors confondu avec le squelette, sinon, le noyau a une épaisseur de deux pixels au plus.

[1] Bat. 403, 20 Av. A. EINSTEIN, 69621 VILLEURBANNE, Tel (33) 72 43 80 96, Fax (33) 72 43 85 29

2°) Détection des arcs du squelette

On utilise une nouvelle fois la distance au bord pour définir un voisinage pour chaque pixel du squelette. En balayant toute l'image, on cherche, en suivant un ordre fixé à l'avance, à relier les points à leur premier plus proche voisin seulement. Lorsque tous les points ont été examinés, on dispose d'un ensemble d'arcs qui parcourent tout le squelette, et sur lequel on va effectuer les opérations suivantes :

3°) Traitement des arcs

a) Détermination du voisinage des arcs

Pour chaque arc, on recherche et on mémorise ses voisins, ceux dont au moins un pixel appartient à un voisinage (toujours défini par la distance au bord) d'un des pixels de l'arc courant.

b) Fragmentation des arcs

Tous les arcs vont maintenant être fragmentés au niveau de chacun de leurs voisins, excepté les barbules, là où la distance entre l'arc et son voisin est la plus petite ; cette étape comprend également la fragmentation des arcs contenant une boucle, sur lesquels on isole la partie cyclique du reste de l'arc.

c) Ebarbulage des petits arcs

Maintenant, on est en mesure de distinguer pour chaque arc, les voisins du début de ceux de la fin, ce que l'on ne pouvait faire avant car les voisins pouvaient se situer au milieu d'un arc.

L'ébarbulage consiste à éliminer les petits arcs qui ont une extrémité isolée (sans voisin).

d) Fusion des fragments d'arc

Cette étape consiste à fusionner deux arcs exclusivement voisins à une extrémité, afin d'obtenir le minimum d'arcs représentant la totalité de la forme.

Les arcs restant sont ensuite polygonisés et on enregistre chaque noeud avec le nombre d'arcs qui y sont connectés et leurs points de contrôle, ainsi que diverses informations utiles comme les angles, les longueurs, la distance des noeuds au bord, la position de chaque noeud par rapport au centre de gravité de la forme...

Enfin, pour compléter le codage, nous y ajoutons le contour polygonisé de la forme car nous estimons que cette information est complémentaire à sa structure.

Le graphe structurel obtenu (figure 1) offre une bonne description schématique de la lettre avec une faible perte d'information, tout en restant relativement insensible aux multiples variations de l'écriture manuscrite.

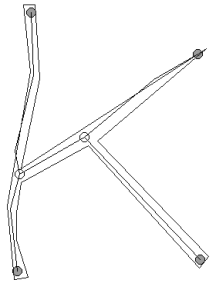

Figure 1

II EXPLOITATION DU GRAPHE STRUCTUREL

L'exploitation du graphe structurel obtenu commence par la recherche d'un isomorphisme optimal entre deux graphes, celui de la lettre à reconnaître et celui d'une des lettres de l'alphabet de référence, afin d'établir des points de comparaison. Ce problème est proche de la stéréovision [BACKER & GERBRANDS, 1987] dans lequel il s'agit de trouver la meilleure mise en correspondance entre les graphes des deux images de la vue stéréoscopique.

Dans une approche utilisant une grammaire de graphe [WONG, 1987], le graphe traduit des relations de proximité ("à coté", "au dessus"...) entre les segments de l'image, tandis que dans notre cas, la signification du graphe est une expression directe de la connectivité d'un ensemble de points extrêmes de l'image. On ne peut formuler le problème de la recherche de l'isomorphisme optimal de graphe comme une recherche dans un arbre [YOU & WONG, 1984], car cette procédure n'est pas

utilisable dans le cas où le nombre de noeuds des deux graphes à apparier est différent. En effet, cette recherche implique des fusions de noeuds et d'arcs, ainsi que des fragmentations, éliminations et concaténations d'arcs du graphe.

Nous avons donc conçu un algorithme original d'appariement de deux graphes structurels quelconques procédant par recherche de similarités entre les arcs, ce qui nous a permis de montrer que

- d'une part, la quantité d'information contenue dans le graphe offre un bon compromis entre la précision de la description et la tolérance aux déformations que l'on observe dans l'écriture manuscrite (figure 1) ;

- d'autre part, malgré la complexité du problème, l'appariement structurel permet d'évaluer une mesure de ressemblance entre les graphes qui rend compte des ambiguïtés naturelles entre les lettres (figures 2a, 2b et 2c) ;

- enfin, la fusion d'une forme à reconnaître avec une forme apprise (figure 3) modélise en quelque sorte la fonction d'interprétation qui nous semble essentielle au processus de la lecture.

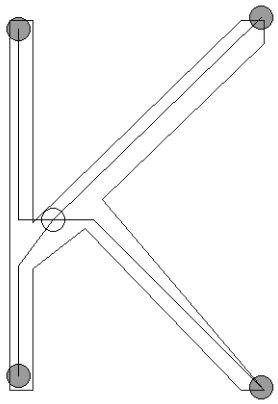

Figure 2a

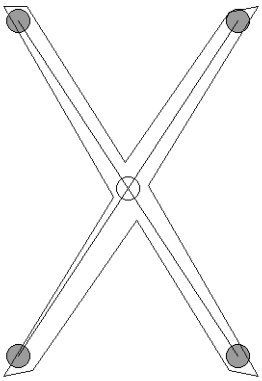

Figure 2b

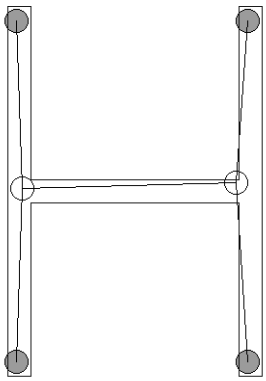

Figure 2c

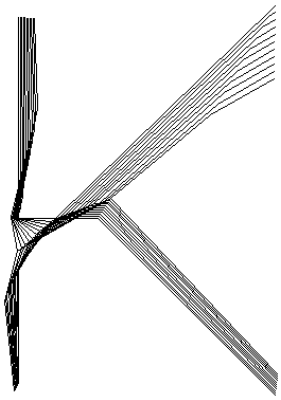

Figure 3
Fusion de la lettre 'K'
avec le modèle 'K'

Distance D(Lettre,Modèle) :
D(K,K)=0.09 ; D(K,X)=0.21 ; D(K,H)=0.27 ; D(K,Y)=35.3 ; D(K,E)=38.5 ...

Les résultats d'appariements obtenus sur un ensemble réduit de lettres (caractères bâtons majuscules) offrent des perspectives intéressantes en ce qui concerne la reconnaissance des mots manuscrits. En réalisant, pour chaque lettre à reconnaître, un appariement structurel avec tous les graphes des lettres du dictionnaire de formes, on obtient une liste de coûts correspondant aux indices de ressemblance, de telle sorte que les lettres les plus proches sont connues et cette information peut être utilisée dans un contexte lexical.

Nous envisageons deux façons d'entreprendre la reconnaissance structurelle des mots manuscrits par les lettres :

- la première, en construisant le graphe sur toutes les lettres (ou entités proches tels que n-gram ou tronçon de lettre) que l'on pourra extraire sur le mot ;

- la seconde, en recherchant dans le graphe complet du mot non segmenté des isomorphismes optimaux de sous-graphes de lettres apprises.

Notre objectif est de retenir un alphabet de référence contenant un nombre minimum nécessaire de lettres de référence (en y incluant quelque variantes) pour une application particulière.

Il nous reste à tester notre méthode à une plus grande échelle pour observer comment évolue la discrimination entre les lettres de référence lorsque leur nombre augmente. Nous estimons que le

graphe structurel que nous avons construit est validé dans la mesure où la distance entre les graphes est égale à la distance entre les lettres originales.

BIBLIOGRAPHIE

BACKER E., GERBRANDS J.J. (1987) Inexact graph matching used in machine vision, NATO ASI Series. Vol. F30, Pattern Recognition Theory and Applications, Edited by P.A. Devijver and J. Kittler, © Springer-Verlag Berlin Heidelberg 1987, pp. 347-356.

WONG A.K.C. (1987) Structural Pattern Recognition : A Random Graph Approach, NATO ASI Series. Vol. F30, Pattern Recognition Theory and Applications, Edited by P.A. Devijver and J. Kittler, © Springer-Verlag Berlin Heidelberg 1987, pp. 323-345.

YOU M., WONG A.K.C. (1984) An Algorithm for Graph Optimal Isomorphism, 7th ICPR Canada, pp. 316-319.