



HAL
open science

Reconnaissance de symboles guidée par une modélisation basée sur les graphes d'adjacence

Hervé Locteau, Sébastien Adam, Eric Trupin, Jacques Labiche, Pierre Héroux

► **To cite this version:**

Hervé Locteau, Sébastien Adam, Eric Trupin, Jacques Labiche, Pierre Héroux. Reconnaissance de symboles guidée par une modélisation basée sur les graphes d'adjacence. Sep 2006, pp.151-156. hal-00119188

HAL Id: hal-00119188

<https://hal.science/hal-00119188>

Submitted on 8 Dec 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 symbole guidée par une modélisation basée sur les graphes de régions adjacentes

Hervé Locteau – Sébastien Adam – Éric Trupin – Jacques Labiche – Pierre Héroux

LITIS, Université de Rouen
F-76800 Saint-Etienne du Rouvray, France

Herve.Locteau@univ-rouen.fr

Résumé : *Cet article introduit une modélisation de symbole rigide faisant intervenir en parallèle deux points de vue, l'un structurel, l'autre statistique. Cette combinaison permet de surmonter les points faibles spécifiques à chaque famille d'approches. Un graphe des régions adjacentes concernant les occlusions est défini et est associé à un graphe de primitives vectorielles connectées. Ces définitions duales donnent lieu à une représentation des occlusions en termes de contour et de contenu par le biais de moments invariants. L'exploitation de cette représentation par un algorithme glouton, suite à l'utilisation d'une méthode originale de prétraitement, est validée à partir du benchmark établi pour le workshop GREC 2003.*

Mots-clés : Graphe des régions adjacentes, représentation de symboles, reconnaissance de symboles.

1 Introduction

La reconnaissance de symboles consiste à localiser et identifier des symboles présents sur un document graphique [LLA 03]. De nombreux types de documents mettent en jeu des symboles connectés aux autres composantes graphiques, rendant la tâche d'extraction de symboles difficile [COR 00]. Parmi les domaines d'application envisagés et les spécificités leur étant associées, nous pouvons distinguer trois grandes classes de documents : (i) les dessins techniques, (ii) les cartes et (iii) une dernière catégorie regroupant les partitions musicales et les formules mathématiques.

Concernant les dessins techniques, les symboles peuvent apparaître au sein d'un ensemble de traits, sans privilégier ni une orientation, ni une échelle. La densité d'informations sur des zones réduites d'un document peut être élevée et la multiplicité des orientations et échelles compliquent alors le processus de segmentation des symboles. Comme tout problème de classification, la reconnaissance de symbole doit être opérée sur des données bruitées du fait de la qualité du document d'origine mais aussi du processus de numérisation.

Les processus de reconnaissance de symboles (rigides) doivent être conçus au vu de contraintes bien établies parmi lesquelles l'invariance aux transformations affines, la segmentation, le bruit ainsi que le nombre de symboles traités. Pour évaluer la qualité des méthodes au regard de ces nombreuses contraintes, le concours de reconnaissance de symboles du GREC 2003 a mis à disposition un protocole d'évaluation de performances [VAL 04].

Une classification commune des modélisations existantes de symboles consiste à opposer les approches structurales (représentations) d'une part aux approches statistiques d'autre part (descripteurs). Dans la littérature ([YAN 04, PEN 04, RAM 03a]), de nombreux modèles structurels impliquent des primitives vectorielles agencées dans une structure où sont définies des relations parmi lesquelles l'interconnection, la perpendicularité, l'intersection, ou le parallélisme. D'autres primitives sont plus rarement utilisées telles que les occlusions ou les contours. Quel que soit le contenu de cette représentation, l'identification consiste à trouver un isomorphisme de sous-graphe entre un graphe prototype et le graphe d'un document. Les approches structurales peuvent être mises en défaut en présence de bruit ou de déformation. En revanche, la segmentation peut être guidée et la prise en compte de l'invariance à la rotation et aux changements d'échelle est relativement naturelle.

En reconnaissance statistique de formes, chaque entité est associé à un vecteur caractéristique d'une portion de l'image. Parmi les descripteurs, citons les études sur les caractéristiques géométriques, les invariants ou encore les transformées ([KHO 90, ADA 02, MUK 01, RAM 03b, TAB 03, CHO 04, YAN 05]). La classification est obtenue en utilisant une technique de partitionnement de l'espace des caractéristiques. Les approches statistiques sont peu sensibles au bruit car elles s'appuient généralement directement sur l'image. Toutefois, l'invariance à la rotation et aux changements d'échelle est parfois délicate à obtenir pour certaines signatures. Enfin, les résultats des méthodes statistiques dépendent de la précision associée à la définition de la région d'intérêt, peu fiable pour des symboles en contexte.

Nous présentons en section 2 la modélisation adoptée pour les symboles. Celle-ci s'appuie sur une double définition des occlusions par le biais des contours et du contenu des occlusions. En section 3, nous proposons à titre d'illustration un algorithme exploitant cette représentation tandis que la section 4 aborde l'usage des caractéristiques statistiques. Nous introduisons en section 5 une méthode générale de prétraitement de documents techniques pour palier les erreurs de segmentation. L'évaluation de notre méthodologie de reconnaissance de symboles est fournie dans la section 6. La section 7 apporte nos conclusions ainsi que les perspectives envisagées.

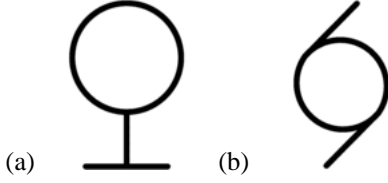


FIG. 1 – Les graphes de régions adjacentes ne permettent pas de distinguer certains symboles : il convient de leur adjoindre une autre structure de graphe impliquant des primitives vectorielles.

2 Modélisation proposée

Dans une chaîne d'analyse de documents, les étapes de bas niveaux ainsi que les images de symboles sont sources d'erreurs de représentation. La modélisation d'une instance d'un objet en utilisant un graphe attribué peut donc différer du modèle idéal de cet objet. La structure de graphe peut avoir plus ou moins de nœuds et d'arcs et les étiquettes associées peuvent être différentes. L'étape de reconnaissance doit donc être tolérante à ce type de biais afin d'identifier un objet comme étant une instance déformée du modèle impliqué.

Nous proposons donc une représentation à base de deux graphes, l'un à partir de primitives vectorielles et l'autre à partir de caractéristiques statistiques. Considérant un graphe de régions adjacentes (GRA), deux formalismes sont donc employés pour les occlusion désignées par les nœuds. À chaque occlusion est associé à la fois :

- une définition des contours de l'occlusion à base de primitives vectorielles,
- une définition du contenu de l'occlusion sous forme d'invariants (pseudo-invariants de Zernike).

L'invariance à la rotation et à l'échelle de la représentation vectorielle des occlusions est obtenue en exprimant les positions relatives des extrémités des primitives vectorielles en coordonnées polaires par rapport au centre de gravité, l'orientation ainsi que la taille de l'occlusion. Les arcs du GRA mettent en évidence les frontières communes à deux régions adjacentes sous forme de liste de primitives vectorielles. Des propriétés relatives telles que la position des centres de gravité, le rapport de surfaces et l'orientation sont également spécifiées. On peut noter que les GRA ne sont pas adaptés pour modéliser certaines formes (cf. Figure 1(a)).

Pour y remédier, toutes les primitives vectorielles sont retranscrites au niveau d'un graphe de primitives interconnectées. Une telle représentation permet de définir précisément les zones d'intérêt lorsqu'on souhaite interpréter un document complet où les symboles sont interconnectés.

Nous précisons par la suite les estimateurs, caractérisant les occlusions et les relations d'adjacences, intervenant durant le processus de reconnaissance de symbole afin d'obtenir une représentation invariante aux transformations affines.

2.1 Invariance à la rotation

L'orientation ϕ d'une forme sur une image f peut être évaluée ($\pm\pi$) à partir des moments centrés de Hu $\mu_{p,q}$ d'ordre $p+q=2$:

$$\phi = \frac{1}{2} \arctan \left(\frac{2\mu_{1,1}}{\mu_{2,0} - \mu_{0,2}} \right) \text{ où } m_{p,q} = \sum_x \sum_y f(x,y)x^p y^q.$$

Cette orientation est obtenue en déterminant les axes d'inertie de la forme mais l'information de direction n'apparaît pas. Puisque la détermination de la direction d'une forme serait arbitraire, nous proposons d'adjoindre à la modélisation des occlusions un descripteur permettant, en présence d'une même forme orientée selon deux angles, de s'affranchir du modulo π lorsqu'on souhaite connaître l'angle de rotation permettant de lier ces formes. Ce descripteur est issu du calcul des moments de Zernike définis par :

$$A_{n,l} = \frac{n+1}{\pi} \int_0^{2\pi} \int_0^1 V_{n,l}^*(\rho, \theta) f(\rho, \theta) \rho d\rho d\theta \quad (1)$$

avec :

- $f(\rho, \theta)$ le niveau de gris de l'image au point de coordonnées polaires (ρ, θ) ,
- $\rho \leq 1$,
- $V_{n,l}(\rho, \theta)$ le polynôme de Zernike d'ordre n et de fréquence l ,
- $n \in \mathbb{N}, l \in \mathbb{Z}, |l| \leq n, (n - |l|)$ pair,
- $z = (a + \mathbf{i}b) \in \mathbb{C}, z^* = a - \mathbf{i}b$.

Les polynômes de Zernike sont définis par :

$$V_{n,l}(\rho, \theta) = R_{n,l}(\rho) \exp(\mathbf{i}l\theta) \quad (2)$$

$$R_{n,l}(\rho) = \sum_{s=0}^{(n-|l|)/2} \frac{(-1)^s \rho^{n-2s} (n-s)!}{s! \left(\frac{n+|l|}{2} - s\right)! \left(\frac{n-|l|}{2} - s\right)!} \quad (3)$$

Partant de l'équation (1), Teague montre que si une forme subit une rotation d'angle ϕ , alors les moments $A_{n,l}'$ issus de cette nouvelle forme peuvent être exprimés en fonction des $A_{n,l}$ issus de l'image initiale par : $A_{n,l}' = A_{n,l} \exp(-\mathbf{i}l\phi)$. En choisissant un couple (n, l) tel que l soit impair, l'appariement d'une occlusion d'un modèle avec une occlusion d'une image fournit un angle de rotation robuste (estimation pondérée par $|A_{n,l}|$) en testant si :

$$A_{n,l}^{forme} = \begin{cases} A_{n,l}^{reference} \exp(-\mathbf{i}l\phi) \\ A_{n,l}^{reference} \exp(-\mathbf{i}l(\phi + \pi)) \end{cases}$$

L'orientation α_{source} d'une occlusion est propagée au niveau de chaque relation de voisinage en spécifiant la position du centre de gravité du nœud destination en fonction du centre de gravité du nœud source d'un arc selon :

$$g_{destination} = g_{source} + s \times d_{source,destination} \times \exp(\mathbf{i}\alpha_{source,destination} + \alpha_{source}) \quad (4)$$

où $\alpha_{source,destination}$ désigne l'orientation relative du segment $g_{source} g_{destination}$ reliant les deux centres de gravité par rapport à l'orientation α_{source} du nœud source ; la distance $d_{source,destination}$ entre g_{source} et $g_{destination}$ relative à l'échelle de source ; le facteur d'échelle s du symbole.

2.2 Invariance au changement d'échelle

L'invariance au changement d'échelle est obtenue en introduisant au niveau des arcs du graphe d'adjacence la notion de taille relative d'une occlusion vis à vis des occlusions adjacentes. Le calcul de surface d'une occlusion

est sensible aux variations d'épaisseur du tracé (introduites par le bruit). Puisqu'à chaque occlusion correspond un cycle au niveau du squelette, la surface d'une région est évaluée en dilatant les occlusions jusqu'à ce que les points de cette forme coïncident avec ceux du cycle du squelette. Durant le processus de reconnaissance de symbole, l'échelle de l'instance vis à vis de son modèle est évaluée par une moyenne pondérée des échelles des occlusions composites. Soient P et R les cycles associés respectivement à l'occlusion d'une image en cours d'analyse ; à une occlusion du modèle d'un symbole, le facteur d'échelle d'une occlusion (élément composite du symbole) $forme$ en fonction de l'occlusion $reference$ est donné par :

$$\begin{aligned} s_{forme} &= \left(\frac{\text{aire}(\text{Dilatation}(forme))}{\text{aire}(\text{Dilatation}(reference))} \right)^{\frac{1}{2}} \\ &= \left(\frac{\text{card}((x,y) \subset P)}{\text{card}((x,y) \subset R)} \right)^{\frac{1}{2}} \end{aligned} \quad (5)$$

3 Reconnaissance de symboles

Dans la littérature, les symboles peuvent être reconnus en utilisant une stratégie ascendante ou descendante. La première famille d'approches tente de faire émerger un symbole à travers des représentations successives, le processus de reconnaissance ne tirant pas profit des connaissances acquises. D'un autre côté, la seconde famille d'approches repose sur des requêtes, les systèmes tentant de mettre en correspondance un modèle avec une partie des données. La reconnaissance consiste alors à vérifier la présence des différents composants du modèle. Les connaissances contextuelles (l'image à interpréter) et des contraintes (le modèle précisé au niveau de la requête) sont utilisées pour émettre puis valider/invalider des hypothèses. En pratique, les systèmes de reconnaissance alternent ces deux stratégies. La détection d'un élément composite permet aux systèmes de générer un ensemble d'hypothèses concernant les autres éléments composites proches de celui qui a été identifié. Ces stratégies mixtes privilégient les représentations structurelles lorsque l'on dispose d'un ensemble de symboles potentiels représentés au sein de documents complets à interpréter.

Nous déclenchons une hypothèse de reconnaissance de symbole lorsqu'une occlusion $loop_i$ est identifiée au sein d'un document comme étant une région d'un GRA (cf. algorithme 1¹). Par le biais des outils présentés dans la section précédente, l'hypothèse de reconnaissance conduit à une transformation $\lambda()$ ainsi qu'à une mise en correspondance entre la région d'amorce et le nœud du GRA candidat G_k .

¹Nous définissons ici les notations employées :

- $G = \{G_k\}_{k \in [1, K]}$ est la base de K modèles de symboles,
- $Node(G_k) = \{node_k^n\}$ et $Edge(G_k) = \{edge_k^e\}$ désignent respectivement l'ensemble des nœuds et des arcs de G_k ,
- $\mathcal{L}()$ est le classifieur statistique des occlusions,
- Λ est l'ensemble des paramètres de la transformation affine $\lambda()$,
- $X \equiv Y$ désigne un isomorphisme entre X and Y ,
- $CG(j)$ est l'ensemble des modèles de symboles candidats, les GRA dont un nœud possède l'étiquette $label_j$

Les relations de voisinages spécifiées au niveau de G_k permettent de générer des contraintes quant à la présence d'occlusions, éventuellement adjacentes également au niveau de l'image analysée, dans le voisinage de $seedLoop$ et pour lesquelles (i) le vecteur Λ est constant et (ii) les propriétés relatives (rapport de surfaces, orientation et positions des centre de gravité) sont respectées. Une hiérarchie des occlusions voisines à identifier permet d'interrompre dès que possible ce processus en privilégiant les régions de plus grandes tailles. Cette stratégie est illustrée par l'algorithme 2¹ : il s'agit d'un algorithme glouton dont l'exploration est contrôlée par le modèle candidat G_k .

Algorithm 1 ¹Déclenchement de la reconnaissance

Input : $Image$
Output : indexation de l' $Image$
for all $loop_i \in Image$ **do**
 for all $label_j \in \mathcal{L}(loop_i)$ **do**
 $CG(j) \leftarrow \{G_m \in G \mid \exists node_m^n \in Node(G_m), label_j \in \mathcal{L}(node_m^n)\}$
 for all $G_k \in CG(j)$ **do**
 for all $node_k^l \in Node(G_k), label_j \in \mathcal{L}(node_k^l)$ **do**
 $RAG_Matching(G_k, node_k^l, loop_i)$
 end for
 end for
 end for
end for

Algorithm 2 ¹RAG_Matching

Inputs : G_k le modèle candidat à identifier
 $seedNode$ un nœud de G_k
 $seedLoop$ une occlusion de l'image telle que $\mathcal{L}(seedLoop) \cap \mathcal{L}(seedNode) \neq \emptyset$
Output : vrai si la propagation des contraintes de G_k est validée
 $Matching \leftarrow \{seedNode\}$
 Évaluer Λ à partir de $(seedNode, seedLoop)$
 $Node(G) \leftarrow \{seedLoop\}$
repeat
 $AN(Matching) \leftarrow \{node_k^l \in Node(G_k) \mid \exists edge_m^e \in Edge(G_m), edge_m^e(source) \in Matching, edge_m^e(target) \in Node(G_k) \setminus Matching\}$
 $node_k^A \leftarrow \arg \max_{n \in AN(Matching)} \{Area(n)\}$
 if $find(loop_i \in Image \mid \mathcal{L}(loop_i) \cap \mathcal{L}(node_k^A) \neq \emptyset, loop_i \Leftrightarrow \lambda(node_k^A))$ **then**
 $Matching \leftarrow Matching \cup \{node_k^A\}$
 $Node(G) \leftarrow Node(G) \cup \{loop_i\}$
 else
 Retourner **faux**
 end if
 Mettre à jour Λ
until $G \equiv G_k$
Retourner **vrai**

Puisque l'approche présentée est équivalente à la recherche d'isomorphisme de sous-graphe, un ordonnance-

ment des hypothèses d'interprétation parmi les modèles $G = \{G_k\}_{k \in [1, K]}$ est effectué afin d'éviter toute reconnaissance prématurée du symbole G_i au lieu du symbole G_j si $G_i \subset G_j$.

4 Apprentissage et classification statistique des occlusions

La classification en reconnaissance de formes se heurte fréquemment au manque de données d'apprentissage. En utilisant l'approche de [KAN 94], des images bruitées à différents niveaux, simulant le bruit introduit par des opérations telles que l'impression, la reproduction ou encore la numérisation, sont générées.

À partir d'un ensemble d'images synthétiques, dont on dispose alors d'une vérité terrain quant aux appariements nécessaires entre les occlusions du modèle G_k et celles de l'instance, nous pouvons extraire sur chaque image bruitée prétraitée (voir section 5), un vecteur caractéristique f_k^j par occlusion auquel est associé une étiquette. À cette étape, si N_k désigne le nombre de nœuds du GRA G_k associé au symbole idéal S_k , nous introduisons $N = \sum_k N_k$ étiquettes distinctes afin de n'introduire aucune connaissance. Après avoir collecté ces données d'apprentissage, une analyse en composantes principales sur les $\{f_k^j\}_k$ est appliquée afin de réduire la dimension de l'espace des caractéristiques de manière non supervisée et par conséquent le nombre de paramètres des modèles employés durant l'étape de classification des occlusions. Chaque nuage de points $\{f_k^j\}_{k,j}$ est approximé par une distribution gaussienne multi-dimension afin d'utiliser la règle de décision empirique de Bayes. Une étude de la matrice de confusion des résultats de classification de ces classificateurs sur une base de validation, construite à partir du même protocole que celui de la base d'apprentissage, conduit à la réduction de la taille de l'alphabet d'étiquettes.

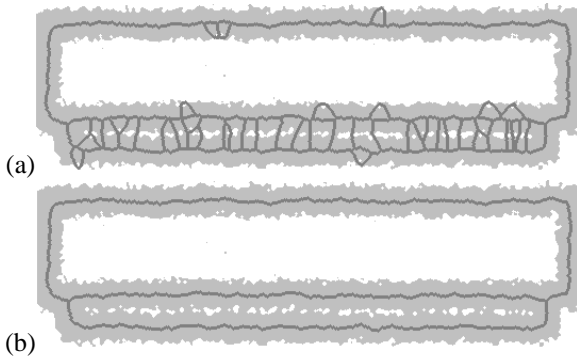


FIG. 2 – Cycles du squelette : (a) les cycles initiaux et (b) les cycles retenus suite à notre filtrage

5 Prétraitement proposé

Une méthode de reconnaissance peut être inadaptée en présence de bruit. Cette section introduit un prétraitement générique original dont le propos est de corriger les erreurs de segmentation (fragmentation des occlusions du modèle, voir figure 2). De manière évidente, la vision humaine intègre la caractérisation d'une famille de documents ainsi que la re-

cherche des spécificités de l'un d'entre eux. Concernant les documents techniques, on peut remarquer que les symboles sont majoritairement représentés par des traits d'épaisseur constante. Nous pouvons donc en déduire qu'avant même de reconnaître un symbole sur un document, nous sommes capables d'en supprimer en partie le bruit. En effet, nous pouvons considérer que les traits les plus fins sont le fait d'une dégradation du document initial.

À partir d'un graphe du squelette d'une forme, le squelette étant construit à partir d'une transformée en distance tel que [BAJ 96], nous construisons un GRA de cycles $G_C(N_C, E_C)$ pour lequel un nœud désigne un cycle géométriquement contraint, i.e. un cycle au sein du graphe du squelette définissant le pourtour d'une unique occlusion; un arc désigne une relation d'adjacence à laquelle est associée une étiquette *voisine* ou *incluse*. Pour chaque couple de régions adjacentes, nous extrayons deux caractéristiques issues de la carte des distances et des branches communes connexes du squelette. À chaque arc $e_i \in E_C$ sont assignées une profondeur minimale (notée $min_d(e_i)$) et une profondeur de référence (notée $ext_d(e_i) = min\{profondeur(pp(e_i), profondeur(dp(e_i))\}$ où pp et dp sont le premier et dernier points) à partir des extrémités de la frontière. Une profondeur globale moyenne \bar{d} ainsi que la variance globale de la profondeur σ_d sont évaluées et utilisées en tant que mesures contextuelles. À partir des $\{min_d(e_i)\}_{e_i \in E_C}$, nous décidons de supprimer des branches du squelette, c'est-à-dire de fusionner des cycles, donc des occlusions, selon la règle suivante :

$$SI ((min_d(e_i) < ext_d(e_i)/2) OU (min_d(e_i) < \bar{d} - max\{\bar{d}/2, \sigma_d\})) ALORS Fusionner(e_i.source, e_i.target)$$

La première condition permet de détecter les irrégularités de l'épaisseur d'un unique trait (un facteur de 1/2 est employé dans la mesure où nous n'utilisons pas d'algorithme de correction des points de jonction). La seconde condition permet d'introduire une connaissance contextuelle pour discriminer les tracés épais des tracés fins au regard de l'image. Ce prétraitement ne nécessite aucun apprentissage et fournit des résultats de segmentation satisfaisant pour des documents techniques caractérisés par la constance de l'épaisseur du tracé pour un même objet.

6 Résultats expérimentaux

Nous avons soumis les images du concours de reconnaissance de symboles du GREC 2003 à notre système dans le but d'en évaluer les performances en fonction du nombre de symboles modèles. Les tests retenus concernent les dégradations binaires ainsi que les changements d'échelle. Dans la mesure où notre système ne s'appuie que sur les occlusions pour amorcer une reconnaissance, nous ne présentons pas d'expérimentations issues des modèles 7 à 9 de dégradation qui engendrent majoritairement une érosion du tracé que nos prétraitements ne corrigent pas. Les caractéristiques utilisées sont les pseudo invariants de Zernike jusqu'à l'ordre 6. Pour chaque symbole, à partir de 10 images par modèle de dégradation, nous obtenons 60 images dégradés d'un même

	Taux	Images idéales	Images ± grandes	Modèles de dégradation binaire						borne théo.
				M1	M2	M3	M4	M5	M6	
Jeu 1 (5 symboles)	Reco.	100	72	96	100	96	72	92	96	100
	Rejet	0	28	4	0	4	28	8	4	0
Jeu 2 (20 symboles)	Reco.	80	67	85	80	76	71	72	77	90
	Rejet	20	33	15	20	24	29	28	23	10
Jeu 3 (50 symboles)	Reco.	72	64.4	68.4	67.6	62.4	62.8	52.8	65.2	82
	Rejet	12	18.4	14.8	14.8	22.4	14.8	18.8	12.8	6
	Erreur	16	17.2	16.8	17.6	15.2	22.4	28.4	22	12

TAB. 1 – Performance (en %) pour les 3 jeux de tests du concours de reconnaissance de symbole du GREC 2003 (chaque symbole y apparaissant 5 fois) et bornes théoriques identifiées par le système. Les expérimentations sont effectuées à la résolution initiale (512 x 512 pour les images idéales et celles reproduisant une dégradation binaire).

symbole que nous répartissons équitablement pour la constitution des bases d'apprentissage et de validation. L'étape de réduction de l'espace des caractéristiques des occlusions permet de se ramener à des classifieurs à 8 paramètres au lieu de 24 tandis que le traitement de la matrice de confusion sur la base de validation permet d'attribuer aux 120 classifieurs 82 étiquettes de classes pour le troisième jeu de tests (cf. section 4). L'ensemble des performances (taux de reconnaissance, de rejet et d'erreur) est reporté au niveau du tableau 1. On peut noter qu'aucune erreur n'est produite pour les deux premiers jeux de test (19.45% en moyenne sur le 3^{ème} jeu).

Introduisons la notion de borne théorique du taux de reconnaissance. Dans l'implémentation actuelle de notre système, la reconnaissance d'un symbole est amorcée par la reconnaissance d'une occlusion. Pour le jeu 2, un symbole ne possédant aucune occlusion, le système rejette l'ensemble des images le représentant. Pour le jeu 3, ce phénomène est reproduit pour deux symboles.

La reconnaissance d'un symbole est établie en s'appuyant sur les relations d'adjacence spécifiées au sein des modèles. Le modèle de la figure 4(b) ne peut donc être exploité. Enfin, des groupes de symboles se voient assigner un même modèle (voir figure 6(a-c)) si on ne considère que le graphe d'adjacence. Il résulte pour les jeux 1, 2 et 3 des bornes théoriques du taux de reconnaissance de 100%, 90% et 82%. Les taux de reconnaissance obtenus indiquent donc que le système présenté est capable de rejet en présence d'images dont les symboles n'ont pas été modélisés.

Certaines erreurs de reconnaissance ou rejets sont la conséquence du processus de segmentation. Certaines occlusions ne sont pas fusionnées par le notre filtrage :

- les occlusions en englobant une autre peuvent être fragmentées au niveau des rapprochements de leur tracé (voir figure 3(a)).
- les occlusions fines peuvent être fragmentées à de multiples localisations (voir figures 3(b) et 3(c)),

Outre les erreurs liées aux symboles ayant un même modèle, les erreurs du système proviennent de la segmentation. La segmentation proposée ne permet pas de mettre en correspondance suffisamment de données du bon modèle (la figure 5 illustre ce type d'erreur). En excluant les erreurs liées au fait que le système exploite, dans la version disponible, un

même modèle pour plusieurs symboles (cf. figure 6), pour 12.22% sur 19.45% des cas, la proposition de reconnaissance est un modèle de symbole n'ayant qu'une occlusion.

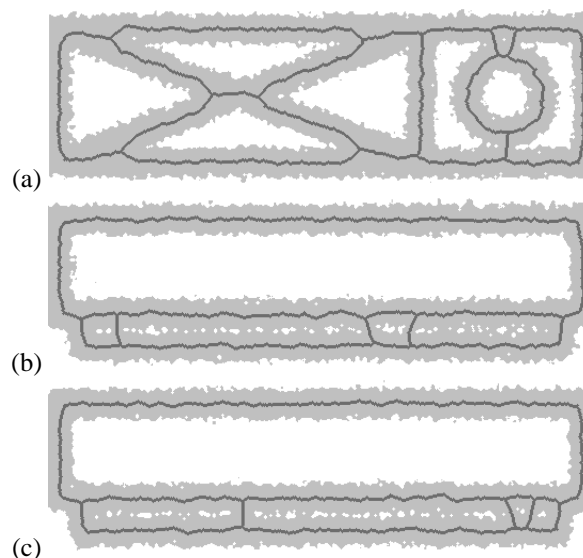


FIG. 3 – « Erreurs de segmentation restantes » suite au filtrage - (a) Occlusions proches liées par une relation d'inclusion : l'occlusion englobante est fragmentée du fait de la croissance de l'occlusion englobée - (b) et (c) : Occlusions fines dont la localisation d'une fragmentation peut être multiple

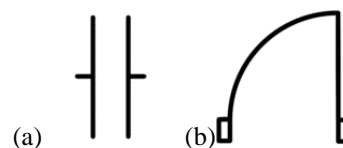


FIG. 4 – Exemples de symboles dont les modèles ne sont pas exploités par le système présenté



FIG. 5 – Confusion de symboles suite aux erreurs de segmentation

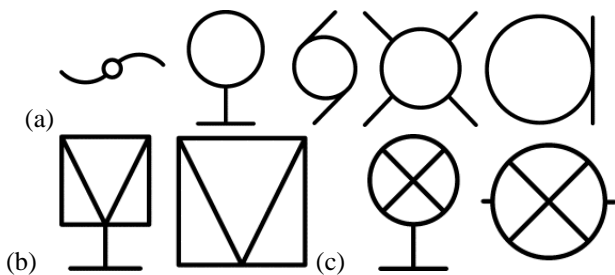


FIG. 6 – Groupes de symboles similaires et non différenciables, identifiés en tant que tels par notre système

7 Conclusion et perspectives

La conception d'un système de reconnaissance de symboles repose sur (i) les prétraitements, (ii) la modélisation des données et (iii) le processus de décision. Pour cette étude, nous nous sommes concentrés sur les deux premiers points et avons proposé une modélisation basée sur les graphes de régions adjacentes dont les nœuds bénéficient d'une double modélisation (statistique et vectorielle). Un premier système, non optimal, est présenté afin de fournir quelques observations quant à la pertinence de la modélisation de symboles rigides. Cette modélisation permet de définir les occlusions à partir de leur contour mais aussi de leur contenu. Les expérimentations ont montré que la différenciation de symboles ne peut être obtenue à partir d'un graphe d'adjacence seul. Notre système doit donc exploiter l'ensemble du graphe de primitives vectorielles. Notre modélisation n'a été validée que sur une partie des dégradations binaires disponibles pour GREC 2003, laissant de côté les modèles de bruit entraînant une érosion du tracé, et par extension, la fusion d'occlusions, l'absence d'amorce pour la reconnaissance de symbole. Il convient donc de pouvoir déclencher une reconnaissance également de manière structurelle en employant une distance d'édition de chaînes cyclique telle que [SÁN 02]. Cette combinaison permet d'émettre et de valider des hypothèses durant le processus d'interprétation d'un document et d'en renforcer la confiance. Ceci permet également d'accroître le pouvoir discriminant des graphes de régions adjacentes. En n'ayant défini qu'un unique modèle par symbole durant l'apprentissage, nous devons tirer parti, dans la phase de reconnaissance, des frontières communes. En effet, les formes fines pouvant être fragmentées en dépit des prétraitements employés et les frontières communes doivent permettre de générer des hypothèses de correction de la segmentation.

Références

- [ADA 02] ADAM S., OGIER J., CARIOU C., , GARDES J., A Scale and Rotation Parameters Estimator Application to Technical Document Interpretation, *Graphics Recognition. Algorithms and Applications*, Springer-Verlag Heidelberg, 2002, pp. 266–272.
- [BAJ 96] SANNITI DI BAJA G., THIEL E., Skeletonization algorithm running on path-based distance maps, *Image and Vision Computing*, vol. 14, 1996, pp. 47–57.
- [CHO 04] CHONG C.-W., RAVEENDRAN P., MUKUNDAN R., Translation and scale invariants of Legendre moments, *Pattern Recognition*, vol. 37, n° 1, 2004, pp. 119–129.
- [COR 00] CORDELLA L. P., VENTO M., Symbol recognition in documents : a collection of techniques ?, *International Journal on Document Analysis and Recognition*, vol. 3, n° 2, 2000, pp. 73–88.
- [KAN 94] KANUNGO T., HARALICK R. M., BAIRD H. S., STUEZLE W., , MADIGAN D., Document Degradation Models : Parameter Estimation and Model Validation, *IAPR Workshop on Machine Vision Applications*, 1994, pp. 552–557.
- [KHO 90] KHOTANZAD A., HONG Y., Invariant Image Recognition by Zernike Moments, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, n° 5, 1990, pp. 489–497.
- [LLA 03] LLADÓS J., VALVENY E., , SÁNCHEZ G., A Case Study of Pattern Recognition : Symbol Recognition in Graphic Documents, *Pattern Recognition in Information Systems*, 2003, pp. 1–13.
- [MUK 01] MUKUNDAN R., ONG S., , LEE P., Image Analysis by Tchebichef Moments, *IEEE Transactions on image Processing*, vol. 10, n° 9, 2001, pp. 1357–1364.
- [PEN 04] PENG B., LIU Y., LIU W., , HUANG G., Sketch Recognition Based on Topological Spatial Relationship, *Structural, Syntactic, and Statistical Pattern Recognition*, 2004, pp. 434–443.
- [RAM 03a] RAMEL J.-Y., VINCENT N., Strategies for Line Drawing Understanding, *International Workshop on Graphics Recognition*, 2003, pp. 13–20.
- [RAM 03b] RAMOS O., VALVENY E., Radon Transform for Lineal Symbol Representation, *International Conference on Document Analysis and Recognition*, 2003, pp. 195–199.
- [SÁN 02] SÁNCHEZ G., LLADÓS J., , TOMBRE K., A mean string algorithm to compute the average among a set of 2D shapes, *Pattern Recognition Letter*, vol. 23, 2002, pp. 203–213.
- [TAB 03] TABBONE S., WENDLING L., , TOMBRE K., Matching of graphical symbols in line-drawing images using angular signature information, *International Journal on Document Analysis and Recognition*, vol. 6, n° 2, 2003, pp. 115–125.
- [VAL 04] VALVENY E., DOSCH P., Symbol Recognition Contest : A Synthesis, *Graphics Recognition : Recent Advances and Perspectives*, vol. 3088 de *Lecture Notes in Computer Science*, pp. 368–385, Springer-Verlag, 2004.
- [YAN 04] YAN L., HUANG G., YIN L., , WENYIN L., A Novel Constraint-Based Approach to Online Graphics Recognition, *Structural, Syntactic, and Statistical Pattern Recognition*, vol. 3138 de *Lecture Notes in Computer Science*, 2004, pp. 104–113.
- [YAN 05] YANG S., Symbol Recognition via Statistical Integration of Pixel-Level Constraint Histograms : A New Descriptor, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, n° 2, 2005, pp. 278–281.