



Vers une approche floue d'encapsulation de graphes: application à la reconnaissance de symboles

Muhammad Muzzamil Luqman, Thierry Brouard, Jean-Yves Ramel, Josep Lladós

► To cite this version:

Muhammad Muzzamil Luqman, Thierry Brouard, Jean-Yves Ramel, Josep Lladós. Vers une approche floue d'encapsulation de graphes: application à la reconnaissance de symboles. Colloque International Francophone sur l'Écrit et le Document, Mar 2010, Sousse, Tunisie. pp.168-184, 2010. <hal-00483706>

HAL Id: hal-00483706

<https://hal.archives-ouvertes.fr/hal-00483706>

Submitted on 17 May 2010

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.

Vers une approche floue d'encapsulation de graphes : application à la reconnaissance de symboles

Muhammad Muzzamil Luqman^{*+} — Thierry Brouard^{*} —
Jean-Yves Ramel^{*} — Josep Lladós⁺

^{*} Laboratoire d'Informatique, Université François Rabelais Tours, 37200 - France

⁺ Computer Vision Center, Universitat Autònoma de Barcelona, 08193 - Spain

{brouard, ramel}@univ-tours.fr, {mluqman, josep}@cvc.uab.es

RÉSUMÉ. Nous présentons une nouvelle méthode de reconnaissance des symboles, reposant sur une approche structurale pour représenter les associations visuelles dans les symboles et un classificateur statistique pour la reconnaissance. Un symbole graphique est vectorisé, ses détails topologiques et géométriques sont codés par un graphe relationnel attribué et une signature est produite. Des intervalles flous calculés d'après les données ont été introduits pour éviter la sensibilité des représentations structurales au bruit. La distribution de probabilités jointes des variables composant les signatures est codée par un réseau bayésien, qui permet de plus de choisir un sous-ensemble de caractéristiques pertinentes. Il est utilisé en apprentissage supervisé pour la reconnaissance des symboles. Les résultats expérimentaux montrent les capacités de reconnaissance et une robustesse au bruit (binaire, vectoriel, contextuel) sur les symboles 2D linéaires, architecturaux et électroniques provenant des bases de données GREC.

ABSTRACT. We present a new methodology for symbol recognition, by employing a structural approach for representing visual associations in symbols and a statistical classifier for recognition. A graphic symbol is vectorized, its topological and geometrical details are encoded by an attributed relational graph and a signature is computed for it. Data adapted fuzzy intervals have been introduced for addressing the sensitivity of structural representations to noise. The joint probability distribution of signatures is encoded by a Bayesian network, which serves as a mechanism for pruning irrelevant features and choosing a subset of interesting features from structural signatures of underlying symbol set, and is deployed in a supervised learning scenario for recognizing query symbols. Experimental results on pre-segmented 2D linear architectural and electronic symbols from GREC databases are presented.

MOTS-CLÉS : intervalle flou, encapsulation de graphes, réseau bayésien, reconnaissance de symboles.

KEYWORDS: fuzzy interval, graph embedding, Bayesian network, symbol recognition.

1. Introduction

Les recherches en reconnaissance de documents graphiques s'intéressent à la localisation et à la classification de formes susceptibles d'apparaître dans des dessins ou documents numérisés. Le champ d'application est très vaste allant de la reconnaissance de logos sur les courriers jusqu'à l'analyse de partitions musicales. (Lladós *et al.*, 2003) mentionnent l'énorme masse de documents existante correspondant à des schémas électroniques, architecturaux et autres cartes plus spécifiques qui a, aujourd'hui, été numérisée et qu'il faudrait indexer. Tous ces dessins sont constitués d'un ensemble fini de formes connues *a priori* (alphabets de symboles) mais pouvant prendre des apparences variées (non rigide) qu'il est alors difficile mais nécessaire de localiser et de reconnaître dans leur contexte pour réaliser l'indexation. De plus, aujourd'hui, les dispositifs nomades autorisant la prise de notes manuscrites en ligne se répandent de plus en plus, augmentant encore la demande de systèmes de traitement de documents graphiques. Une description détaillée sur les domaines applicatifs de la reconnaissance de graphiques est disponible dans (Chhabra, 1998) ainsi que dans (Lladós *et al.*, 2002).

D'un point de vue scientifique, les états de l'art proposés par (Chhabra, 1998), (Cordella *et al.*, 2000), (Lladós *et al.*, 2002) et (Tombre *et al.*, 2006), fournissent un panorama complet et détaillé des travaux réalisés dans ce domaine depuis une quinzaine d'années démontrant les efforts importants consacrés à la mise en place de méthodes statistiques, syntaxiques, structurelles et hybrides utilisables sur de tels documents. Les approches syntaxiques ont d'abord proposé d'exploiter des grammaires couplées à des analyseurs syntaxiques (Bunke *et al.*, 2001) et ont été suivies par un usage assez intensif de méthodes structurelles (Lladós *et al.*, 2002). Les approches structurelles se distinguent des approches statistiques de par la structure des données manipulées afin de représenter les formes à reconnaître. Les approches structurelles exploitent bien souvent des attributs symboliques stockés sous forme de chaînes ou de graphes tandis que les approches statistiques représentent les formes par un ensemble de valeurs numériques stockées dans un vecteur. (Cordella *et al.*, 2000) ont produit une étude détaillée des méthodes employant ces différents types d'approches dans le contexte de l'analyse de symboles graphiques mettant en exergue le pouvoir de représentation des méthodes symboliques et structurelles. Cependant, ils notent également que ces méthodes semblent d'une complexité supérieure et sont également plus sensibles aux distorsions et bruits (Bunke *et al.*, 2001). D'un autre côté, l'utilisation de vecteurs numériques comparés à l'aide de classificateurs statistiques permet de bénéficier d'outils mathématiques très performants et bien connus, offrant de plus la possibilité de mise en place de techniques d'apprentissage. Ainsi, un délaissement des approches structurelles au bénéfice des méthodes statistiques s'est produit notamment durant les compétitions de reconnaissance de symboles pré-segmentés (GREC) laissant de côté l'analyse et l'indexation de dessins complets nécessitant d'exploiter l'information topologique.

Ces dernières années, la volonté de produire à nouveau des systèmes capables d'indexer ou de faire des recherches sur des documents complets semble renverser la

tendance. Afin de contribuer à cette nouvelle dynamique, le travail présenté dans cet article propose d'exploiter les avantages des méthodes structurales et statistiques en essayant d'améliorer la phase de transformation de la représentation structurale des données en un vecteur numérique grâce à l'utilisation de fonctions d'appartenance floues. Afin de présenter cette proposition, l'article est organisé de la manière suivante : la section 2 présente l'architecture générale du système. La section 3 décrit chacune des parties du système de manière détaillée. Les résultats expérimentaux obtenus sont présentés en section 4 et l'article se termine par une conclusion mettant en avant les futures directions de nos recherches.

2. Méthode proposée et travaux comparables

L'objectif de cette partie est de mettre en avant les points forts de notre proposition en les positionnant par rapport aux travaux antérieurs exploitant une méthodologie similaire.

Notre méthode est une méthode hybride combinant l'approche structurale et l'approche statistique. Elle propose d'exploiter à la fois le fort pouvoir de représentation des méthodes structurales à base de graphes (section 2.1) et l'efficacité calculatoire des méthodes de classification statistiques (section 2.2).

2.1. Signatures structurales et encapsulation de graphes

(Ventura *et al.*, 1994) ont été les premiers à introduire le concept de signature vectorielle dans le domaine de la reconnaissance de graphiques : ils proposent d'extraire les relations géométriques existantes entre les différents segments constituant les symboles. Le comptage du nombre de relations de chaque type (pré-défini) leur permettait de produire une signature numérique utilisée par la suite pour détecter le nombre d'occurrences d'un symbole particulier dans un dessin. Les travaux de ce type se sont ensuite succédés et récemment une synthèse sur l'usage de signature vectorielle en reconnaissance de graphiques a été produite (Rusinol, 2009). Parmi les travaux recensés, on note les travaux de (Dosch *et al.*, 2003), (Rusinol *et al.*, 2006), (Qureshi *et al.*, 2007), (Zhang *et al.*, 2007) et (Coustaty *et al.*, 2007). Les outils exploités sont variés comme par exemple la transformée de Hough pour détecter les lignes. Afin d'étudier ensuite les relations spatiales entre ces lignes, les treillis de Galois et les graphes d'adjacence de régions sont fréquemment rencontrés. On peut noter que la plupart du temps, seuls les segments rectilignes sont considérés pour produire les signatures, sauf dans les travaux de (Zhang *et al.*, 2007) qui exploitent également les primitives de types arc et cercle. (Qureshi *et al.*, 2007) propose de vectoriser les schémas pour produire un graphe relationnel attribué (ARG) à partir duquel il génère une signature vectorielle (G-Signature). La comparaison est la plupart du temps réalisée à l'aide d'une distance euclidienne par comparaison avec un ensemble de symboles modèles.

L'invariance aux transformations (rotation et changement d'échelle) dépend de la manière dont sont construites les signatures (utilisation par exemple des longueurs relatives plutôt que des longueurs absolues).

Concernant la résistance aux distorsions et aux dégradations, la plupart des travaux dénotent qu'il reste de nombreux progrès à faire pour arriver à concurrencer les méthodes purement statistiques.

Mais, les recherches réalisées dans le domaine de la comparaison de graphes ont fait naître de nouveaux espoirs grâce à l'utilisation de techniques comme le sondage de graphes (Lopresti *et al.*, 2004) et l'encapsulation de graphes (Luo *et al.*, 2003, Neuhaus *et al.*, 2006). Ces travaux proposent également de coupler les approches structurelles avec les approches statistiques de la même manière que le proposent les méthodes de signatures vectorielles dans le domaine spécifique de la reconnaissance de symboles.

C'est justement pour contribuer à ces travaux récents que nous proposons d'introduire une approche floue lors de la phase de construction d'une signature numérique partant d'un graphe attribué afin d'obtenir une meilleure résistance aux dégradations et de diminuer la complexité.

Pour cela, nous sommes repartis des travaux de (Qureshi *et al.*, 2007) proposant de construire une signature vectorielle à partir des résultats d'une vectorisation de l'image suivie par une phase de construction d'un graphe relationnel attribué permettant de stocker les relations existantes entre les primitives constituant les dessins. Les détails de la signature proposée initialement (G-Signature), construite à partir d'un sondage du graphe obtenu sont disponibles dans (Qureshi *et al.*, 2007) et sont accompagnés des résultats expérimentaux.

Notre première contribution réside dans l'exploitation de fonctions d'appartenance floues au lieu d'intervalles fixes et arbitraires lors de la réalisation du sondage du graphe (Luqman *et al.*, 2009).

Nos choix ont été motivés par les travaux de (Mitra *et al.*, 2005) sur les ensembles flous démontrant l'intérêt de ces approches pour augmenter la robustesse d'un système.

Outre cette gestion de l'incertitude, la seconde contribution de notre travail se situe dans l'exploitation de réseaux bayésiens comme classifieurs permettant la mise en place d'un mécanisme d'apprentissage plus performant comme l'explique (Valveny *et al.*, 1999).

2.2. Réseaux bayésiens

Les réseaux bayésiens sont des modèles graphiques probabilistes (MGP) et sont définis par leur structure et leur paramètres. La structure est formée par un graphe acyclique dirigé (GAD) qui représente les relations qui existent entre les variables du domaine. Les paramètres du réseau sont constitués des distributions de probabilités conditionnelles (DPC) associées à chacun des nœuds. À l'image d'autres modèles graphiques, un réseau bayésien encode la distribution des probabilités jointes d'un ensemble de variables aléatoires et peut être employé afin de répondre à toutes les questions possibles utilisant ces variables (action qui porte le nom d'inférence). Une

introduction aux réseaux bayésiens peut être consultée dans (Charniak, 1991) et dans (Heckerman, 2008).

Les réseaux bayésiens ont déjà été appliqués avec succès à un grand nombre de problèmes en apprentissage et en reconnaissance des formes. Ils sont bien connus pour leur performance et leur aptitude à faire des prédictions valides en situation d'incertitude. Mais à notre connaissance, ils sont très peu utilisés pour la reconnaissance de symboles graphiques. Récemment, dans (Barrat *et al.*, 2007), un classifieur naïf de Bayes a été utilisé sous l'angle purement statistique afin de reconnaître des symboles graphiques pré-segmentés. Leur système utilise trois descripteurs de forme : les descripteurs de Fourier génériques, les descripteurs de Zernike et la R-signature de dimension 1. Une sélection de caractéristiques pertinentes est réalisée (méthode projective) afin de former un vecteur de caractéristiques. Cela peut également permettre de réduire le nombre de variables du domaine et par conséquent le nombre de nœuds du réseau bayésien. Bien que performant, le classifieur naïf de Bayes suppose une forte relation d'indépendance entre les variables attributs, la variable classe étant connue.

Nous pensons que la puissance des réseaux bayésiens n'est que partiellement exploitée. Par exemple, la découverte de dépendances entre toutes les paires possibles de variables à partir des données associées est préférable à l'utilisation de relations pré-définies. Ceci peut aussi servir à détecter les variables non pertinentes au profit de celles menant à une meilleure discrimination des symboles présents (section 3.3 et section 3.4).

2.3. Originalité de notre approche

Notre approche correspond à une adaptation originale des méthodes de classification reposant sur les réseaux bayésiens pour la classification de symboles graphiques. Concernant les phases de capture de l'information et de représentation des données nous exploitons un graphe attribué encodant les informations géométriques, spatiales et relationnelles capturées à l'aide des primitives issues de la vectorisation.

La transformation du graphe en vecteur numérique (encapsulation de graphes) que nous proposons essaie de conserver les informations de structure des formes et de gérer l'imprécision des mesures grâce à l'utilisation de fonctions d'appartenance floues. Nous pensons que la mise en place d'intervalles chevauchants permettra d'augmenter significativement la robustesse du système vis-à-vis des distorsions et du bruit susceptibles d'apparaître dans de telles images comme stipulé dans (Mitra *et al.*, 2005).

Pour la phase de reconnaissance, nous avons choisi d'utiliser un réseau bayésien construit à partir d'une base d'apprentissage selon la proposition récente faite par (De-laplace *et al.*, 2006) qui propose d'exploiter un algorithme génétique pour déterminer la structure du réseau.

Les formes à reconnaître sont classifiées en utilisant la règle d'inférence probabiliste bayésienne classique et à partir de la distribution des probabilités jointes.

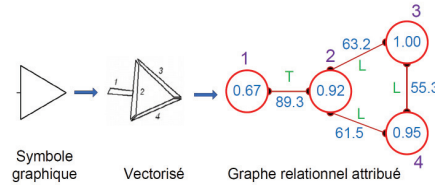


Figure 1. Phase de mise en place de la représentation.

Les caractéristiques exploitées pour produire le vecteur numérique (signature vectorielle) ont été choisies avec attention afin d’éviter toute redondance et toute perte d’information. Nous avons essayé de limiter au maximum le nombre de caractéristiques puisqu’il est reconnu que les réseaux bayésiens obtiennent de meilleurs résultats lorsque la dimensionnalité (nombre de nœuds) reste faible.

Notre signature reste de plus générique et robuste. Elle peut être utilisée quel que soit le type de dessins comportant des symboles linéaires 2D et ce pour des tâches de reconnaissance classique, pour effectuer de la localisation de symboles ou pour effectuer une indexation d’une base d’images de taille conséquente.

3. Description détaillée

Dans cette section nos propositions sont présentées selon un découpage couramment utilisé pour décrire les systèmes de reconnaissance de symboles. Ces phases ont été mises en évidence par (Cordella *et al.*, 2000) dans leur état de l’art. Ce découpage, que nous avons légèrement adapté pour qu’il corresponde à notre méthode, se compose de 3 phases successives : la construction d’une représentation, la production d’une description (à partir de fonctions d’appartenance floues dans notre cas) et la classification.

3.1. Phase de mise en place de la représentation

Cette phase primordiale correspond dans notre système à la construction des graphes relationnels attribués selon la méthode décrite dans (Qureshi *et al.*, 2007), et résumée figure 1. La topologie et les relations spatiales existantes entre les primitives de base constituant les dessins sont capturées et stockées dans le graphe. Le dessin est tout d’abord vectorisé puis représenté par un ensemble de primitives auxquelles sont associés différents labels (figure 1). Ces primitives deviennent les nœuds du graphe et un attribut numérique représentant la «longueur relative (normalisée entre 0 et 1)» ainsi qu’un attribut symbolique définissant le «type (vecteurs pour les régions pleines et quadrilatères pour les régions fines)» leurs sont associés. Les arcs du graphe décrivent le type de relation détectée entre les primitives à l’aide de deux attributs : un «type (L, X, T, P, S)» et un «angle relatif compris entre 0° et 90°».

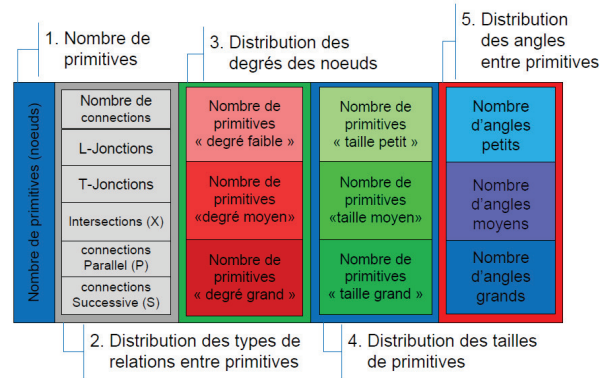


Figure 2. La signature "floue".

3.2. Phase de production de la signature "floue"

L'objectif de cette étape est de transformer le graphe relationnel attribué en un vecteur de caractéristiques numériques encapsulant au mieux toutes les informations contenues dans le graphe. L'objectif est aussi d'introduire une meilleure résistance aux distorsions et aux irrégularités de formes que l'on trouve souvent dans les images numérisées.

Pour cela, nous introduisons une approche intégrant de l'incertitude (que nous nommons abusivement "floue") durant la phase de sondage du graphe. Les valeurs prises par les attributs numériques des nœuds et des arcs du graphe sont réparties dans des classes (petit, moyen, grand) selon des fonctions d'appartenance définissant des intervalles se chevauchent comme décrit figure 3.

Le processus de transformation du graphe que nous proposons ici s'inspire des travaux décrit dans (Luqman *et al.*, 2009) en reprenant les même caractéristiques (sondes) extraites du graphe qui sont de différents types (voir figure 2) :

- G1 : nombre de primitives (nœuds)
- G2 : distribution des types de relation entre primitives (arcs)
- G3 : distribution des degrés des nœuds
- G4 : distribution des tailles de primitives (nœuds)
- G5 : distribution des angles entre primitives (arcs)

Le choix de ces caractéristiques s'appuie bien évidemment sur nos réflexions mais aussi sur les travaux antérieurs notamment ceux de (Rusinol *et al.*, 2006). Les groupes de caractéristiques G1 et G2 encodent la structure du symbole et l'arrangement des primitives composant le symbole. Ces caractéristiques permettront donc de distinguer les symboles ayant des formes différentes. Le groupe G3 décrit plutôt la complexité des symboles (faible ou grand nombres de relations entre les primitives le composant).

Les groupes G4 et G5 caractérisent l'homogénéité interne d'un symbole (variabilité de taille et d'angle entre les primitives de base). Les sondes correspondants aux groupes G1 et G2 s'obtiennent assez directement en parcourant et comptabilisant les nœuds et arcs du graphe.

Les intervalles associés aux groupes G4 et G5 sont obtenus simplement en divisant en trois parties de taille identique les plages de valeurs pouvant être occupées par la longueur relative et l'angle des primitives. Les intervalles obtenus sont décrits sur les figures 4 et 5.

Pour les caractéristiques du groupe G3, une base d'apprentissage est exploitée pour construire des intervalles de densités. La technique utilisée exploite les histogrammes des densités selon le procédé décrit dans (Leray *et al.*, 2004). Cette technique a été initialement proposée dans (Colot *et al.*, 1994) pour la discrétisation de données continues et est basée sur le critère AIC (*Akaike information criterion*). À partir d'un histogramme comportant au départ m bandes, le découpage optimal en n bandes ($n < m$) est déterminé en fusionnant successivement deux bandes adjacentes selon une fonction de coût associée au critère AIC. Les intervalles obtenus sont ensuite élargis afin d'obtenir un chevauchement entre les classes pour introduire l'incertitude (flou). Ces intervalles serviront ensuite à déterminer combien un graphe comporte de nœuds de degré faible, moyen ou élevé (trois sondes de notre signature). Les chevauchements entre intervalles sont introduits de manière à mieux gérer (absorber) les irrégularités dues aux bruits et distorsions qui pourraient affecter la signature. On associe à chaque intervalle une fonction d'appartenance floue telle que décrite sur la figure 3 à fin de définir la contribution apportée par une primitive en fonction de la valeur prise par un attribut. Nous disposons donc de 16 primitives, notées f1 (nombre de primitives), f2(nombre de connections), f3(L-Jonctions), f4(T-Jonctions), f5(Intersections (X)), f6(connections parallèles (P)), f7(connections successives (S)), f8(nombre de primitives «degré faible»), f9(nombre de primitives «degré moyen»), f10(nombre de primitives «degré grand»), f11(nombre de primitives «taille petit»), f12(nombre de primitives «taille moyen»), f13(nombre de primitives «taille grand»), f14(nombre d'angles petits), f15(nombre d'angles moyens) et f16 (nombre d'angles grands).

3.3. Phase d'apprentissage

Après avoir représenté l'ensemble des images de symboles de la base d'apprentissage par des graphes attribués, les signatures sont générées puis stockées sous la forme de vecteurs numériques. Chaque caractéristique est discrétisée indépendamment et séparément des autres. Les réseaux bayésiens sont ensuite appris.

L'apprentissage du réseau bayésien est réalisé en deux temps. La structure du réseau est d'abord apprise grâce à l'utilisation d'un algorithme génétique selon la méthode proposée dans (Delaplace *et al.*, 2006). Bien qu'il s'agisse d'une méthode évolutionnaire de construction, dans notre cas, les résultats obtenus ont toujours été très stables fournissant toujours la même structure de réseau lors d'invocations multiples sur une même base d'apprentissage.

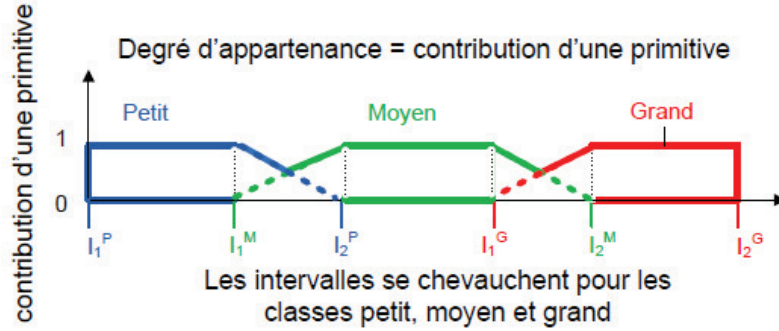


Figure 3. Fonction d'appartenance floue (définissant la contribution d'une primitive en fonction de la valeur prise). I_1^M et I_2^M définissent l'intervalle pour la classe moyen.

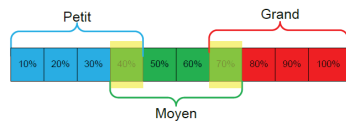


Figure 4. Intervalles associés à la longueur relative.

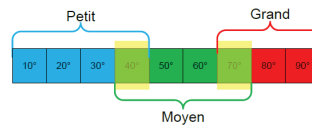


Figure 5. Intervalles associés à l'angle des primitives.

Chaque caractéristique de notre signature devient un nœud du réseau. L'objectif de la phase d'apprentissage de structure est alors de déterminer la structure du réseau la plus représentative des dépendances existantes entre toutes les paires de variables. La figure 6 montre l'une des structures de réseau apprises durant nos expérimentations.

La seconde phase de l'apprentissage est celle de l'estimation des paramètres, lesquels correspondent aux distributions de probabilités conditionnelles $\Pr(\text{nœud}_i | \text{parents}_i)$ associées aux nœuds du réseau. Ces distributions quantifient les dépendances relationnelles existantes entre les nœuds. Ces paramètres sont obtenus avec la méthode d'estimation du maximum de vraisemblance. Le problème de l'évitement des probabilités nulle est résolu en utilisant des probabilités *a priori* de Dirichlet. Le réseau obtenu représente la distribution des probabilités jointes associée aux signatures de symboles. Les propriétés d'indépendances conditionnelles (Charniak, 1991) du réseau nous ont permis d'ignorer les caractéristiques inutiles pour décrire les ensembles de données.

L'indépendance conditionnelle d'un nœud dans le réseau est directement exploitée durant l'inférence probabiliste et aide à ne pas tenir compte des variables inutiles lors du calcul des probabilités *a posteriori* pour les différentes classes de symboles (section 3.4).

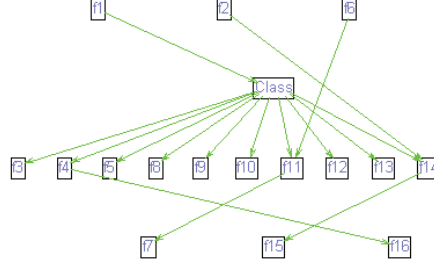


Figure 6. Exemple de structure de réseau bayésien apprise.

3.4. Phase de classification (reconnaissance de symboles)

Afin de reconnaître un symbole nous appliquons l'inférence probabiliste bayésienne sur la distribution de probabilités jointes encodée par le réseau. Nous utilisons pour cela la technique appelée *junction tree* (Jensen *et al.*, 1990) qui est la méthode exacte la plus populaire dans ce domaine et qui est proposée dans (Leray *et al.*, 2004). Cette technique consiste à utiliser la signature du symbole à reconnaître (appelée évidence) et à la propager dans le réseau afin de calculer les probabilités *a posteriori* pour chaque classe de symbole. L'équation 1 rappelle la règle de Bayes de notre système : la probabilité *a posteriori* d'une classe c_i sachant une évidence e est calculée d'après la vraisemblance $\Pr(e|c_i)$, la probabilité *a priori* $\Pr(c_i)$ et la vraisemblance marginale $\Pr(e)$. Cette dernière grandeur (voir éq. 3) sert à normaliser la probabilité *a posteriori*, c-à-d. que le résultat du calcul soit compris entre 0 et 1.

$$\Pr(c_i|e) = \frac{\Pr(e, c_i)}{\Pr(e)} = \frac{\Pr(e|c_i) \times \Pr(c_i)}{\Pr(e)} \quad [1]$$

où,

$$e = f1, f2, f3, \dots, f16 \quad [2]$$

$$\Pr(e) = \Pr(e, c_i) = \sum_{i=1}^k \Pr(e|c_i) \times \Pr(c_i) \quad [3]$$

Les probabilités *a posteriori* sont calculées pour toutes les k classes de symboles. Le symbole à reconnaître est affecté à la classe qui maximise cette probabilité *a posteriori*.

4. Expérimentations

L'organisation de concours de reconnaissance de symboles au cours de ces dix dernières années (Aksoy *et al.*, 2000, Valveny *et al.*, 2003, Dosch *et al.*, 2006, Valveny

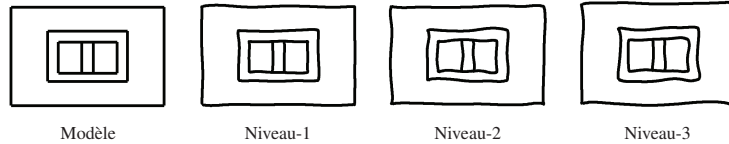


Figure 7. Un modèle de symbole (provenant de la base GREC2005) avec des exemples déformés utilisés pour simuler le dessin manuel.

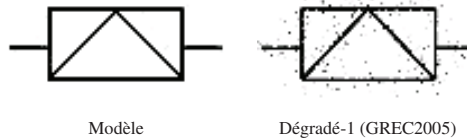


Figure 8. Un modèle de symbole (provenant de la base GREC2005) avec un exemple dégradé utilisé pour simuler la photocopie, l'impression ou la numérisation.

et al., 2007) a permis à la communauté de disposer aujourd'hui d'une base conséquente d'images test pour l'évaluation des nouvelles méthodes. Ces compétitions ont été organisées afin de tester la robustesse face au passage à l'échelle, aux dégradations binaires et vectorielles sur des symboles pré-segmentés (symboles de plans électroniques ou architecturaux).

Différentes versions de cette base ont été exploitées durant les concours successifs du colloque GREC: en 2005 (Dosch *et al.*, 2006) et 2007 (Valveny *et al.*, 2007) les bases étaient similaires tandis que la base GREC2003 (Valveny *et al.*, 2003) est un sous-ensemble de la base GREC2005.

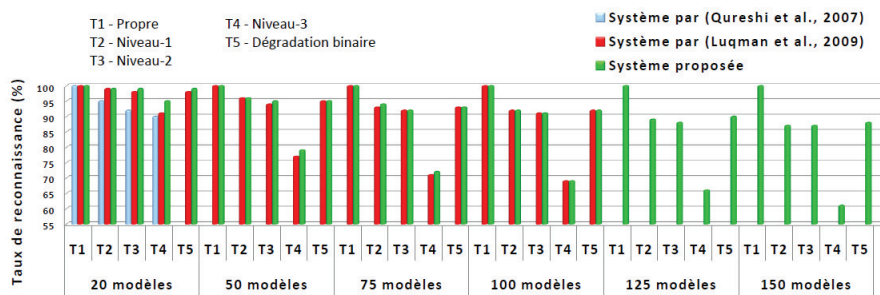
4.1. Reconnaissance de symboles dégradés par bruits binaires ou vectoriels

Nous avons mené des expérimentations sur une base de symboles synthétiques provenant de la base GREC2005¹. Afin d'obtenir une évaluation réaliste des performances de notre méthode intégrant une approche floue par rapport aux approches plus classiques nous avons produit plusieurs bases comportant 20, 50, 75, 100, 125 et 150 classes de symboles. Nous avons également généré nos propres bases d'apprentissage et de test basées sur les déformations et dégradations proposées lors de GREC2005 pour nos expérimentations. Rappelons que dans tous les cas, nous utilisons uniquement des symboles propres (non dégradés) pour l'apprentissage.

Pour chaque classe, un modèle parfait associé à 36 versions transformées par des rotations aléatoires et 12 par des changements d'échelle constituent la base d'apprentissage. Comme les caractéristiques que nous avons choisies ont déjà été démontrées invariantes aux rotations et changements d'échelle dans (Qureshi *et al.*, 2007, Luqman

1. <http://symbolcontestgrec05.loria.fr/symboldescription.php/>; visitée le 11 février 2010

Nombre des classes (modèles)		20	50	75	100	125	150
Symboles propres (rotation et changement d'échelle)		100%	100%	100%	100%	100%	100%
Hand-drawn déformation	Niveau-1	99%	96%	94%	92%	89%	87%
	Niveau-2	99%	95%	92%	91%	88%	87%
	Niveau-3	95%	79%	72%	69%	66%	61%
Dégradation binaire		99%	95%	93%	92%	90%	88%

Tableau 1. Résultats expérimentaux (bruits binaires et vectoriels).**Figure 9.** Comparaison des résultats avec ceux de (Qureshi et al., 2007) et (Luqman et al., 2009).

et al., 2009) et à cause du fait que les réseaux bayésiens sont réputés performants sur les bases comportant un grand nombre d'exemples, notre système a été aussi testé face à des symboles transformés par différents niveaux de dégradations vectorielles et de dégradations binaires selon le protocole défini dans les concours GREC (figure 7 et figure 8). Chaque test réalisé comporte 10 images requêtes pour chacune des classes de symboles utilisées lors du test.

Le tableau 1 résume les résultats fournis par notre système. Un taux de reconnaissance de 100% est obtenu pour les symboles non dégradés illustrant à nouveau l'invariance à la rotation et aux changements d'échelle de notre méthode même après intégration des intervalles flous. Notre méthode fait mieux que les participants aux concours GREC pour les tests de passage à l'échelle et fait aussi bien lorsque les niveaux de dégradations et déformations restent faibles, résultats consultables dans (Valveny et al., 2003, Dosch et al., 2006). Les taux de reconnaissance diminuent lorsque les niveaux de dégradations ou de déformations deviennent trop importants: l'étape de vectorisation devient alors moins efficace comme indiqué dans (Lladós et al., 2002), et la représentation sous forme de graphe de primitives devient inexploitable pour construire une signature numérique.

La figure 9 compare nos résultats avec ceux d'autres travaux. Dans (Qureshi et al., 2007) une signature proche de la notre, non floue, et un classifieur de type k-ppv reposant sur une distance euclidienne sont utilisés. Dans (Luqman et al., 2009) on utilise la même signature, mais sans approche floue avec réseaux bayésiens. Ces résultats permettent de juger des améliorations très significatives apportées par l'approche floue

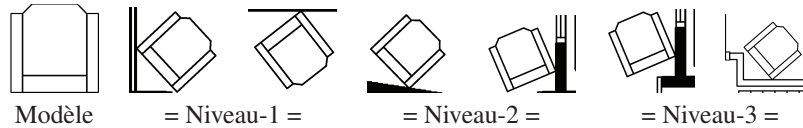


Figure 10. *Un fauteuil avec des exemples bruités (bruit contextuel).*

et les réseaux bayésiens : le système proposé peut traiter beaucoup plus nombre des classes. Les taux de reconnaissance obtenus montrent clairement la robustesse et le pouvoir de généralisation du système face aux déformation et dégradations réalistes et pouvant apparaître dans ces types de documents.

4.2. Reconnaissance de symboles en contexte (bruit contextuel)

Un second ensemble de tests a été réalisé sur une base de symboles synthétiques obtenus par extraction de régions d'intérêt dans un corpus d'images de documents complets proposés par (Delalandre *et al.*, 2007). Ces expérimentations visent à prouver la robustesse de notre méthode sur des images se rapprochant le plus possible de celles rencontrées dans les applications réelles. À notre connaissance et à ce jour, aucune autre publication ne propose de telles évaluations pour juger des performances applicatives réelles des méthodes proposées de reconnaissance de symboles.

Deux sous-ensembles d'images test ont été constitués à partir, d'une part, d'extraits de plans architecturaux comportant un symbole parmi 16 classes possibles et d'autre part, d'extraits de schémas électriques comportant un symbole parmi 21 possibles. Ces classes correspondent à celles exploitées dans les bases GREC2005 (Dosch *et al.*, 2006).

De la même manière que pour les expérimentations précédentes, l'apprentissage est réalisé à partir de symboles propres (non dégradés et sans contexte). Pour chaque classe, 36 symboles modèles sont obtenus par rotation et 12 symboles sont obtenus par changement d'échelle (réalisés avec ImageMagick²). Les images test sont composées d'images comportant différents niveaux de bruit contextuel comme décrit figure 10 et auxquelles différentes transformations affines (rotation et changement d'échelle) sont de plus appliquées aléatoirement. Cette base de test est disponible gratuitement sur Internet³.

Le tableau 2 résume les résultats de nos expérimentations sur ces images. Aucun pré-traitement de débruitage ou de correction n'a été appliqué sur les images. La vectorisation est appliquée directement sur les images brutes. Après construction du graphe et production de la signature floue correspondante, celle-ci est fournie directement au réseau bayésien pour classification.

2. <http://www.imagemagick.org/> ; visité le 11 février 2010

3. <http://mathieu.delalandre.free.fr/projects/sesyd/queries.html> ; visité le 11 février 2010

	Bruit	Nombre des classes (modèles)	Nombre de symboles requête (par classe)	Taux de reconnaissance	Taux de reconnaissance (top-5)
Symboles architecturaux	Niveau-1	16	100	89%	96 %
	Niveau-2	16	100	82%	93 %
	Niveau-3	16	100	78%	91 %
	moy.			83%	93%
Symboles électroniques	Niveau-1	21	100	72%	95%
	Niveau-2	21	100	69%	93%
	Niveau-3	21	100	65%	92%
	moy.			69%	93%

Tableau 2. Résultats expérimentaux (bruit contextuel).

Les résultats obtenus sur les schémas électriques sont moins bons que ceux obtenus sur les symboles architecturaux, nous pensons que cela est dû principalement à la présence de cercles et arcs plus difficilement vectorisables que les segments de droite. Concernant les résultats dans leur globalité, nous estimons que les taux obtenus sont très encourageants et témoignent de la capacité de notre méthode à exploiter avec une certaine souplesse l'information structurelle encapsulée de manière floue dans la signature vectorielle. Ces résultats démontrent également que cette signature permet de discriminer correctement des symboles en contexte.

5. Conclusion

Les méthodes structurelles sont les plus adaptées pour la représentation d'images de graphiques. Les méthodes statistiques constituent quant à elles des outils de classification très performants et rapides. En offrant un mécanisme permettant de convertir une représentation structurelle en un vecteur numérique adéquat, il devient possible de bénéficier des avantages des deux catégories d'approches. Notre proposition exploite des techniques de discrétisation et de fonction de appartenance floue pour produire une signature plus souple car intégrant des données incertaines.

Les caractéristiques de cette signature traduisent le contenu d'un graphe en termes de structure, de complexité et d'homogénéité des primitives (nœuds et arcs) qui le constituent. L'introduction du flou dans le mécanisme de production de la signature numérique constitue un moyen original et efficace de résoudre le problème de la sensibilité aux distorsions et aux bruits habituellement associée aux méthodes structurelles. Cela n'enlève aucune de ses propriétés d'invariance aux changements d'échelle et à la rotation.

Ensuite, une adaptation originale des mécanismes d'apprentissage des réseaux bayésiens nous permet de définir, dans un premier temps, la structure du réseau puis dans un deuxième temps, ses paramètres probabilistes. L'inférence probabiliste bayésienne permet alors de classer des symboles inconnus de manière pertinente même en contexte bruité.

Les tests expérimentaux effectués démontrent une amélioration significative des résultats au niveau de la robustesse et de la généralisation sans avoir à recourir à des algorithmes de prétraitement ou de nettoyage des images initiales. De part l'introduction de l'approche floue et des réseaux bayésiens, les résultats obtenus sur les bases du concours GREC sont au moins de niveau équivalent aux meilleurs résultats obtenus durant les compétitions ce qui nous a permis d'envisager de nouvelles évaluations sur des bases encore plus complexes. Ces dernières expérimentations, réalisées sur des symboles entourés d'un bruit contextuel, ont confirmé la bonne résistance de notre signature à l'information *parasite* apportée par l'environnement d'un symbole.

Nous pensons que ces résultats pourraient encore être améliorés si nous incorporions des symboles bruités, déformés ou en contexte dans les bases d'apprentissage (ce que nous avons refusé de faire jusqu'à présent). Le système peut être étendu facilement : le nombre de classes peut être important, les formes de nature variée (nombreux domaines applicatifs) et les temps d'exécution restent très raisonnables (la reconnaissance est pratiquement instantanée). Dans le futur, nous pensons adapter ces travaux afin de produire un système complet d'indexation et recherche d'occurrences de symboles (*spotting*) dans des dessins complets.

Remerciements

Ce projet a été partiellement supporté par le gouvernement Pakistanais sous la forme d'une bourse de doctorat attribué par la Higher Education Commission (ref. PD-2007-1/Overseas/FR/HEC/222).

6. Bibliographie

- Aksoy S., Ye M., Schauf M., Song M., Wang Y., Haralick R. M., Parker J. R., Pivovarov J., Royko D., Sun C., Farneback G., « Algorithm Performance Contest », *ICPR*, vol. IV, p. 870-876, 2000.
- Barrat S., Tabbone S., Nourissier P., « A Bayesian classifier for symbol recognition », *GREC*, vol. 7, HAL - CCSD, October 24, 2007.
- Bunke H., Gunter S., Jiang X., « Towards Bridging the Gap between Statistical and Structural Pattern Recognition : Two New Concepts in Graph Matching », *ICAPR*, vol. 2013 of *LNCS*, p. 1-11, 2001.
- Charniak E., « Bayesian Networks without Tears », *AI Magazine*, vol. 12, n° 4, p. 50-63, 1991.
- Chhabra A. K., « Graphic Symbol Recognition: An Overview », *LNCS*, vol. 1389, p. 68-79, 1998.
- Colot O., Olivier C., C. P., A. E. M., « Information criteria and abrupt changes in probability laws », *Signal Processing VII : Theories and Applications*, p. 1855-1858, 1994.
- Cordella L. P., Vento M., « Symbol recognition in documents: a collection of techniques ? », *IJDAR*, vol. 3, n° 2, p. 73-88, 2000.
- Coustaty M., Guillas S., Visani M., Bertet K., Ogier J.-M., « On the Joint Use of a Structural Signature and a Galois Lattice Classifier for Symbol Recognition », *GREC*, vol. 5046 of *LNCS*, Springer, p. 61-70, 2007.
- Delalandre M., Pridmore T., Valveny E., Locteau H., Trupin E., « Building Synthetic Graphical Documents for Performance Evaluation », *GREC*, vol. 5046 of *LNCS*, Springer, p. 288-298, 2007.
- Delaplace A., Brouard T., Cardot H., « Two Evolutionary Methods for Learning Bayesian Network Structures », *CIS*, vol. 4456 of *LNCS*, Springer, p. 288-297, 2006.

MM. Luqman, T. Brouard, JY. Ramel & J. Lladós

- Dosch P., Lladós J., « Vectorial Signatures for Symbol Discrimination », *GREC*, vol. 3088 of *LNCS*, Springer, p. 154-165, 2003.
- Dosch P., Valveny E., « Report on the Second Symbol Recognition Contest », *GREC*, vol. 3926 of *LNCS*, Springer, p. 381-397, 2006.
- Heckerman D., « A Tutorial on Learning with Bayesian Networks », *Innovations in Bayesian Networks*, vol. 156 of *Studies in Computational Intelligence*, Springer, p. 33-82, 2008.
- Jensen F. V., Lauritzen S. L., Olesen K. G., « Bayesian updating in causal probabilistic networks by local computations », *Computational Statistics Quarterly*, vol. 4, p. 269-282, 1990.
- Leray P., François O., « BNT Structure Learning Package: Documentation and Experiments », *Laboratoire PSI - INSA Rouen- FRE CNRS 2645*, November 15, 2004.
- Lladós J., Sanchez G., « Symbol recognition using graphs », *ICIP*, p. 49-52, 2003.
- Lladós J., Valveny E., Sanchez G., Martí E., « Symbol Recognition: Current Advances and Perspectives », *LNCS*, vol. 2390, p. 104-128, 2002.
- Lopresti D. P., Wilfong G. T., « A fast technique for comparing graph representations with applications to performance evaluation », *IJDAR*, vol. 6, n° 4, p. 219-229, 2004.
- Luo B., Wilson R. C., Hancock E. R., « Spectral embedding of graphs », *Pattern Recognition*, vol. 36, n° 10, p. 2213-2230, October, 2003.
- Luqman M. M., Brouard T., Ramel J.-Y., « Graphic Symbol Recognition Using Graph Based Signature and Bayesian Network Classifier », *ICDAR*, vol. 10, IEEE Computer Society, p. 1325-1329, 2009.
- Mitra S., Pal S. K., « Fuzzy sets in pattern recognition and machine intelligence », *Fuzzy Sets and Systems*, vol. 156, n° 3, p. 381-386, 2005.
- Neuhaus M., Bunke H., « Edit Distance-based Kernel Functions for Structural Pattern Classification », *Pattern Recognition*, vol. 39, n° 10, p. 1852-1863, 2006.
- Qureshi R. J., Ramel J.-Y., Cardot H., Mukherji P., « Combination of Symbolic and Statistical Features for Symbols Recognition », *ICSCN*, p. 477-482, 2007.
- Rusinol M., « Geometric and Structural-based Symbol Spotting. Application to Focused Retrieval in Graphic Document Collections », *PhD thesis, Universitat Autònoma de Barcelona*, 2009.
- Rusinol M., Lladós J., « Symbol Spotting in Technical Drawings Using Vectorial Signatures », *GREC*, vol. 3926 of *LNCS*, Springer, p. 35-46, 2006.
- Tombre K., Tabbone S., Dosch P., « Musings on Symbol Recognition », *GREC*, vol. 3926 of *LNCS*, Springer, p. 23-34, 2006.
- Valveny E., Dosch P., « Symbol Recognition Contest: A Synthesis », *GREC*, vol. 3088 of *LNCS*, Springer, p. 368-386, 2003.
- Valveny E., Dosch P., Fornes A., Escalera S., « Report on the Third Contest on Symbol Recognition », *GREC*, vol. 5046 of *LNCS*, p. 321-328, 2007.
- Valveny E., Martí E., « Deformable Template Matching within a Bayesian Framework for Hand-Written Graphic Symbol Recognition », *GREC*, vol. 1941 of *LNCS*, Springer, p. 193-208, 1999.
- Ventura A., Schettini R., « Graphic Symbol Recognition Using a Signature Technique », *ICPR*, vol. 2, p. 533-535, 1994.
- Zhang W., Liu W., « A New Vectorial Signature for Quick Symbol Indexing, Filtering and Recognition », *ICDAR*, vol. 9, p. 536-540, 2007.