



ATELIER ECOI 2006

Extraction de **C**onnaissance à partir d'**I**mages

proposé par



Systemes Intelligents de Perception Université Paris 5



Centre de Recherche en Informatique de Paris 5

Nicole Vincent

Nicolas Loménie

SOMMAIRE

Nicolas Loménie

* Titre : Présentation de l'Atelier ECOI – EGC 2006 - Extraction de COonnaissance à partir d'Images - Le contexte des ronds-points dans les images satellites

* Auteurs : Nicolas Loménie, Nicole Vincent et Guray Erus (Laboratoire SIP-CRIP5, paris)

* Résumé :

Nous présentons la motivation de cet atelier dans ce type de manifestation. Les experts des domaines Image et Connaissance commencent de plus en plus à avoir les même problématiques : on le perçoit avec l'émergence des ontologies en analyse d'image. D'une façon plus générale, il semble que l'approche statistique a pour l'instant épuisé la plupart de ses armes, et que le besoin se fait sentir de revenir aux considérations structurelles pour fournir une valeur ajoutée aux systèmes de traitement d'images à venir. Dans le contexte satellitale, par exemple, la très haute résolution change les besoins méthodologiques. Détecter et discriminer des ponts et des ronds-points dans une image satellite à 1 m de résolution se conçoit difficilement sans le support d'un modèle qualitatif (structurel) de l'objet à reconnaître. D'autre part, la construction de ce modèle doit être la moins contraignante possible (par exemple en terme d'étiquetage pour l'apprentissage) pour envisager un déploiement réaliste vers les utilisateurs ou acheteurs de ces images et les aider à naviguer dans des images de très grande taille quotidiennement en fonction d'un besoin spécifique. Tous ces problèmes qui revisitent la version statistique de la reconnaissance de formes et de l'analyse d'image s'appuient sur la structure de graphe sémantique et seront exposés selon différents angles au cours de cet atelier.

Noël RICHARD

* Titre : Image : entre pixels et graphes sémantiques

* Auteurs : Noël Richard, Denis Arrivault, Christine Fernandez-Maloigne, Philippe Bouyer (Laboratoire SIC, RC-SOFT)

* Résumé : Dans le cadre de la valorisation des bases d'images patrimoniales, nous travaillons sur les structures de données associées au contenu pixellique de l'image. Les techniques traditionnelles utilisent des structurent basées sur la construction de pyramides irrégulières à partir de la carte de segmentation. Afin de prendre en compte la variabilité des représentations, deux évolutions de ce type de structure sont présentées : les structures à base de graphes aléatoires et celles à bases de graphes d'attributs hiérarchiques flous. La validation de ces approches est présentée suivant deux axes. Le premier sur des images issues de bases spécialisées (Columbia et médiévales) pour la validation de l'apport hiérarchique et spatial dans la description. Le second est développé pour des bases de symboles complexes manuscrits, ceci afin de s'affranchir des variations liées à la segmentation. Cette seconde série de travaux permet de comparer les approches par graphes aléatoires et graphes d'attributs flous. Dans une troisième partie, la démarche d'intégration de l'aspect sémantique est présentée avec la technique de validation de cette intégration.

Sébastien Sorlin

* Titre : Une mesure générique de la similarité de deux graphes

* Auteurs : Sébastien Sorlin, Olfa Sammoud, Christine Solnon et Jean Michel Jolion (Laboratoire LIRIS, Lyon)

* Résumé :

Mesurer la similarité d'images est un problème qui se pose dans de nombreuses applications, que ce soit pour rechercher des images, les classifier ou encore les confronter à des modèles pour les "reconnaître". Quand les images sont représentées par des graphes, on peut mesurer la similarité de deux images en appariant les sommets des graphes associés aux images de façon à retrouver leurs caractéristiques communes. Différentes sortes d'appariements de graphes, impliquant différentes mesures de similarité, ont été proposées, e.g., l'isomorphisme de (sous-)graphes pour mesurer l'équivalence ou l'inclusion de deux graphes, et la distance d'édition de graphes pour mesurer le coût de transformation d'un graphe en un autre. Ces différents appariements sont "univoques" dans le sens où un sommet d'un graphe ne peut être apparié qu'à au plus un sommet de l'autre graphe.

Nous proposons ici une nouvelle mesure de similarité de graphes basée sur un appariement multivoque des sommets des graphes, chaque sommet pouvant être apparié à un ensemble de sommets de l'autre graphe. Cet aspect nous permet notamment de comparer des graphes décrivant des images à différents niveaux de granularité, et de prendre en compte les problèmes de sur- et de sous-segmentation des images. Notre mesure de similarité est générique dans le sens où elle est paramétrée par des fonctions de similarité qui permettent d'exprimer des connaissances de similarité propres à l'application considérée. En particulier, nous montrons que toutes les mesures de similarité proposées jusqu'alors peuvent être vues comme un cas particulier de notre mesure. Nous nous intéressons également au problème du calcul de cette similarité, et nous proposons et comparons trois algorithmes : un algorithme glouton, un algorithme basé sur une recherche locale Taboué réactive et un algorithme basé sur l'optimisation par colonies de fourmis.

Stéphanie Guillas

* Titre : Comment utiliser le treillis de Galois en Reconnaissance d'Images

* Auteurs : Stéphanie Guillas, Karell Bertet, Jean-Marc Ogier (Laboratoire L3I, La Rochelle)

* Résumé :

Le treillis de Galois ou treillis de concept a été utilisé à de nombreuses reprises comme outil de classification, notamment dans le domaine de la fouille de données. L'objectif de cette présentation est de proposer une description générique de son utilisation en classification, afin de mettre en évidence le fait que le treillis de Galois est plutôt approprié à des données détériorées. Nous ciblons ensuite les points clé de cette approche pour permettre une adaptation du treillis de Galois dans le cas spécifique où les données sont des images de symboles bruités.

Présentation de l'Atelier ECOI – EGC 2006

Extraction de CONnaissance à partir d'Images

Le contexte des ronds-points dans les images satellites

Nicolas Loménie, Nicole Vincent, Guray Erus

SIP-CRIP5, Université René Descartes, lomenie@sip-crip5.org

1 Introduction

Voici la seconde édition de cet atelier à la croisée des chemins : l'Image et la Connaissance. Comme nous l'expliquions dans le cadre de l'atelier 2005, la communication entre les deux communautés en est encore à ses balbutiements. Après le raisonnement spatial, le thème fédérateur choisi cette année est l'utilisation de la représentation par graphes au sens large à la fois en apprentissage et en classification.

Ainsi les trois contributions essentielles présentées ici abordent cette problématique : Stéphanie Guillas au sein du laboratoire L3I utilise une représentation par graphe particulièrement intéressante dans le cadre de données détériorées que l'on appelle les treillis de Galois ou treillis de concept. Noël Richard au sein du laboratoire SIC en collaboration avec RC-SOFT explore la gamme des graphes aléatoires, flous et pyramidaux pour contrer le problème bien connu de robustesse introduit par la nécessité de segmenter l'image pour obtenir ces graphes sémantiques. Les limites des processus bas-niveau comme la segmentation sont encore loin d'autoriser des bonds significatifs au niveau de l'intégration de la sémantique. Enfin, Sébastien Sorlin du laboratoire LIRIS affronte le délicat problème à la fois de la mesure de similarité entre graphes et de l'explosion combinatoire de ces représentations, notamment en expérimentant des méthodes dites méta-heuristiques comme les colonies de fourmis.

Dans cet atelier seront donc abordés les trois thèmes centraux de la Connaissance dans les Images :

- la représentation structurelle et sémantique : théorie des graphes ;
- les limites de la segmentation structurelle en régions, actuellement peu générique et robuste : analyse d'image ;
- le problème de complexité numérique : théorie de l'optimisation.

Tous ces thèmes sont loin d'être maîtrisés pour concevoir des systèmes industriels alors que, comme le notait Daniel Duclos du groupe SAGEM dans l'atelier précédent (Loménie et al, 2005), « les enjeux industriels, pour les applications, civiles ou militaires, notamment en robotique extérieure, en analyse d'images aéroportées, en guidage d'engins, ... sont majeurs, particulièrement dans des contextes où des décisions doivent être prises sans l'intervention d'un opérateur humain pour valider les résultats de traitements. »

2 Contexte de l'atelier

Dans ce contexte, nous travaillons au laboratoire SIP sur les images satellites à très haute résolution depuis quelques années. Nous avons épuisé un certain nombre de techniques essentiellement statistiques pour essayer de détecter et classifier les ponts et les ronds-points dans une image. Or, cette démarche fonctionne à terme sur la base d'images d'apprentissage et en réglant les paramètres adéquatement mais échoue à coup sûr si l'on change soit la base, soit les conditions d'acquisition. Cette limite empêche tout développement d'envergure des algorithmes proposés. D'autre part, cette démarche est tributaire d'un très grand nombre d'exemples d'apprentissage, ce qui d'un point de vue utilisateur (les acheteurs d'images satellites) n'est pas vraiment envisageable.

Partant des travaux de Sylvie Philipp du laboratoire ETIS (Philipp-Folliguet, 2005), nous nous intéressons de plus en plus à un système d'apprentissage bouclé sur l'utilisateur pour indexer structurellement les objets cartographiques dans une image satellite de taille 24000x24000 pixels. C'est ce que nous tentons d'expliquer dans la suite.

Les systèmes de détection d'objets cartographiques sur les images satellitaires développés au sein du laboratoire SIP utilisent un modèle structurel de l'objet recherché pour déterminer l'appartenance des candidats à une certaine classe d'objet. Ce modèle est pour l'instant généré manuellement, soit en utilisant des règles fixes (Sanz et al., 2003), soit en précisant des contraintes géométriques (Sanz et al., 2004). Pour rendre l'approche générique à différents objets cartographiques sans avoir à entreprendre l'étape de modélisation humaine à chaque fois, l'objectif de notre travail est de générer les modèles structurels d'objets cartographiques automatiquement à partir d'une base d'images satellitaires segmentées type SPOT5-THR à 2.5 mètres de résolution.

Nous avons mis entre parenthèses l'approche système expert à base de règles décrivant les agencements de régions étiquetées par un apprentissage radiométrique (Sanz et al., 2003)(Loméniet et al., 2003). en particulier en raison de la difficulté de l'étiquetage pour l'apprentissage supervisé des classes de régions (eau, verdure...).

Au fur et à mesure de notre réflexion, nous avons estimé que la construction du modèle qualitatif/structurel de l'objet ne pouvait pas se faire de façon réaliste par rapport à l'usage sur une grande base d'exemples. Nous proposons donc un système semi-automatique devant fonctionner sur un petit nombre d'exemples fournis de façon brute sous la forme de vignette non étiquetées par un utilisateur final. Le modèle structurel sur lequel nous nous concentrons à présent utilise des ARGs décrivant essentiellement la structure géométrique et topologique 2D de l'objet composite à étudier. C'est ce que nous décrivons dans ce qui suit.

Avant de continuer, nous pensons qu'il faut souligner quelques points :

- Sur un problème aussi complexe, à l'heure actuelle, l'intervention de l'utilisateur final est nécessaire ;
- En conséquence, pour concevoir un outil générique (adapté à plusieurs types de structures composites), les exemples ne doivent pas être très nombreux (quelques unités);

- Un processus d'apprentissage de plus en plus fin de ce que l'on cherche avec l'utilisateur dans la boucle est souhaitable ;
- Les exemples devront être centrés dans l'imagette à quelques pixels près ce qui est le cas dans les exemples d'apprentissage;
- La méthodologie doit permettre à un utilisateur d'explorer une très grande image satellite (24000x24000) en désignant seulement quelques vignettes d'intérêt dans l'image, ce qui nous rapproche d'un problème d'indexation dans lequel la structure de l'image est plus importante que les caractéristiques radiométriques (ce qui se fait encore peu même dans un domaine aussi exploré que l'indexation d'image) ;
- Il n'y a pas de distorsion géométrique importante.

Comme le note certains experts en analyse d'images, ce qui manque aux processus de traitement d'image actuels c'est une modélisation plus qualitative de la scène à analyser. A l'inverse, on pourrait essayer d'utiliser comme aide une modélisation qualitative de l'objet composite d'intérêt donné graphiquement par l'utilisateur. C'est l'idée que nous développons à présent.

Parallèlement aux images radiométriques I, nous utiliserons des graphes relationnels attribués (ARGs) pour décrire structurellement l'information contenue dans l'image I. Un ARG décrira une configuration de régions et de primitives géométriques adjacentes dans l'image. Ce graphe sera la structure d'accueil de nos modèles qualitatifs de scènes et d'objets. Ces graphes sont cachés dans les images I réelles et doivent apparaître après un processus de segmentation adaptée à notre problématique. L'idée est qu'il doit exister un ARG plus ou moins commun caché dans les vignettes d'apprentissage qui constitue le modèle qualitatif de l'objet recherché.

Le but de ce volet méthodologique est de détecter des objets possédant des structures classiques de la géométrie euclidienne (parallèles, rectangles, cercles) connectées (dans un ARG par exemple). Il y a trois parties à développer dans ce volet qui correspondent aux trois modules qui ressortent de la figure 1.

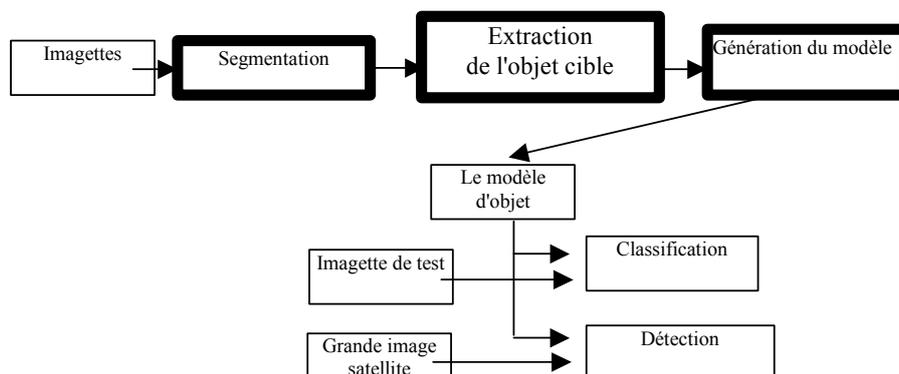


Figure 1. Architecture globale du système

L'entrée du système est un nombre prédéterminé d'images échantillons de l'objet cible dans lequel l'objet est centré, plus éventuellement un modèle simple G_M de l'objet recherché dessiné par l'utilisateur et traduit en un ARG (voir figure 2).

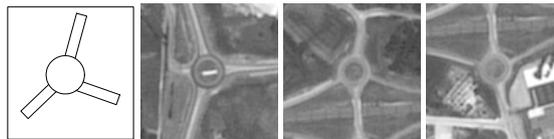


Figure 2. Entrées du système

Ces images seront segmentées (et labellisées) en utilisant un algorithme de segmentation géométrique. Dans chaque image segmentée (sous la forme d'ARGs), les parties appartenant à l'objet cible seront extraites en utilisant les similarités entre les ARGs. Les images sur lesquelles on extrait les objets cibles seront utilisées pour la génération d'un modèle de l'objet utilisant des formes géométriques simples comme des rectangles et des cercles. Comme premier objectif, le modèle sera utilisé pour la classification d'une image exemple dans une des classes d'objet données. Ensuite, le modèle sera utilisé dans la détection d'objets similaires dans une grande image satellite.

2.1 Module 1 : La segmentation géométrique des vignettes

Dans cette partie, qui pourrait constituer un outil à part entière pour la communauté de l'image, on veut orienter la segmentation région classique par les primitives géométriques présentes dans l'image et éventuellement par le modèle qualitatif proposé en entrée.

L'algorithme s'appuierait sur une sur-segmentation de l'image par la transformé de Ligne de Partage des Eaux (LPE), suivi d'un algorithme de fusion des régions (Guigues, 2003). La fusion sera orientée par une fonction de distance déterminée par les deux contraintes suivantes:

- Radiométrique: La distance radiométrique maximale (Comme c'est le cas pour la plupart des méthodes existantes) qui mesure la similarité radiométrique des régions adjacentes.
- Géométrique: Apres la détection des primitives géométriques dans l'image (nous les appellons les contours potentiels), la distance de séparation par des contours potentiels .

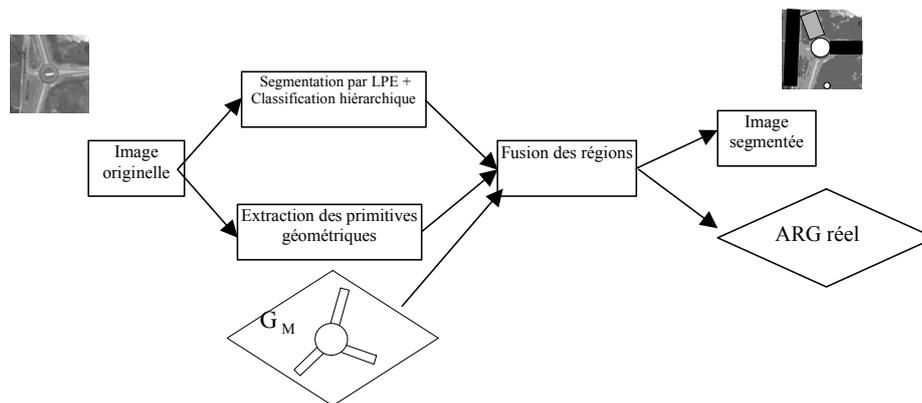


Figure 3. Segmentation géométrique

2.2 Module 2 : Extraction de l'objet cible par rapport au fond de l'image

C'est sans doute la partie la plus délicate à modéliser. Un point extrêmement dur de nombreuses applications qui traitent de l'image. Comment segmenter ce qui est **structurellement commun** à plusieurs images ? Nous pensons que l'on doit pouvoir s'aider des ARGs qualitatifs cachés dans chacune de ces images. La difficulté est d'extraire les meilleurs ARGs possibles dans chaque image pour avoir une chance de trouver des points structurels (qualitatifs) commun entre les images. En effet, en sortie de l'étape de segmentation géométrique, on doit pouvoir extraire un ARG représentant la structure de l'image I dans lequel on trouvera des régions sans formes spécifiques et des régions ayant des attributs géométriques euclidiens forts (circularité, parallélisme). L'ARG modèle en serait alors un sous-ARG plus ou moins commun à toutes les images exemples et serait une amélioration de l'ARG modèle dessiné par l'utilisateur (voir module 3).

Soient deux vignettes images d'apprentissage I_1 et I_2 . En sortie de l'étape 1, on a les deux images segmentées I_{S1} et I_{S2} en régions et leurs ARGs associés (ou Graphes Réels) G_{R1} et G_{R2} . L'idée est d'extraire les objets d'intérêts dans l'image I_{S1} et I_{S2} en cherchant les similarités entre G_{R1} et G_{R2} pour obtenir les images I_{E1} et I_{E2} dans lesquelles les objets cibles sont extraits. A ce point là, à notre connaissance, il n'existe aucune méthodologie applicable telle quelle même si nous puissions des idées de la méthodologie adoptée dans (Taillandier, 2004). Parallèlement on a deux ARGs exemples G_{E1} et G_{E2} pour l'objet cible (figure 4).

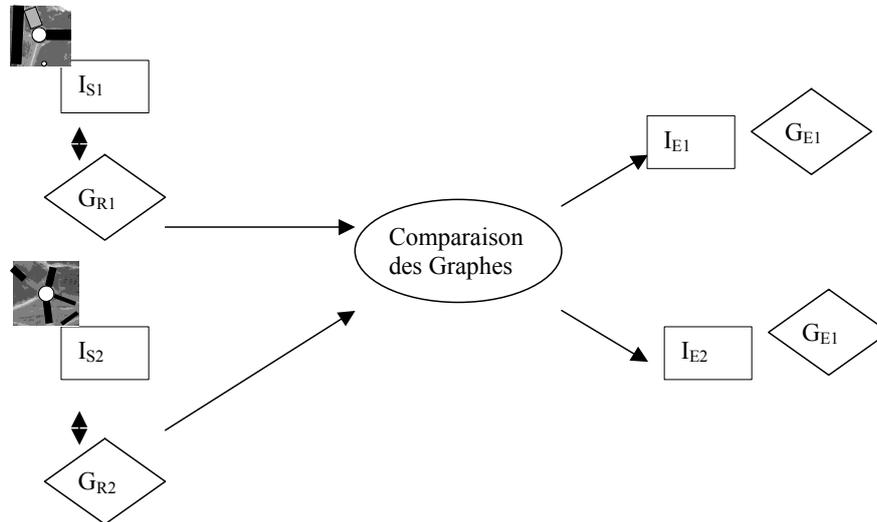


Figure 4. Schéma d'extraction des objets cibles à partir de plusieurs images

2.3 Module 3 : Apprentissage du modèle qualitatif de l'objet cible

Cette partie a déjà donné lieu à publications (Erus G et Loménie N., 2005) qui expliquent comment à partir de plusieurs ARGs exemples G_{Ei} d'un objet cible obtenir un ARG modèle G_M qui viendra mettre à jour le graphe modèle G_M initié par l'utilisateur au début de l'étape 1. Ces travaux s'appuyant sur les recherches développées dans (Cordella et al., 2002)(Bunke et al., 1998).

3 Conclusion

En conclusion, nous avons élaboré une stratégie réaliste du point de vue de l'utilisateur pour la caractérisation structurale et la détection d'objets composites dans les image satellitales dont les trois modules sont décrits dans les parties précédentes et qui reposent sur deux pré-requis essentiels : peu d'exemples et bouclage avec l'utilisateur dans l'esprit d'un système d'indexation de type « online query by content ». Notre laboratoire qui est plus orienté analyse d'image bas-niveau commence à s'intéresser à ces techniques fortement liées à la sémantique du processus. Aussi, grâce aux trois articles de fond qui suivent cette courte mise en situation, nous espérons que les lecteurs de ces actes seront éclairés sur les problématiques de recherche qui sous-tendent nos ambitions.

Je reprendrai la conclusion de Daniel Duclos (Loménie et al., 2005) exprimée dans l'atelier 2005, plus que jamais pertinente et qui caractérise l'esprit de cette réunion des réflexions menées par les experts en Image et en Connaissance : « Force est de constater que le chemin à parcourir reste long et qu'il passe par une collaboration étroite entre l'ensemble des communautés intéressées par l'acquisition, la représentation et l'exploitation de la Connaissance. Cette collaboration sera sans nul doute particulièrement riche et stimulante, à la hauteur des challenges à relever, comme l'ont montré toutes les tentatives menées jusqu'ici dans cette voie. »

Références

- Bunke H. and Shearer K.** (1998), A Graph distance metric based on the Maximal Common Subgraph, *Pattern Recognition Letters*, Vol. 19, 1998.
- Cordella L. P., Foggia P., Sansone C., Vento M.** (2002), Learning structural shape descriptions from examples, *Pattern Recognition Letters* 23(12): 1427-1437.
- Erus G., Loménie N.** (2005), Automatic Learning of Structural Models of Cartographic Objects, *IAPR - Workshop on Graph-based Representations in Pattern Recognition*, Poitiers.
- Guigues L.** (2003), Modèles multi-échelles pour la segmentation d'images, thèse de doctorat, IGN.
- Loménie N., Trias R., Barbeau J.** (2003), Integrating Textural and Geometric Information for an automatic bridge detection, *IGARSS'03*, Toulouse, France, juillet 2003.
- Loménie N., Stamon G., Duclos D.** (2005), Introduction à l'atelier ECOI 2005 : Extraction de Connaissances à partir d'Images, *EGC'05*, Actes des Ateliers – Atelier ECOI, Ed. CEPADUES.
- Philipp-Folliguet S.** (2005) , Recherche d'Images par le Contenu : requête globale ou partielle, apprentissage en ligne, *EGC'05*, Actes des Ateliers – Atelier ECOI, Ed. CEPADUES.
- Sanz R. et Loménie N.** (2003), Automatic bridge detection in high-resolution satellite images, *3rd International Conference on Computer Vision Systems (ICCVS'03)*, ser. LNCS. Graz, Austria : Springer, pp. 172-181, Apr. 2003.
- Sanz R.T, Loménie N. et Barbeau J.** (2004), Using textural and geometric information for an automatic bridge detection system, *Advanced concepts for intelligent vision systems, ACIVS'04*, Bruxelles, Belgium, pp. 172-181, Aug. 2004.
- Taillandier F.** (2004), Reconstruction du bâti en milieu urbain : une approche multi-vues, Thèse de l'Ecole Polytechnique.

Image: entre pixels et graphes sémantiques

Noël Richard¹, Denis Arrivault^{1,2}, Christine Fernandez-Maloigne¹, Philippe Bouyer²

¹Laboratoire SIC, SP2MI, 86962 FUTUROSOCOPE,

courriel: nom@sic.univ-poitiers.fr

²RC-SOFT, Domaine de la combe, BP39, 16710 Saint Yriex *

Résumé

Dans le cadre de la valorisation des bases d'images patrimoniales, nous travaillons sur les structures de données à associer au contenu pixellique de l'image. Les techniques traditionnelles utilisent des structures basées sur la construction de pyramides irrégulières à partir de la carte de segmentation. Afin de prendre en compte la variabilité des représentations, deux évolutions de ce type de structure sont présentées : les structures à base de graphes aléatoires et celles à base de graphes d'attributs hiérarchiques flous. La validation de ces approches est présentée suivant deux axes. Le premier sur des images issues de bases spécialisées (Columbia et médiévale) pour la validation de l'apport hiérarchique et spatial dans la description. Le second est développé pour des bases de symboles complexes manuscrits (hiéroglyphes et caractères grecs anciens), ceci afin de s'affranchir des variations liées à la segmentation. Cette seconde série de travaux permet de comparer les approches par graphes aléatoires et graphes d'attributs flous. Dans une troisième partie, la démarche d'intégration de l'aspect sémantique est présentée avec la technique de validation de cette intégration.

1 Introduction

L'analyse d'images s'est essentiellement développée autour des méthodes statistiques et du traitement du signal. Longtemps les développements conduits se sont basés sur le contenu pixellique de l'image dans ses aspects spatiaux et/ou fréquentiels, ou bien sur des modèles probabilistes de contenu. De nombreuses voies ont exploré les informations pixelliques pouvant être extraites à partir des informations de couleur, de texture ou de forme. Les applications de ces développements dans les systèmes d'indexation de grandes bases d'images ont cependant montré les limites de ces approches, notamment dans leur incapacité à saisir le contenu sémantique de l'image. Pour s'approcher de cette capacité, il est indispensable travailler sur l'organisation spatiale du contenu dans ses aspects de localisation relative et d'inclusion. Différentes solutions existent pour cela, exploitant généralement des structures de représentation à partir de graphes.

Comparer deux images ou une image avec un modèle est une opération de base de tout système de traitement d'images. La construction de métriques d'images à partir de graphes exploite des méthodes d'appariement qui permettent d'extraire le meilleur sous-graphe commun à deux représentations. Cette étape préliminaire permet ensuite d'établir une mesure de similarité. L'adaptation sémantique de ces mesures impose d'approcher la notion de différence perceptuelle. Santini[19] exprime ce type de similarité comme une dichotomie liée à deux éléments, l'un visant à trouver l'ensemble des attributs significatifs et l'autre exprimant une métrique dans l'espace des attributs.

*Remerciements : Merci à Eric Palazzo, directeur du CESCMI, et à Aurélia Bolot, responsable de la photothèque pour l'expertise et l'accès aux images de peintures médiévales.

Pour Santini, l'intégration de l'aspect perceptuel exploite un formalisme flou. Le découpage du problème en deux groupes permet de choisir où intégrer cette adaptation : soit dans l'espace des attributs, soit dans la métrique. Pour exprimer notre travail dans ce cadre, nous présentons tout d'abord différentes voies de représentation du contenu spatial : pyramide de graphes, graphes aléatoires et graphes d'attributs hiérarchiques flous. Nous proposons quelques résultats permettant d'apprécier l'apport de ce type d'information et les limites encourues. Afin de quantifier correctement cet apport, le problème a été dérivé sur une base de caractères ne présentant pas de difficultés de segmentation. Enfin, nous expliquons nos développements actuels sur l'intégration de connaissances "expert" et les contraintes de validation qui sont posées par ce travail spécifique.

2 Capter l'organisation spatiale

La structuration spatiale du contenu d'une image prend comme point de départ une carte de segmentation. Ce qui sous-tend que celle-ci soit optimale face au contenu sémantique de l'image, dès lors la structure topologique utilisée doit pouvoir intégrer au minimum les relations d'adjacence entre les objets. Suivant les structures utilisées, les relations d'inclusion et de localisation relative entre les objets sont également exploitées. Les graphes d'adjacences de région (ARG) ont été les premières structures mises en oeuvre pour l'analyse d'images, cependant elles furent rapidement remplacées par des structures plus adaptées à la description spatiale ou topologique comme les cartes combinatoires[7], les pyramides irrégulières[17], pyramides de graphes[10], ...

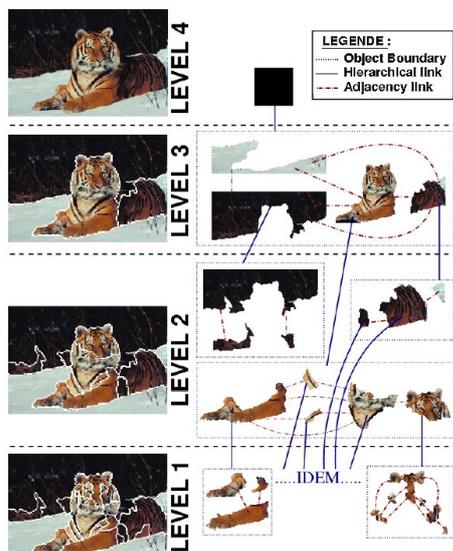


FIG. 1 – Décomposition par pyramides de graphes.



FIG. 2 – Recherche dans base médiévale

2.0.1 Graphes aléatoires

Pour comparer deux contenus d'images à partir de graphes, il est possible d'utiliser des algorithmes spéciaux d'appariement et de mesure de similarité [13, 18]. Cependant la segmentation

d'images ne produit pas forcément une organisation directement plaquée sur l'organisation sémantique du contenu. Pour s'affranchir de cette difficulté, deux stratégies existent : soit sous-segmenter l'image (directement ou par fusion contrôlée topologiquement), soit introduire de la souplesse dans le système d'appariement et de mesure de la similarité. Dans ce cas, une solution naturelle consiste à utiliser une description exploitant des graphes aléatoires[1]. L'intérêt de ce type d'approche est d'arriver à exprimer la variabilité des entités graphiques sous forme de distributions : distributions d'association, distributions d'attributs sur les noeuds et arcs. Par définition, un graphe aléatoire correspond à un graphe dont chaque élément est une variable aléatoire :

définition : Un graphe aléatoire est une paire $R = (W, B)$ telle que :

1. W représente l'ensemble des noeuds aléatoires avec W un n-uplet $(\alpha_1, \alpha_2, \dots, \alpha_n)$ où chaque α_i , appelé noeud aléatoire, est une variable aléatoire ;
2. B , représente l'ensemble des arêtes aléatoires avec B un m-uplet $(\beta_1, \beta_2, \dots, \beta_m)$ où chaque β_i , appelé arête aléatoire, est une variable aléatoire ;
3. Pour chaque graphe $G = (N, A)$ réalisation de R , il est associé un isomorphisme de sous graphe : $\phi : G \rightarrow R$ et une probabilité $p(G, \phi) = Pr(R = M(G), \phi \in \Phi)$. La famille des isomorphismes de sous-graphe Φ est telle que :

$$p(G, \phi) \geq 0, \forall G \in \Gamma$$

$$\sum_{\Gamma} p(G, \phi) = 1 \quad (1)$$

A partir de cette définition et en exprimant chaque attribut d'arête et de noeud comme une variable aléatoire (moyennant une base d'apprentissage correctement constituée), il est possible de construire pour chaque classe a un graphe modèle G_a , qui soit un graphe aléatoire (figure 3).

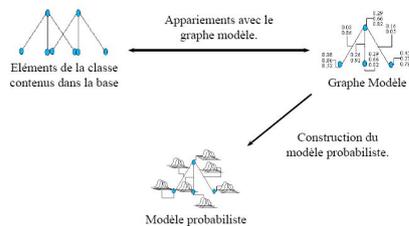


FIG. 3 – Apprentissage de graphes aléatoires.

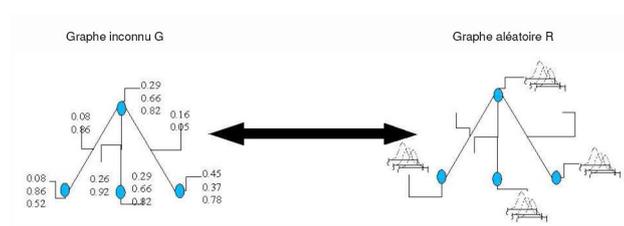


FIG. 4 – Reconnaissance de graphe aléatoire.

L'exploitation de cette construction nécessite l'estimation de la probabilité ϕ d'affectation du graphe G au graphe modèle G_a de la classe a (figure 4). L'appariement préliminaire au calcul de la similarité exploite des mesures de vraisemblance entre noeuds et des mesures de vraisemblances entre arêtes. La phase de décision nécessite quant à elle des hypothèses d'indépendance entre les noeuds, les arêtes et des distributions conditionnelles des arêtes relatives aux noeuds (en dehors des extrémités concernées par les arêtes elles-mêmes).

Théoriquement les graphes aléatoires présentent de nombreux points d'intérêt, ne serait ce que par le formalisme embarqué. Cependant les hypothèses liées sont difficilement vérifiables (indépendance des attributs entre eux, complétude de l'affectation des noeuds et arcs à la suite de l'appariement, ...).

2.0.2 Les graphes d'attributs hiérarchiques flous

La meilleure alternative actuelle exploite des graphes d'attributs, dont les attributs sont des arbres linguistiques flous (figure 6)[2]. L'objectif est d'offrir une précision de description croissante avec la profondeur d'analyse de l'arbre. L'intégration de variables linguistiques permet d'introduire un niveau de description sémantique, l'association avec des degrés d'appartenance flous assure la souplesse nécessaire pour capter la variabilité des objets. Ainsi suivant le niveau de précision souhaité, les deux objets de la figure 5 peuvent être soit similaires, soit différents.

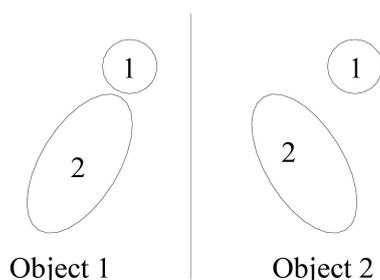


FIG. 5 – Deux objets avec similarité

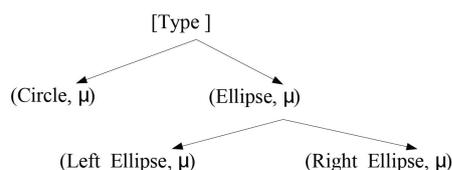


FIG. 6 – attribut linguistique arborescent

Man and Poo[16] ont proposé une mesure de similarité entre deux graphes d'attributs flous appariés par un isomorphisme. Dans [2], Arrivault propose l'extension de la mesure aux graphes d'attributs hiérarchiques de profondeur d . L'association de l'ensemble à une fonction d'appariement adaptée aux graphes d'attributs flous permet de disposer des éléments de mesure de similarité entre deux graphes[8].

Cette approche offre une solution naturelle à la mesure de similarité, mais ne permet pas de déterminer un graphe modèle automatiquement. Dans notre cadre supervisé, une heuristique simple permet de spécifier que le graphe modèle d'une classe est le graphe le plus simple afin de limiter l'effet des artefacts. Cependant sous cette formulation, le modèle correspond à l'individu le plus représentatif et non l'individu moyen tel que le préconise le graphe aléatoire.

3 Résultats

3.1 Cas de l'image

En combinant une approche par graphe avec une approche multirésolution au sein d'un graphe hiérarchique ou d'une pyramide de graphes, nous obtenons une structure autorisant la plupart des modes de recherches/comparaison [11]. L'aspect multirésolution a pour objet de rendre notre structure indépendante de la taille de l'objet dans la scène, en stockant de façon hiérarchique les relations d'inclusion et les informations associées. L'aspect graphe autorise quant à elle la représentation des relations d'adjacences sans autres informations topologiques.

Dombre[10] a étudié l'apport du graphe pyramidal, issu des pyramides irrégulières, pour l'indexation des bases d'images couleur. Il montre les gains dus à l'intégration de l'information spatiale. La figure 9 présente les courbes précision¹ obtenues par recherche uniquement basée sur les attributs globaux à l'image(pointillé), sur une combinaison entre ces derniers et les attributs issus

¹Résultats sur la base Columbia, attribut associé aux noeuds : histogramme couleur exprimé dans l'espace Lab.

de la pyramide de graphes (batons rompus), enfin sur seulement ceux issus de la pyramide (trait continu)[12].

Dans un cadre de bases spécialisées (base "peinture" du CESCO²), la question posée porte sur la capacité à rechercher des éléments sémantiques particuliers dans la base (mains, drapés, ...). Une première évaluation portant sur la recherche d'apôtres dans une sous-partie de la base montre la capacité du système à retrouver les images de plans élargis et celles de détails comportant le même contenu (figure 2, image requête en haut à gauche). Les attributs associés à chaque noeud correspondent à l'information de couleur (histogramme de la région, associé à une distance EMD), de forme (Fourier, Zernicke, circularité, rectangularité et excentricité) et de texture (attributs de Haralick issus de la matrice de Cooccurrence). Bien qu'il soit délicat sur cette base d'établir un benchmark afin d'évaluer l'efficacité du système, les résultats obtenus permettent d'espérer avancer vers le contenu sémantique de l'image.

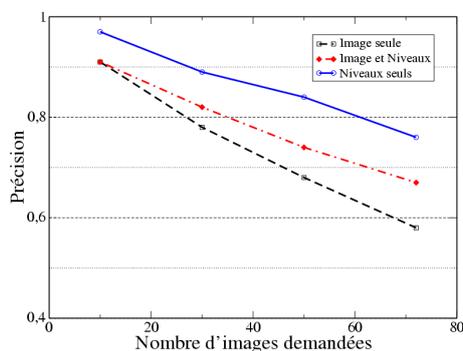


FIG. 7 – Gain par apport spatial

Cependant le point noir de la proposition est le coût combinatoire des techniques d'appariement de graphes ou de parties de graphes [20]. Même des techniques de mesure de similarité construites à partir d'outils à optimisation non linéaire présentent des coûts rédhibitoires pour des bases d'images conséquentes[13]. Le principal problème pour la confrontation avec un graphe sémantique, ou un graphe modèle, réside dans la segmentation qui ne conduit pas forcément à rendre adjacentes des régions sémantiques voisines[5]. La résolution de cette question passe souvent par une sous-segmentation de l'image contrôlée par un critère de type Borsotti[6]. Si l'information couleur est souvent prépondérante dans la décision, elle est également très sensible aux conditions d'illumination. Face à ces variations, l'assemblage topologique de l'objet représente la principale information invariante, malheureusement bruitée par les étapes de prétraitements et de segmentation.

3.2 Cas d'un symbole

Afin de s'affranchir des imprécisions liées à la segmentation, une étude spécifique a été menée sur des bibliothèques de symboles manuscrits complexes (hiéroglyphes, grecs ancien). Les symboles étant binaires, les seules paramètres influençant le résultats sont ceux de la squeletisation et de l'extraction de points d'intérêt entre les extrémités des segments. Dans une telle étude, la structure de graphe décrit la forme du symbole comme un assemblage de primitives simples (figure 8) ainsi que les positions relatives des primitives entre-elles.

²Centre d'Études Supérieures de la Civilisation Médiévale, UMR 6589, 24 rue de la chaîne, 86000 Poitiers.

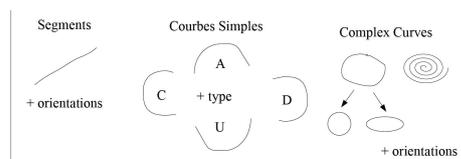


FIG. 8 – Types de primitives graphiques

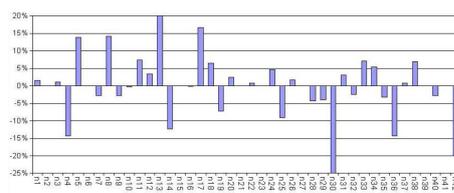


FIG. 9 – Différence en % de bonne détection entre une approche structurale et une approche par coopération (%coop-%stat)

3.2.1 Apport des graphes aléatoires

Cette étude a permis d'aborder l'apport réel de l'information structurale face à l'information statistique de forme classique, issue des descripteurs de région et de forme (géométrique, Zernicke, Fourier). La première gamme de résultats montre, par exemple, la complémentarité de l'approche par graphe avec une approche statistique. L'approche par graphe développée dans [3] exploite une coopération entre graphes aléatoires et attributs statistiques par gestion d'un rejet en ambiguïté à la suite d'une classification statistique. La base exploitée possède 42 classes de caractères grecs anciens manuscrits (4 scribes) pour un total de 17641 caractères.

Cette gamme de résultat montre également les limites de l'approche par graphe aléatoire : le point négatif provient du trop faible gain apporté et de la limite de la caractérisation des distributions par des modèles gaussiens. Ce second point est un handicap lourd pour des bases où certaines classes sont très faiblement représentées. Le point positif provient de la complémentarité avec des approches classiques : les erreurs sont distribuées, il n'y a pas redondance d'information. Néanmoins pour exploiter au mieux cet apport, une précision plus importante dans la description structurale est nécessaire.

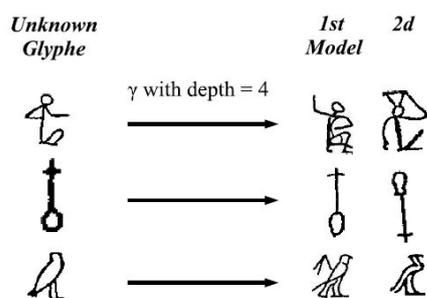


FIG. 10 – Exemples de structures retrouvées par GAHF



FIG. 11 – Exemple de frise géométrique et de frise végétale

3.2.2 Apport des graphes d'attributs hiérarchiques flous(GAHF)

Le GAHF permet d'embarquer de façon plus formelle les caractéristiques structurales de l'objet. Son apport s'évalue notamment sur des structures plus complexes, difficilement exploitables par des attributs statistiques usuels. Dans le contexte des hiéroglyphes, la complexité provient de

la construction des symboles avec différents niveaux d'inclusion et de bouclage. Mais surtout sur le point de vue de l'apprentissage, la complexité provient des 6000 classes potentielles. De plus pour nombre de classes il existe un très faible nombre de représentants (<10), avec des différences intervenant pour partie sur des détails associés à la sémantique (poitrine, couvre-chef, ...).

La figure 10 montre comment un classifieur utilisant un GAHF exploite la structure des symboles pour la reconnaissance (296 modèles issus d'un scripteur, reconnaissance sur une page écrite par un autre scripteur, 4 niveaux de profondeur utilisés dans l'arbre d'attributs). La constatation essentielle est que le système préserve dans sa décision le niveau de complexité des symboles (ce qui correspond également aux groupes sémantiques : humains, objets, animaux). Les réponses apportées pour l'objet "calvaire" montrent également comment le système peut trouver des ressemblance même par symétrie horizontale. Dans ce cas, le matching des primitives initiales connectées entre elles de façon similaire assure un bon score de reconnaissance face à celle des positions relatives codées par les arêtes du graphe.

4 L'obligation d'introduire une connaissance dans le traitement

Par un travail spécifique sur des objets graphiques de complexité limitée, il est possible de quantifier l'apport d'une approche structurelle pour l'analyse. Cependant dans un cadre d'analyse d'images plus global, la complexité combinatoire augmente exponentiellement. Dans le cas des symboles, la question est de trouver le graphe ou la structure qui s'apparie au mieux avec le graphe requête. Dans le cas de l'image la question est plus spécifiquement de trouver le sous-graphe (le sous-ensemble de l'image) qui s'apparie au mieux avec les graphes d'objets connus. Si dans un cadre spécialisé le nombre d'objets connus est défini et parfois limité, cela ne peut être le cas pour les bases multimédia. De plus, l'accroissement de la définition des images génère également un accroissement de la précision sémantique souhaitée dans la recherche, soit un accroissement de la taille et de la profondeur du graphe.

Cette problématique soulève deux questions de fond : comment choisir et construire les modèles et deuxièmement, comment réduire le coût de la recherche et de l'appariement ?

4.1 Le graphe sémantique

Dans le cadre des bases d'images spécialisées, le contenu présente une unicité sémantique, sans que cela conduise à une unicité de contenu pixellique (problème de calibration, de variabilité du vivant, ...). La construction de modèles issus du savoir de l'expert est alors une voie de résolution possible [15, 21]. Ainsi l'approche développée par l'équipe d'Isabelle Bloch ([9] par exemple) exploite un modèle déformable, qui permet d'intégrer l'attachement aux données pixelliques en maintenant un contrôle topologique par le biais du contrôle spatial porté par un graphe hiérarchique attribué. Dans le cadre de l'imagerie médicale, l'atlas ainsi représenté est tridimensionnel et l'apprentissage basé sur une base d'images segmentées par des experts.

Nos travaux sur la valorisation de bases d'images patrimoniales, en liaison avec le CESCO, nous portent vers l'intégration de connaissances expertes dans le système d'analyse. Pour cela, l'intégration de modèles sémantiques associés aux images est une nécessité. Cependant dans ce cadre de peintures à contenu essentiellement liturgique (forte codification), l'aspect artistique et l'évolution temporelle des techniques induisent une variabilité forte. Ainsi, l'information principale pour le decryptage de l'image (segmentation) est l'information couleur, alors que celle-ci ne présente qu'un intérêt modéré dans la reconstruction du contenu sémantique face à l'organisation

spatiale³. Pour cela, notre approche s'appuie d'une part sur une base de cas, portant la trace de la variabilité, associée à une base de modèles. Compte tenu du domaine, les modèles sont décrits par des objets de base et un ensemble de règles associées, l'ensemble autorise ainsi une généralisation des concepts, ainsi que la possibilité d'affiner la connaissance liée au modèle (figure 12). Cette dernière contrainte est une nécessité pour répondre aux attentes des experts (comparaison des représentations des mains par exemple (figure 13), travail sur les thénars)

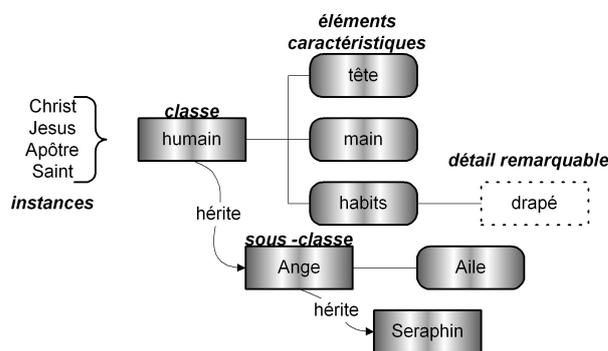


FIG. 12 – Extrait de description sémantique pour les images médiévales

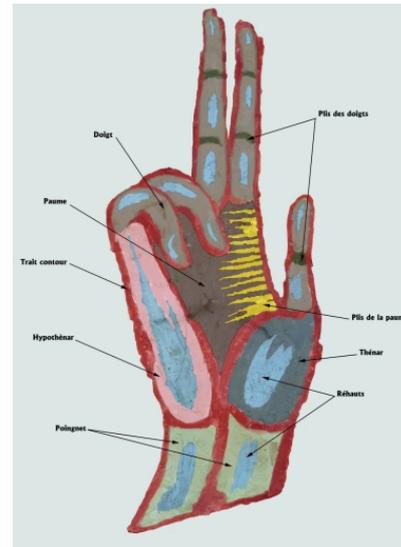


FIG. 13 – Niveaux de détails utilisés par l'expert

Valider l'approche Afin de pouvoir évaluer les algorithmes développés, un travail préliminaire est en cours pour l'adapter à la recherche de frises dans les peintures. L'intérêt réside dans l'extraction de structures plus simples, et surtout répétitives dans la même image. L'enjeu est la réduction de la variabilité liée à l'acquisition, ainsi que celle liée aux évolutions du support ou simplement de représentation. Il s'agit donc d'une recherche intra-images, à l'opposé de l'application finale qui est inter-images. Deux niveaux de frises sont concernés, les frises à base de motifs géométriques et les frises de motifs végétaux (figure 11).

5 Conclusion

Il existe de nombreuses structures possibles pour représenter le contenu d'une image ou la structure de symboles graphiques, les problématiques de choix sont généralement peu étudiées. Nous essayons de dériver les questions posées par ces choix et les paramètres de contrôle pour valider et optimiser l'apport de telles structures. Le problème du modèle est une question récurrente pour réduire les coûts de recherche, cependant les questions posées par la construction automatique de modèles de graphes n'ont pas de réponses complètes actuellement. A cette question, il faut également associer celle de la captation des variations des objets numérisés. Les variations sont de deux types, celles liées à l'objet et celles liées aux conditions d'acquisition. S'il est imaginable

³Un saint Jean s'identifie au fait qu'il porte un livre, une tablette ou un parchemin, à la différence de saint Pierre qui porte des clefs.

de réduire l'influence du second type par une maîtrise de la calibration pour les bases d'images spécialisées, en revanche pour le multimedia cela est impossible.

Après avoir travaillé sur la recherche d'invariants de couleur, de texture ou de forme, il apparaît que ceux-ci ne peuvent être que d'un intérêt limité. Depuis quelques années, cette quête d'invariance s'oriente plus naturellement vers des descripteurs sémantiques de bas niveau associés à des métriques floues[14]. De façon similaire, afin de conjuguer l'efficacité de discrimination avec la souplesse d'adaptation aux variabilités des objets, l'intégration d'une grammaire linguistique pour la description des relations de voisinage et des primitives offre toutes les garanties d'un dialogue simplifié avec l'expert ou l'utilisateur.

Les travaux décrits ici portent sur la valorisation de bases patrimoniales par l'accès au contenu graphique sous une forme sémantique. La particularité de notre approche réside dans l'utilisation systématique d'une base de données comme mémoire du système. Cette mémoire numérique représente à la fois les connaissances apprises avec l'expert et la mémoire des cas déjà rencontrés. La gestion de connaissances est embarquée au travers de modèles génériques et de règles d'organisation spatiales. Une implantation est à l'étude sur la base d'un moteur de règles (type Prosper ou OntoDB). Le problème de l'intégration de la sémantique dans l'image n'est pas de pouvoir étiqueter des objets, ce qui peut s'exprimer comme un apprentissage particulier[4], mais de pouvoir disposer de modèles incrémentables associés à une base de cas qui intègre la variabilité des objets connus. Cette base de cas permet dès lors d'étendre, de façon supervisée, les connaissances liées aux modèles et d'aller plus loin dans la valorisation de la base. Ce que n'intègre pas encore notre travail, c'est une mesure de la validité des modèles face à la base de cas, pour éviter la dérive du système.

La question de la réduction du coût d'utilisation de ces approches n'a pas été développée ici, elle fait l'objet de travaux connexes. Plusieurs voies sont étudiées, en premier lieu celles qui visent à associer à chaque attribut une mesure de précision permettant de quantifier son importance relative et d'une mesure de pertinence dans la décision. En parallèle de cette approche, l'intégration d'extracteurs de points d'intérêts visuels doit permettre de déterminer un ensemble de points préliminaires d'études. L'enjeu est d'accroître la qualité de la décision tout en diminuant le nombre d'éléments à apparier, dans une phase de réduction de l'ensemble de recherche. Cette réduction initiale s'associe à une maîtrise des informations de contexte pour la sélection de modèles pertinents.

Références

- [1] A.K.C.Wong and M. You. Entropy and distance of randomgraphs with application to structural pattern recognition. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 7(5) :599–609, Septembre 1985.
- [2] D. Arrivault, N. Richard, and P. Bouyer. A fuzzy hierarchical attributed graph approach for handwritten hieroglyphs description. In A. Gagalowicz and W. Philips, editors, *11th International Conference on Computer Analysis of Images and Patterns (CAIP05)*, number LNCS 3691, pages 748–755, Rocquencourt, Septembre 2005. INRIA, Springer.
- [3] D. Arrivault, N. Richard, C. Fernandez-Maloigne, and P. Bouyer. Collaboration entre approches statistique et structurelle pour la reconnaissance de caractères anciens. In *8ème Colloque International Francophone sur l'Ecrit et le Document*, page a compléter, La Rochelle (France), Juin 2004.
- [4] K. Barnard, P.Duygulu, J.F.G. De Freitas, and D.A. Forsyth. Object recognition as machine translation : exploiting image database clustering models. pages 000–000, -, 2002. -.

- [5] I. Bloch, H. Maître, and M. Anvari. Fuzzy adjacency between image objects. *International Journal of Uncertainty, Fuzziness and knowledge-Based Systems*, 5(6) :615–653, 1997.
- [6] M. Borsotti, P. Campadelli, and R. Schettini. Quantitative evaluation of color image segmentation results. *Pattern Recognition Letters*, 19 :741–747, 1998.
- [7] L. Brun and W. Kropatsch. Defining regions within the combinatorial pyramid framework. In *Proceedings of the computer vision winter workshop*, pages 198–207, 2002.
- [8] K. Chan and Y. Cheung. Fuzzy-attribute graph with application to chinese character recognition. In *SMC*, volume 22, pages 402–410, 1992.
- [9] O. Coliot. *Représentation, évaluation et utilisation de relations spatiales pour l'interprétation d'images*. Phd, ENST-département TSI, Septembre 2003.
- [10] J. Dombre. *Systèmes de représentation multi-échelles pour l'indexation et la restauration d'archives médiévales couleur*. PhD thesis, Université de Poitiers, décembre 2004.
- [11] J. Dombre, N. Richard, and C. Fernandez-Maloigne. Content-based image retrieval and high level representations. In *IST/SPIE Internet Imaging III*, pages 134–143, San José, California, USA, January 2002.
- [12] J. Dombre, N. Richard, and C. Fernandez-Maloigne. Nouvelle utilisation de l'information spatiale pour l'indexation d'images basées contenu. *Annales des Télécommunications*, 57(9-10) :943–957, Septembre-Octobre 2002.
- [13] Steven Gold and Anand Rangarajan. A graduated assignment algorithm for graph matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(4) :377–388, 1996.
- [14] B. Ionescu, P. Lambert, D. Coquin, and L. Darlea. Color-based semantic characterization of cartoons. In *IEEE International Symposium on Signals, Circuits and Systems*, volume 1, pages 223–226, IASI(Roumanie), Juillet 2005.
- [15] R. Kikinis, M. Shenton, D. Iosifescu, R. McCarley, P. Saiviroonporn, H. Hokama, A. Robatino, D. Metcalf, C. Wible, C. Portas, R. Donnino, and F. Jolesz. A digital brain atlas for surgical planning model driven segmentation and teaching. In *IEEE Transactions on Visualization and Computer Graphics*, volume 2, pages 232–241, Septembre 1996.
- [16] G. Man and J. Poon. A fuzzy-attributed graph approach to handwritten character recognition. In *Second IEEE international Conference on Fuzzy Systems*, pages 570–575, San Francisco, Mars 1993.
- [17] A. Montanvert, P. Meer, and A. Rosenfeld. Hierarchical image analysis using irregular tessellations. In *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, volume 13, pages 307–316, Avril 1991.
- [18] H. S. Ranganath and L. J. Chipman. Fuzzy relaxation approach for inexact scene matching. In *International Conference on Image and Vision Computing (IVC)*, volume 10, pages 631–640, Septembre 1992.
- [19] S. Santini and R. Jain. Similarity measures. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(9) :871–883, 1999.
- [20] K. Shearer, H. Bunke, S. Venkatesh, and D. Kieronska. Efficient graph matching for vidéo indexing. *1st IAPR, int. Workshop on Graph based representations - Springer-verlag, computing suppl.*, 12 :53–62, 1998.
- [21] N. Zlatoff, B. Tellez, and A. Baskurt. Image understanding and scene models : a generic framework integrating domain knowledge and gestalt theory. In *IEEE International Conference on Image Processing (ICIP)*, pages 2355–2358, Singapour, Octobre 2004.

Mesurer la Similarité de Graphes

Sébastien Sorlin, Olfa Sammoud, Christine Solnon et Jean-Michel Jolion*

* LIRIS, CNRS UMR 5205, Lyon

{sebastien.sorlin,christine.solnon,jean-michel.jolion}@liris.cnrs.fr , olfa.sammoud@isg.rnu.tn

Résumé. Mesurer la similarité d'images est un problème qui se pose dans de nombreuses applications, que ce soit pour rechercher des images, les classifier ou encore les confronter à des modèles pour les "reconnaître". Quand les images sont représentées par des graphes, on peut mesurer la similarité de deux images en appariant les sommets des graphes associés aux images de façon à retrouver leurs caractéristiques communes. Différentes sortes d'appariements de graphes, impliquant différentes mesures de similarité, ont été proposés, *e.g.*, l'isomorphisme de (sous-)graphes pour mesurer l'équivalence ou l'inclusion de deux graphes, et la distance d'édition de graphes pour mesurer le coût de transformation d'un graphe en un autre. Ces différents appariements sont "univoques" dans le sens où un sommet d'un graphe ne peut être apparié qu'à au plus un sommet de l'autre graphe.

Nous proposons ici une nouvelle mesure de similarité de graphes basée sur un appariement multivoque des sommets des graphes, chaque sommet pouvant être apparié à un ensemble de sommets de l'autre graphe. Cet aspect nous permet notamment de comparer des graphes décrivant des images à différents niveaux de granularité, et de prendre en compte les problèmes de sur- et de sous-segmentation des images. Notre mesure de similarité est générique dans le sens où elle est paramétrée par des fonctions de similarité qui permettent d'exprimer des connaissances de similarité propres à l'application considérée. En particulier, nous montrons que toutes les mesures de similarité proposées jusqu'alors peuvent être vues comme un cas particulier de notre mesure. Nous nous intéressons également au problème du calcul de cette similarité, et nous proposons et comparons trois algorithmes : un algorithme glouton, un algorithme basé sur une recherche locale Taboué réactive et un algorithme basé sur l'optimisation par colonies de fourmis.

1 Introduction

Les graphes sont souvent utilisés pour représenter des images (Ambauen et al. (2003)) : les sommets représentent les régions de la scène et les arcs représentent les relations binaires entre les régions. Dans ce contexte, comparer des images consiste à comparer les graphes qui les représentent, *i.e.*, appairer leurs sommets afin d'identifier leurs points communs (Conte et al. (2004)). Cela peut se faire par la recherche d'un appariement exact de ces graphes (*i.e.*, une

relation d'isomorphisme de graphe ou de sous-graphe) afin de montrer que les deux graphes sont identiques ou qu'un graphe est inclus dans un autre. Cependant, les images sont souvent bruitées et présumer de l'existence d'une telle relation est une hypothèse trop forte. Par conséquent des appariements à tolérance d'erreurs (tels que la recherche du plus grand sous-graphe commun aux deux graphes ou le calcul de la distance d'édition de graphes) ont été proposés (Bunke (1997), Conte et al. (2004)). Le problème consiste alors à trouver un meilleur appariement, *i.e.*, celui qui préserve un maximum de sommets et d'arcs.

Plus récemment, trois différents articles ont proposé d'aller un cran plus loin dans la comparaison de graphes en introduisant la notion d'appariements multivoques, *i.e.*, des appariements où le sommet d'un graphe peut être apparié à un ensemble de sommets de l'autre graphe :

- Dans Champin et Solnon (2003), les graphes représentent les objets de conception d'une application de CAO. Dans ce contexte, en fonction de la granularité de représentation des objets, un composant d'un objet (*i.e.*, un sommet du graphe) peut jouer le même rôle qu'un ensemble de composants d'un autre objet. Par conséquent, les auteurs introduisent une mesure de similarité de graphes basée sur des appariements multivoques où un sommet d'un graphe peut être associé à plusieurs sommets d'un autre graphe.
- Dans Boeres et al. (2004), l'appariement de graphes est utilisé pour comparer des images médicales de cerveaux à un modèle d'image de cerveau. Dans ce contexte, l'image modèle a un aspect schématique facile à segmenter alors que l'image médicale, souvent bruitée, est généralement sur-segmentée. Par conséquent, une région de l'image modèle peut correspondre à plusieurs régions de l'image réelle et la comparaison de ces images est basée sur un appariement multivoque des deux graphes où un sommet du modèle peut être apparié à plusieurs sommets de l'image réelle. Dans Deruyver et al. (2006), la mise en correspondance de plusieurs sommets d'un graphe image à un même sommet d'un graphe modèle est utilisé pour une segmentation sémantique d'images.
- Ambauen et al. (2003) définit une nouvelle distance d'édition de graphes qui introduit deux nouvelles opérations d'édition –l'éclatement et la fusion de sommets– afin de prendre en compte le fait que les images peuvent être sur- ou sous-segmentées.

Dans Sorlin et Solnon (2005), nous avons montré que la mesure de Champin et Solnon (2003) est "plus générale" que toutes les autres mesures de similarité de graphes existantes : il est possible de la paramétrer afin de la rendre équivalente aux mesures existantes. Cependant, cette mesure est parfois difficile à paramétrer. Dans cette mesure, deux sommets (ou arcs) sont jugés d'autant plus similaires qu'ils partagent d'étiquettes communes. Or, dans de nombreuses applications (en particulier la reconnaissance d'images), certaines propriétés (telles que la taille ou la couleur d'une région) sont des nombres et ne peuvent pas être représentés par des étiquettes. En outre, dans Champin et Solnon (2003) la similarité de deux graphes dépend des étiquettes retrouvées par un appariement et non pas de l'appariement lui-même. Par conséquent, le calcul de certaines mesures de similarité (*e.g.*, Boeres et al. (2004)) est difficile avec cette mesure.

Objectifs et plan. Un premier objectif de ce papier (section 2) est de proposer une nouvelle mesure de similarité de graphes ayant le même pouvoir d'expression que la mesure de Champin et Solnon (2003) mais plus simples à paramétrer. En section 3 nous montrons que toutes les mesures de similarité de graphes existantes peuvent être vues comme des cas particuliers de notre mesure. Dans une seconde partie (section 4), nous nous intéressons au problème du

calcul de cette similarité. Nous proposons en particulier un algorithme de recherche locale Taboue réactive et un algorithme à base de colonie de fourmis et nous présentons quelques résultats expérimentaux.

2 Une nouvelle mesure de similarité de graphes

Appariement de graphes. Un graphe orienté G est défini par un couple $G = (V, E)$, où V est un ensemble fini de sommets et $E \subseteq V \times V$ est un ensemble d'arcs orientés. Pour comparer deux graphes, une première étape consiste à appairer leurs sommets afin d'identifier les points communs à ces deux graphes. Nous considérons ici des appariements multivoques, *i.e.*, chaque sommet d'un graphe peut être apparié à un ensemble –éventuellement vide– de sommets de l'autre graphe. Plus formellement, étant donnés deux graphes $G = (V, E)$ et $G' = (V', E')$, **un appariement** m de ces deux graphes est une relation entre V et V' (*i.e.*, $m \subseteq V \times V'$) telle que $(u, v) \in m$ ssi le sommet u de G est apparié au sommet v de G' . Nous notons $m(v)$ (resp. $m(u, v)$) l'ensemble des sommets (resp. arcs) appariés à v (resp. (u, v)) :

$$\begin{aligned} \forall v \in V, m(v) &= \{v' \in V' \mid (v, v') \in m\} \\ \forall v' \in V', m(v') &= \{v \in V \mid (v, v') \in m\} \end{aligned}$$

Et par conséquent :

$$\begin{aligned} \forall (u, v) \in E, m(u, v) &= \{(u', v') \in E' \mid u' \in m(u), v' \in m(v)\} \\ \forall (u', v') \in E', m(u', v') &= \{(u, v) \in E \mid u \in m(u'), v \in m(v')\} \end{aligned}$$

Fonctions de similarité de sommets et d'arcs. Pour chaque sommet v (resp. arc (u, v)) et chaque ensemble de sommets s_v (resp. d'arcs s_e) auxquels v (resp. (u, v)) peut être apparié, nous devons évaluer à quel point les propriétés de v (resp. de (u, v)) sont retrouvées. Plus formellement, nous supposons l'existence de deux fonctions sim_v et sim_e :

$$\begin{aligned} sim_v &: (V, \wp(V')) \cup (V', \wp(V)) \rightarrow]-\infty, 1] \\ sim_e &: (E, \wp(E')) \cup (E', \wp(E)) \rightarrow]-\infty, 1] \end{aligned}$$

Ces fonctions dépendent de l'application considérée et sont utilisées pour exprimer les connaissances de similarité et les contraintes que l'appariement recherché doit respecter. Ces fonctions retournent 1 quand toutes les propriétés du sommet ou de l'arc sont retrouvées (la similarité pour l'élément considéré est alors maximum) et $-\infty$ quand l'appariement effectué est incohérent (*i.e.*, viole des contraintes). Les valeurs intermédiaires sont utilisées lorsque seules quelques propriétés sont retrouvées et/ou qu'il est nécessaire de pénaliser l'appariement effectué.

Similarité de graphes. Etant donné un appariement m de deux graphes G et G' , la similarité de ces deux graphes par rapport à m est défini comme un ratio entre les propriétés retrouvées par l'appariement m et le nombre de sommets et d'arcs des deux graphes (afin de normaliser

Mesurer la similarité de graphes

la similarité de deux graphes entre $] -\infty, 1]$:

$$sim_m(G, G') = \frac{\sum_{v \in V \cup V'} sim_v(v, m(v)) + \sum_{(u,v) \in E \cup E'} sim_e((u,v), m(u,v))}{|V \cup V'| + |E \cup E'|}$$

$sim_m(G, G')$ définit la similarité de deux graphes par rapport à un appariement. La similarité entre deux graphes est définie comme la similarité induite par le meilleur appariement (*i.e.*, celui induisant la plus grande similarité) :

$$sim(G, G') = \max_{m \subseteq V \times V'} sim_m(G, G')$$

Notons que le mot "similarité" est ici utilisé dans son sens le plus commun : la similarité de deux graphes est élevée si les graphes se "ressemblent" et est égale à 1 (le maximum) quand les graphes sont "identiques". En effet, selon le choix des fonctions sim_v et sim_e , la similarité de graphes telle que nous l'avons définie peut ne respecter aucune propriété mathématique d'une fonction de similarité "classique". En général, notre similarité ne respecte pas l'inégalité triangulaire, elle peut être assymétrique, la similarité d'un graphe avec lui même peut être différente de 1 et la similarité de deux graphes peut être négative.

3 Comparaison avec d'autres mesures de similarité

En définissant correctement les fonctions sim_v et sim_e de similarité des sommets et des arcs, notre mesure de similarité permet la résolution de nombreux problèmes d'appariements de graphes. Nous présentons dans cette section quelques uns de ces problèmes et montrons brièvement comment il est possible de les résoudre grâce à notre mesure. Nous invitons le lecteur désirant plus de précisions à se référer à Sorlin et Solnon (2005) (la mesure utilisée est celle de Champin et Solnon (2003) mais nous montrons ici que notre nouvelle mesure est équivalente à celle de Champin et Solnon (2003)).

Isomorphisme de graphes. Le problème de l'isomorphisme entre deux graphes $G = (V, E)$ et $G' = (V', E')$ tels que $|V| = |V'|$ consiste à trouver une application bijective $\phi : V \rightarrow V'$ telle que $(u, v) \in E$ ssi $(\phi(u), \phi(v)) \in E'$. Cela revient à montrer que les deux graphes sont structurellement identiques.

Le problème consiste à trouver un appariement univalent (*i.e.*, chaque sommet est apparié à au plus un autre sommet) préservant tous les sommets et tous les arcs des deux graphes. Si la fonction sim_v (resp. sim_e) est telle qu'elle retourne 0 si le sommet (resp. l'arc) n'est pas apparié, 1 s'il est apparié à un seul sommet (resp. arc) et $-\infty$ s'il est apparié à plusieurs éléments (afin d'éviter les appariements multivoques), nous pouvons montrer que la similarité entre G et G' est égale à 1 ssi les graphes sont isomorphes.

Isomorphisme de sous-graphe partiel. Le problème de l'isomorphisme de sous-graphe partiel entre deux graphes $G = (V, E)$ et $G' = (V', E')$ tels que $|V| \leq |V'|$ consiste à trouver

une application injective $\phi : V \rightarrow V'$ telle que $(u, v) \in E \Rightarrow (\phi(u), \phi(v)) \in E'$. Cela revient à montrer que la structure du graphe G est incluse dans le graphe G' .

Le problème consiste à trouver un appariement univalent préservant tous les sommets et tous les arcs du graphe G . Si la fonction sim_v (resp. sim_e) se comporte comme pour l'isomorphisme de graphes pour les sommets (resp. les arcs) de G et qu'elle renvoie $-\infty$ si un sommet (resp. un arc) de G' est apparié à strictement plus d'un autre sommet (resp. arc) et 1 sinon, nous pouvons montrer que la similarité entre G et G' est égale à 1 ssi tous les éléments de G se retrouvent dans G' par un appariement qui définit une application injective des sommets de G dans les sommets de G' (i.e., ssi G est isomorphe à un sous-graphe partiel de G').

Nous avons montré dans Sorlin et Solnon (2005) qu'il est également possible de choisir les fonctions de similarité de façon à modéliser le problème de l'isomorphisme de sous-graphe non-partiel. Pour des raisons de place, nous ne détaillerons pas cette modélisation et nous invitons le lecteur à ce référer au papier original.

Plus grand sous-graphe partiel. Etant donnés deux graphes $G = (V, E)$ et $G' = (V', E')$, le problème du plus grand sous-graphe partiel commun consiste à trouver la taille du plus grand graphe G'' (en terme de nombre de sommets et d'arcs) qui soit un sous-graphe partiel de G et de G' (i.e., G'' est inclus dans G et dans G' à un renommage des sommets près). Le plus grand sous-graphe partiel commun à deux graphes peut être vu comme l'intersection de ces deux graphes et Bunke et Jiang (2000) définissent une mesure de similarité de graphes en faisant le ratio entre la taille de ce sous-graphe et la taille des graphes considérés. Deux graphes sont alors jugés d'autant plus similaires que leur intersection est grande (Lin (1998); Tversky (1977)).

Le problème consiste à trouver un appariement univalent qui préserve un nombre maximum de sommets et d'arcs des deux graphes. Si on choisit les mêmes fonctions sim_v et sim_e que pour le problème de l'isomorphisme de graphes, l'appariement qui maximise la similarité identifie le plus grand sous-graphe partiel commun aux deux graphes.

De la même façon, il est montré dans Sorlin et Solnon (2005) que le calcul de la similarité de deux graphes permet de résoudre le problème du plus grand sous-graphe non partiel commun à deux graphes et de calculer la distance d'édition de graphes Bunke (1997) entre deux graphes.

Appariements multivoques. Notre mesure de similarité est basée sur un appariement multivoque des graphes. Par conséquent, nous avons pu montrer qu'elle permettait le calcul des mesures de similarité proposées par Boeres et al. (2004), Ambauen et al. (2003) et Champin et Solnon (2003). Pour des raisons de place nous ne détaillons que la modélisation de la mesure de Champin et Solnon (2003) et nous renvoyons le lecteur à Sorlin et Solnon (2005) pour les autres mesures.

Champin et Solnon (2003) proposent une mesure de similarité de graphes multi-étiquetés (i.e., un ensemble d'étiquettes est associé à chaque sommet et chaque arc). Deux graphes sont jugés d'autant plus similaires que leur nombre d'étiquettes communes (par rapport à un appariement multivoque de leurs sommets) est grand. Cette mesure est générique car elle est paramétrée par une fonction f de pondération des étiquettes communes aux deux graphes et une fonction g pénalisant les éclatements de sommets. Notre nouvelle mesure est équivalente à cette mesure : en effet, la fonction sim_v (resp. sim_e) peut calculer l'intersection entre les éti-

quettes du sommet (resp. de l'arc) considéré et l'union des étiquettes des sommets (resp. arcs) auxquels il est apparié afin d'extraire l'ensemble des étiquettes communes aux deux graphes. Cet ensemble constitue l'argument de la fonction f de la mesure de Champin et Solnon (2003). Par conséquent, il est possible de reproduire le comportement de la fonction f grâce aux fonctions sim_v et sim_e . La fonction g de pondération des sommets éclatés peut également être calculée par la fonction sim_v . Notre nouvelle mesure de similarité peut donc être équivalente à la mesure de Champin et Solnon (2003) et les preuves d'équivalence avec les mesures de similarité de graphes de Boeres et al. (2004) et Ambauen et al. (2003) réalisées dans Sorlin et Solnon (2005) s'appliquent à notre nouvelle mesure. Notons cependant que la modélisation de ces mesures est beaucoup plus simple avec notre nouvelle mesure qu'avec celle de Champin et Solnon (2003).

4 Calcul de la mesure de similarité

Le calcul de notre mesure de similarité est un problème NP-difficile (il est plus général que la recherche du plus grand sous-graphe commun à deux graphes). L'explosion combinatoire rend les méthodes de recherches complètes (*i.e.*, basées sur une exploration exhaustive de l'espace de recherche et des méthodes de filtrage de cet espace) limitées à de très petits graphes (quelques sommets seulement). Par conséquent, nous proposons trois algorithmes incomplets (*i.e.*, ne garantissant pas l'optimalité de la solution trouvée mais ayant une complexité faiblement polynomiale) permettant le calcul de la similarité de deux graphes. Ces algorithmes, initialement proposés pour le calcul de la mesure de similarité de graphes de Champin et Solnon (2003) s'adaptent parfaitement à notre nouvelle mesure : selon le problème considéré, les graphes manipulés ne sont plus nécessairement étiquetés ce qui améliore considérablement le temps d'exécution de ces algorithmes.

Algorithme glouton. Dans Champin et Solnon (2003), un algorithme glouton de recherche du meilleur appariement de deux graphes est proposé. L'algorithme démarre d'un appariement vide (*i.e.*, aucun sommet n'est apparié). A chaque itération, un nouveau couple de sommets est ajouté à l'appariement courant. Le couple ajouté est choisi aléatoirement parmi ceux dont l'ajout maximise la similarité courante. Ce processus est itéré tant que l'ajout d'un couple à l'appariement améliore sa similarité. Cet algorithme a une complexité faiblement polynomiale et, comme il est non déterministe, il peut être exécuté plusieurs fois afin de garder le meilleur appariement trouvé.

Algorithme de recherche locale Taboue réactive. L'algorithme glouton retourne un appariement m "localement optimal" : ajouter ou enlever un couple de sommets à m ne peut l'améliorer. Cependant, il est parfois possible d'améliorer m en lui ajoutant et/ou en lui supprimant plusieurs couples de sommets. Une recherche locale cherche à améliorer une solution courante en explorant son voisinage : les voisins d'un appariement m sont les appariements obtenus en ajoutant ou en supprimant un seul couple de sommets à m . En démarrant d'un appariement initial, une recherche locale explore l'espace de recherche en se déplaçant de voisin en voisin jusqu'à l'obtention de la solution optimale ou l'utilisation du nombre maximum d'itérations autorisé. A chaque itération, le prochain voisin à explorer est choisi selon une heuristique.

La recherche Taboue [Glover (1989); Petrovic et al. (2002)] est une des meilleures heuristiques connues pour le choix du prochain voisin. Le voisin qui maximise la similarité est toujours sélectionné mais, afin de ne pas rester autour d'un maximum local en réalisant toujours les mêmes mouvements, une liste Taboue est utilisée. Cette liste mémorise les k derniers mouvements réalisés (*i.e.*, les k derniers couples de sommets ajoutés ou enlevés) afin d'interdire les mouvements inverses durant k itérations (*i.e.*, ajouter un couple de sommets précédemment supprimé ou supprimer un couple de sommets précédemment ajouté). Ce mécanisme permet de s'échapper des maxima locaux.

La longueur k de la liste taboue est un paramètre critique : si la liste est trop longue, l'algorithme converge trop lentement et si elle est trop courte, l'algorithme est incapable de quitter les maxima locaux et d'améliorer la solution. Pour résoudre ce problème de paramétrage, nous proposons dans Sorlin et Solnon (2005) d'utiliser la recherche locale taboue réactive de Battiti et Protasi (2001) pour le calcul de la similarité de deux graphes. La longueur de la liste est adaptée dynamiquement pendant la recherche en fonction du besoin courant de diversification et d'intensification de la recherche. Une table de hachage mémorisant la clé de hachage de tous les appariements explorés est utilisée. Quand un même appariement est exploré deux fois (*i.e.*, quand une collision survient dans la table de hachage), la recherche est diversifiée en allongeant la liste taboue. Si au contraire aucune collision ne survient pendant un nombre fixé d'itérations, la recherche est intensifiée en raccourcissant la liste. Plus de détails peuvent être trouvés dans Sorlin et Solnon (2005).

Algorithme d'optimisation par colonie de fourmis. Le principe des algorithmes d'optimisation à base de colonie de fourmis (ACO) [Dorigo et Caro (1999); Dorigo et Stützle (2004)] consiste à reformuler le problème à résoudre en un problème de recherche d'un chemin optimal dans un graphe (appelé graphe de construction) et à utiliser des fourmis artificielles pour trouver les bons chemins de ce graphe. A chaque itération de l'algorithme (appelée cycle), chaque fourmi de la colonie construit aléatoirement un chemin du graphe (*i.e.*, une solution du problème) et de la phéromone est déposée sur les meilleurs chemins découverts lors de ce cycle. Lors des cycles suivants, les fourmis construisent de nouveaux chemins avec une probabilité dépendant de la phéromone déposée lors des cycles précédents et d'une heuristique propre au problème considéré. La colonie de fourmis converge alors peu à peu vers les meilleurs solutions. Des mécanismes de contrôle de la quantité de phéromone déposée sur le graphe de construction déterminent l'efficacité d'un algorithme ACO : la quantité de phéromone déposée sur les composants des solutions peut être bornée afin de réguler l'effort de diversification de la recherche et une évaporation de la phéromone au cours du temps est mise en place afin "d'oublier" les mauvaises solutions construites.

Nous proposons dans Sammoud et al. (2005, 2006) un algorithme ACO pour le calcul de la similarité de deux graphes. Les sommets du graphe de construction sont tous les couples de sommets pouvant appartenir à un appariement. Chaque chemin du graphe construit par les fourmis est alors un appariement possible entre les graphes. Nous proposons une heuristique adaptée aux problèmes d'appariements de graphes et nous comparons deux stratégies phéromonales : une où la phéromone est déposée sur les sommets du graphe de construction (exprimant l'intérêt d'ajouter le couple de sommets correspondant à un appariement) et une autre où la phéromone est déposée sur les arcs (exprimant alors l'intérêt d'ajouter un couple de sommets à un appariement sachant qu'un autre couple appartient déjà à cet appariement).

Notre algorithme ACO est également couplé à une recherche locale qui améliore les solutions construites à chaque cycle et permettant une convergence plus rapide de l'algorithme. Nous invitons le lecteur voulant approfondir la question à lire Sammoud et al. (2005, 2006).

Résultats expérimentaux. Dans Sorlin et Solnon (2005); Sammoud et al. (2005, 2006) nous testons et comparons nos algorithmes sur des problèmes d'isomorphisme de graphes et de sous-graphes Foggia et al. (2001), sur les problèmes d'appariements non-bijectifs de graphes de Boeres et al. (2004) ainsi que sur des problèmes de mesures de similarité de graphes basées sur un appariement multivoque des sommets Champin et Solnon (2003). Les résultats montrent que, sur les problèmes d'appariements de graphes les plus simples (*i.e.*, les problèmes d'isomorphisme de (sous-)graphes), nos algorithmes peuvent être utilisés mais ne sont pas compétitifs avec des méthodes dédiées (e.g. Nauty McKay (1981) ou VFLIB Cordella et al. (2001)). Sur les problèmes plus difficiles (notamment ceux basés sur des appariements multivoques), nos algorithmes sont très compétitifs et permettent de résoudre des problèmes ayant plus d'une centaine de sommets (contre une petite dizaine pour les approches complètes existantes). L'algorithme glouton, très rapide, est le moins performant de nos trois algorithmes : il ne permet généralement pas de trouver la solution optimale des instances difficiles. L'algorithme Tabou réactif permet en général de trouver les solutions optimales des instances difficiles. Enfin, notre algorithme ACO est globalement meilleur que Tabou sur les instances très difficiles mais nécessite en général un temps d'exécution plus long que l'algorithme Tabou.

5 Conclusion

Nous proposons une nouvelle mesure de similarité de graphes : basée sur un appariement multivoque des sommets des graphes, elle permet la prise en compte des problèmes de sur- et sous-segmentation en reconnaissance d'images. Notre mesure est paramétrée par des fonctions de similarité des sommets et des arcs choisies en fonction de l'application considérée. Ces fonctions rendent notre mesure générique dans le sens où elle est équivalente à la mesure de Champin et Solnon (2003) (montrée générique dans Sorlin et Solnon (2005)) et que toutes les autres mesures de similarité de graphes existantes peuvent être vues comme un cas particulier de notre mesure. Nous proposons trois algorithmes de calcul de cette similarité et en particulier un algorithme de recherche locale Tabou réactive et un algorithme basé sur l'optimisation à base de colonie de fourmis meilleur que Tabou mais nécessitant un temps d'exécution plus long. Ces algorithmes permettent le calcul de la similarité de graphes de plus de 100 sommets alors que les méthodes complètes existantes ne permettent, dans le cas général, la comparaison que de très petits graphes.

Références

- Ambauen, R., S. Fischer, et H. Bunke (2003). Graph Edit Distance with Node Splitting and Merging, and Its Application to Diatom Identification. In *IAPR-TC15 Wksp on Graph-based Representation in Pattern Recognition, LNCS 2726*, pp. 95–106.
- Battiti, R. et M. Protasi (2001). Reactive local search for the maximum clique problem. In Springer-Verlag (Ed.), *Algorithmica*, Volume 29, pp. 610–637.

- Boeres, M., C. Ribeiro, et I. Bloch (2004). A randomized heuristic for scene recognition by graph matching. In *WEA 2004*, pp. 100–113.
- Bunke, H. (1997). On a relation between graph edit distance and maximum common subgraph. *PRL : Pattern Recognition Letters 18*, 689–697.
- Bunke, H. et X. Jiang (2000). *Graph Matching and Similarity*, Volume Teodorescu, H.-N., Mlynek, D., Kandel, A., Zimmermann, H.-J. (eds.) : Intelligent Systems and Interfaces, Chapter 1. Kluwer Academic Publishers.
- Champin, P.-A. et C. Solnon (2003). Measuring the similarity of labeled graphs. In *5th International Conference on Case-Based Reasoning (ICCBR 2003)*, pp. 80–95. Lecture Notes in Artificial Intelligence 2689-Springer-Verlag.
- Conte, D., P. Foggia, C. Sansone, et M. Vento (2004). Thirty years of graph matching in pattern recognition. *International Journal of Pattern Recognition and Artificial Intelligence 18(3)*, 265–298.
- Cordella, L., P. Foggia, et M. V. C. Sansone (2001). An improved algorithm for matching large graphs. *3rd IAPR-TC15 Workshop on Graph-based Representations in Pattern Recognition*.
- Deruyver, A., Y. Hodé, et J.-M. Jolion (2006). Pyramide adaptative et graphe sémantique : Un processus de segmentation dirigé par la connaissance. In *RFIA '06*, pp. To Appear.
- Dorigo, M. et G. D. Caro (1999). The ant colony optimization meta-heuristic. In *New Ideas in Optimization*, pp. 11–32. McGraw Hill, London, UK.
- Dorigo, M. et T. Stützle (2004). *Ant Colony Optimization*. MIT Press.
- Foggia, P., C. Sansone, et M. Vento (2001). A database of graphs for isomorphism and subgraph isomorphism benchmarking. *3rd IAPR-TC15 Workshop on Graph-based Representations in Pattern Recognition*, 176–187.
- Glover, F. (1989). Tabu search - part I. *Journal on Computing*, 190–260.
- Lin, D. (1998). An Information-Theoretic Definition of Similarity. In *proc. of ICML 1998, 15th Inter. Conf. on Machine Learning*, pp. 296–304. M. Kaufmann.
- McKay, B. (1981). Practical graph isomorphism. *Congressus Numerantium*.
- Petrovic, S., G. Kendall, et Y. Yang (2002). A Tabu Search Approach for Graph-Structured Case Retrieval. In *STAIRS 2002*, pp. 55–64.
- Sammoud, O., C. Solnon, et K. Ghédira (2005). An ant algorithm for the graph matching problem. *5th European Conference on Evolutionary Computation in Combinatorial Optimization (EvoCOP 2005) Lecture Notes in Computer Science 3448 - Springer Verlag*, 213–223.
- Sammoud, O., S. Sorlin, C. Solnon, et K. Ghédira (2006). A comparative study of ant colony optimization and reactive search for graph matching problem. *6th European Conference on Evolutionary Computation in Combinatorial Optimization (EvoCOP 2006) Lecture Notes in Computer Science - Springer Verlag*, To appear.
- Sorlin, S. et C. Solnon (2005). Reactive tabu search for measuring graph similarity. In L. Brun et M. Vento (Eds.), *5th IAPR-TC-15 workshop on Graph-based Representation in Pattern Recognition*, pp. 172–182. Springer Verlag.
- Tversky, A. (1977). Features of Similarity. In *Psychological Review*, Volume 84, pp. 327–352. American Psychological Association Inc.

Summary

Many applications involve matching two graphs in order to identify their common features and compute their similarity. Different kinds of graph matchings, giving rise to different graph similarity measures, have been proposed. A first goal of this paper is to propose a new graph similarity measure based on a multivalent matching of the graph vertices. We show that this measure is generic in the sense where it can be used to compute all other known graph similarity measures. In a second part, we address the problem of computing this measure and we describe three algorithms: a greedy algorithm, that quickly computes sub-optimal solutions, a reactive Tabu search algorithm, that may improve these solutions and an ant colony optimisation algorithm. Some experimental results are given.

Comment utiliser le treillis de Galois en reconnaissance d'images?

Stéphanie Guillas, Karell Bertet et Jean-Marc Ogier

Laboratoire L3I, Université de La Rochelle, av M. Crépeau, 17042 La Rochelle Cedex 1, France
{sguillas, kbertet, jmogier}@univ-lr.fr

Résumé

Le treillis de Galois ou treillis des concepts a été utilisé à de nombreuses reprises comme outil de classification, notamment dans le domaine de la fouille de données. L'objectif de ce papier est de proposer une description générique de son utilisation en classification, afin de mettre en évidence le fait que le treillis de Galois est plutôt approprié à des données détériorées. Nous ciblons ensuite les points clé de cette approche pour permettre une adaptation du treillis de Galois dans le cas spécifique où les données sont des images de symboles bruités.

Mots Clef

Classification, treillis de Galois, extraction de connaissances, apprentissage.

1 Introduction

Nous proposons une première réflexion concernant l'exploitation des treillis de Galois dans un processus de reconnaissance des formes. Une étude récente [MNN05] propose une comparaison de différentes méthodes de classification supervisée basées sur un treillis de Galois, où les expérimentations présentées montrent clairement que le treillis de Galois offre un cadre intéressant en classification, malgré une complexité théorique exponentielle dans le pire des cas. En effet, les treillis nous semblent appropriés, de par leurs facultés de validation de primitives pertinentes au regard d'un contexte particulier. Des pistes de réflexion sont proposées et font l'objet d'expérimentations menées au sein de notre laboratoire, dans le cadre d'une problématique de reconnaissance de symboles. La reconnaissance (figure 1) est classiquement réalisée en deux étapes : l'apprentissage qui fait l'objet de la partie suivante, et la classification qui sera présentée par la suite.

2 Apprentissage

La phase d'apprentissage consiste à organiser l'information extraite d'un ensemble d'objets sous forme d'un treillis de Galois. Nous nous plaçons dans le cas où les objets sont des images de symboles décrites par des vecteurs numériques (ou signatures) de même taille extraits à partir des images. Il est nécessaire d'avoir des données pré-

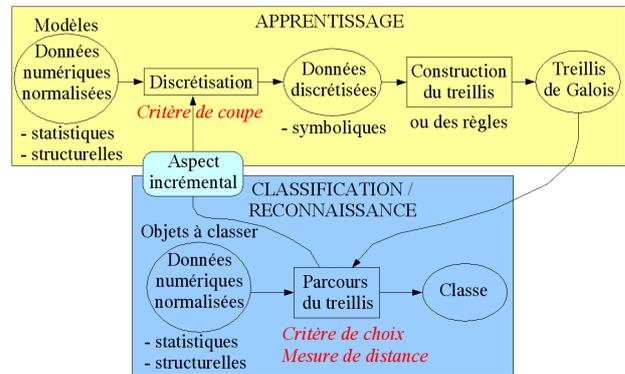


FIG. 1 – Description de la reconnaissance des formes

ablement normalisées pour que leur représentation soit équivalente. Cette phase d'apprentissage se décompose en 2 étapes comme l'illustre la figure 1 :

- une étape de *discrétisation* des données numériques : où les données sont réparties dans des intervalles disjoints. La discrétisation est essentielle à la construction du treillis de Galois. Un paramètre de *critère de coupe* est nécessaire à la construction des intervalles.
- une étape de *construction du treillis* à partir des données discrétisées ne nécessitant aucun paramétrage.

2.1 Discrétisation

La discrétisation consiste à organiser les données numériques $p = (p_i)_{i \leq n, p \in O}$ représentant différents objets en intervalles discrets afin d'obtenir une caractéristique spécifique de chaque classe d'objets.

La discrétisation est réalisée à partir des signatures que l'on peut organiser sous forme d'une table de données à double entrée (fig. 2). Au début, on construit pour chaque caractéristique i un intervalle $x = I_i$ regroupant l'ensemble V_x des valeurs p_i des objets $p \in O$. On peut ainsi initialiser la relation d'appartenance R qui s'en déduit. Notons qu'après cette étape d'initialisation, chaque objet $p \in O$ est en relation selon R avec l'unique intervalle $x = I_i$ pour chaque caractéristique i .

Il s'agit ensuite de sélectionner un *intervalle* $x \in I$ à couper, et un *point de coupe* dans cet intervalle $v_j \in V_x$ parmi les n valeurs $V_x = v_1, \dots, v_n$ triées par ordre croissant de l'intervalle, puis de couper l'intervalle x en deux intervalles x' et x'' avec $V_{x'} = v_1 \leq \dots \leq v_j$

et $V_{x''} = v_{j+1} \leq \dots \leq v_n$. Chaque objet sera alors en relation d'appartenance R avec l'un des deux intervalles créés, ce qui permet de différencier les deux sous-ensembles d'objets formés. On réitère le procédé de découpage des intervalles tant que l'on n'est pas capable de distinguer chacune des classes. La sélection de l'intervalle à couper est dépendante d'un *critère de coupe* à définir.

Lorsque chaque classe est caractérisable par un ensemble d'intervalles qui lui est propre, on obtient une table discrétisée (fig. 3) contenant l'ensemble des objets $p \in O$ et des intervalles $I = I_1 \times I_2 \times \dots \times I_m$ avec I_i les intervalles obtenus pour chaque caractéristique $i = 1 \dots m$. Toute caractéristique k qui n'a jamais été sélectionnée pour être discrétisée, contient un seul intervalle ($|I_k| = 1$) en relation avec tous les objets, elle n'est donc pas discriminante, et peut être retirée de la table discrétisée. Cette table permet de déduire la relation d'appartenance R , et par conséquent, pour un objet $p = (p_1, p_2, \dots, p_m) \in O$ où p_i est la valeur pour la caractéristique $i = 1 \dots m$, de connaître l'ensemble I_p des intervalles associés à p .

Exemple 1 Prenons l'exemple de la table 2. Elle présente les données normalisées de 10 objets répartis suivant 4 classes. La signature caractérisant chaque objet est composée de 3 caractéristiques (a , b et c). Après discrétisation par un critère de coupe défini à partir d'un calcul d'entropie, on obtient la table de la figure 3. Chacune des caractéristiques a été sélectionnée et découpé une seule fois, elles sont donc toutes conservées.

Classe	Ident.	Signature		
		a [0-20]	b [0-20]	c [0-20]
1	1	1	4	15
	2	0	0	18
2	3	1	12	13
	4	0	16	15
	5	3	12	11
3	6	8	16	15
	7	6	20	20
	8	15	12	15
4	9	18	4	0
	10	20	12	2

FIG. 2 – Signatures des 10 objets

2.1.1 Critère de coupe

De nombreux critères permettent de sélectionner l'intervalle à diviser et son point de coupe. Le choix de ce paramètre est déterminant pour l'apprentissage. Nous recherchons un intervalle $x \in I$, dont les valeurs $V_x = (v_1 \dots v_n)$ sont triées par ordre croissant, qui *maximise*

Classe	Ident.	Intervalles					
		$a1$ [0-3]	$a2$ [6-20]	$b1$ [0-4]	$b2$ [12-20]	$c1$ [0-2]	$c2$ [11-20]
1	1	X		X			X
	2	X		X			X
2	3	X			X		X
	4	X			X		X
	5	X			X		X
3	6		X		X		X
	7		X		X		X
	8		X		X		X
4	9		X	X		X	
	10		X		X	X	

FIG. 3 – Table discrétisée avec le critère de coupe d'entropie

un critère, pour une valeur v_j donnée. L'intervalle sera coupé entre v_j et v_{j+1} . Parmi les critères, citons la *distance maximale*, l'*entropie* et le *coefficient de Hotelling*.

Le *critère de distance maximale* choisit l'intervalle qui possède un écart maximal entre deux valeurs consécutives. C'est un critère non supervisé. Il est en effet possible que la coupe de l'intervalle soit faite entre deux objets de la même classe. Ce critère de discrétisation peut donc entraîner un très grand nombre de découpages. Le *critère d'entropie* permet quant à lui de séparer au mieux les classes, car il évalue le degré de mélange des classes (critère supervisé). Cependant il ne prend pas en compte l'écart entre les différentes classes. Enfin le *critère de Hotelling* choisit l'intervalle qui maximise la distance entre les différentes classes (i.e. la variance inter-classe) et minimise l'éparpillement au sein de chacune des classes (i.e. la variance intra-classe). Il prend en compte les classes ainsi que leur organisation, et notamment des distances qui les séparent.

Tout critère supervisé s'annule lorsque les classes sont correctement séparées. Ainsi, lorsque la table discrétisée sépare les classes, les critères d'entropie et de Hotelling sont nuls, alors que la distance maximale (critère non supervisé) peut encore être utilisée. En poursuivant la discrétisation par ce critère, la table obtenue contiendra plus d'intervalles et une description plus fine des classes.

Notons la possibilité d'intégrer des données symboliques aux données numériques. L'intégration de ces données consiste à calculer une extension de la relation d'appartenance R pour ajouter ces données au treillis. Cette intégration peut aussi être réalisée au cours de l'initialisation de la relation R , avant la discrétisation ; et ainsi servir à affiner le critère de coupe.

2.2 Construction du treillis de Galois

Après la phase de discrétisation vient la construction du treillis. Cette étape est totalement déterminée par la relation d'appartenance R obtenue. Aucun critère ou paramètre n'est à prendre en compte pour la construction du graphe étant donné qu'il représente toutes les combinaisons possibles de relations R entre objets et intervalles.

2.2.1 Description

Un treillis de Galois est composé d'un ensemble de *concepts* reliés par inclusion, et qui forment ainsi un graphe possédant les propriétés d'un treillis.

On associe à un ensemble d'objets $A \subseteq O$ l'ensemble $f(A)$ des intervalles en relation avec les objets de A :

$$f(A) = \bigcap_{p \in A} I_p = \{x \in I \mid pRx \forall p \in A\}$$

Duallement, pour un ensemble d'intervalles $B \subseteq I$, on définit l'ensemble $g(B)$ des objets en relation avec les intervalles de B :

$$g(B) = \{p \in O \mid pRx \forall x \in B\}$$

Les fonctions f et g définies entre objets et intervalles forment une *correspondance de Galois*.

Un *concept formel* est un couple objets-intervalles en relation selon R . Plus formellement, c'est un couple (A, B) avec $A \subseteq O$, $B \subseteq I$, $f(A) = B$ et $g(B) = A$.

Le concept minimal au sens de la relation R contient tous les objets O . Il s'agit du concept $(O, f(O))$. L'ensemble $f(O)$, généralement vide, correspond aux intervalles partagés par tous les objets. Duallement, le concept maximal est $(g(I), I)$.

Exemple 2 Sur la table de la figure 3, le concept (\emptyset, I) est le concept maximal du treillis de Galois car $g(I) = \emptyset$ (aucun objet n'est en relation avec tous les intervalles à la fois). A l'inverse, si on applique la fonction f sur l'ensemble O de tous les objets, on obtient un ensemble vide d'intervalles car nous avons supprimé les intervalles en relation avec l'ensemble des objets. Le concept (O, \emptyset) est donc le concept minimal du treillis.

Les concepts correspondant aux relations objets-intervalles sont ordonnées par inclusion pour former le treillis de Galois de la relation R . Sa représentation est unique. Il existe de nombreux algorithmes de génération du treillis de Galois : Bordat [Bor86], Ganter [Gan84], Godin et al. [GMA91] et Nourine et Raynaud [NR99] qui a la meilleure complexité théorique (complexité quadratique par élément du treillis produit).

La principale limite de l'utilisation du treillis de Galois est due à son coût à la fois en temps et en espace. En effet, la taille du treillis est bornée par $2^{|S|}$ dans le pire des cas, et par $|S|$ dans le meilleur des cas. Notons cependant que

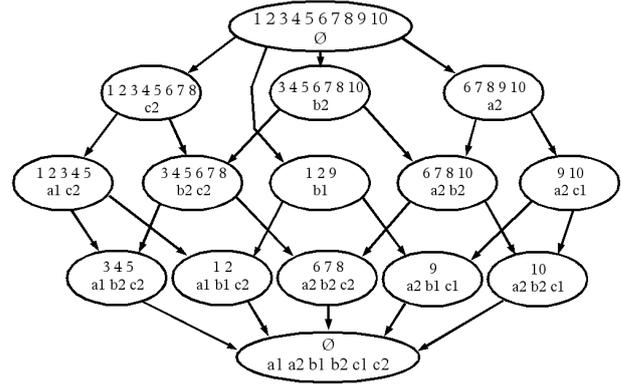


FIG. 4 – Treillis de Galois

sa taille reste raisonnable en pratique, comme l'illustrent les expérimentations qui en ont déjà été faites.

Afin de limiter cette complexité exponentielle, notons la possibilité de générer non pas le treillis lui-même, mais une *représentation* du treillis. Parmi les nombreuses représentations proposées dans la littérature, citons la représentation par un *système de règles d'implication* [BN04, TB01] que l'on retrouve en analyse de données. Une telle représentation permet, outre sa propriété de représentativité condensée du treillis, d'éviter la génération complète du treillis en proposant une *génération à la demande* des concepts du treillis qui sont nécessaires au cours de la phase de reconnaissance (i.e. de classification).

3 Classification

Lorsque la phase d'apprentissage est achevée et que le treillis de Galois est généré, il est possible de passer à l'étape de classification. Le principe est de déterminer la classe de nouveaux objets, c'est à dire de reconnaître la classe d'objets qui peuvent être plus ou moins détériorés comme l'illustre la figure 1.

3.1 Principe de la navigation

Le treillis de Galois peut être vu comme un espace de recherche dans lequel on évolue en fonction des caractéristiques validées. Il s'agit à partir du *concept minimal* $(O, f(O))$ où toutes les classes des objets sont *candidates* à la reconnaissance et aucun intervalle n'est validé, de progresser étape par étape au sein du treillis de Galois par validation de nouveaux intervalles et par conséquent réduction de l'ensemble d'objets, jusqu'à un *concept final* où les objets restants, qui sont en relation avec tous les intervalles validés durant le parcours du graphe, sont tous de la même classe (fig. 5). En pratique, la navigation est effectuée dans le diagramme de Hasse, c'est à dire la réduction réflexive et transitive du treillis de Galois.

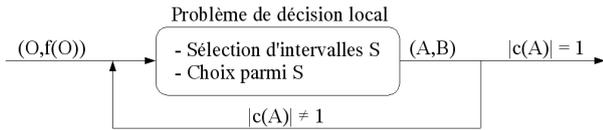


FIG. 5 – Navigation dans le treillis de Galois

3.1.1 Description d'une étape élémentaire de classification

Une étape élémentaire de classification consiste à partir d'un concept courant à sélectionner un ensemble d'intervalles S et à faire un choix parmi ces intervalles afin de progresser vers un nouveau concept courant. Plus précisément, S est une famille d'intervalles obtenue à partir des n successeurs $(A_1, B_1), \dots, (A_n, B_n)$ du concept courant (A, B) par :

$$S = \bigcup_{i=1}^n B_i \setminus B = \{X_1, \dots, X_n\}$$

S vérifie les propriétés suivantes :

- $X_i \cap X_j = \emptyset, \forall i, j \leq n, i \neq j$
- $|X_i \cap I_j| \leq 1, \forall i \leq n, \forall j \leq m$, c'est à dire que X_i ne peut contenir 2 intervalles issus d'une même caractéristique i .

Il s'agit ensuite de choisir un ensemble d'intervalles parmi X_i , et donc définir un *critère de choix* pour la sélection de ces intervalles. Plus précisément :

Choisir X_i parmi $S = \{X_1, \dots, X_n\}$.

Ce choix est essentiel à toute étape élémentaire de classification, base de la navigation dans le treillis et de la reconnaissance d'un symbole s . Ce choix est donc dépendant des données, et non de la structure. Plus précisément, il est déterminé par un *critère de choix* défini à partir d'une *mesure de distance* entre s et un intervalle.

Exemple 3 Dans l'exemple (fig. 4), le choix à partir du concept minimal doit être fait parmi $S = \{\{c_2\}, \{b_1\}, \{b_2\}, \{a_2\}\}$. Notons qu'il existe dans le treillis de Galois plusieurs chemins (scénarii de reconnaissance) permettant d'aboutir à une même classe. C'est cette propriété qui nous motive à utiliser le treillis de Galois dans le cadre d'une classification d'objets détériorés.

3.1.2 Mesure de distance

Faire un choix parmi S nécessite l'utilisation d'une *mesure de distance* entre la valeur s_i du symbole s à classifier et un intervalle $x \in I_i$. Dans notre expérimentation, nous avons utilisé le ratio entre la distance euclidienne de s_i au centre de l'intervalle et la demi-longueur de celui-ci :

$$d(s_i, x) = \frac{\sqrt{(s_i - x_{milieu})^2}}{\sqrt{(x_{borneInf} - x_{milieu})^2}}$$

Par abus de notation, notons $d(s, x)$ au lieu de $d(s_i, x)$ cette mesure de distance, afin de pouvoir l'étendre à un ensemble d'intervalles $X \subseteq I$:

$$d(X) = \frac{1}{|X|} \sum_{x \in X} d(s, x)$$

3.1.3 Critères de choix

Etant donnés $S = \{X_1, \dots, X_n\}$ la famille des intervalles sélectionnés, et $d(s, X)$ la mesure de distance entre l'objet s à classifier et un ensemble d'intervalles $X \subseteq I$, il s'agit de mettre en place un critère de choix pour sélectionner X_i parmi S . De nombreux critères de choix peuvent être définis, dont une liste exhaustive n'est pas envisageable. Voici quelques exemples simples :

- 1 choisir i tel que $d(s, X_i)$ est minimal.
- 2 choisir i tel que $|X_i \cap I_k| = |\{x \in X_i \cap I_k\}|$ est maximal, avec I_k l'ensemble des k premiers intervalles de S triés par ordre croissant selon la distance $d(s, x)$.
- 3 choisir i tel que $|\{x \in X_i \text{ tel que } d(s, x) < d_c\}|$ est maximal, avec d_c une constante.

Notons que le critère n°1, défini de manière globale sur chaque X_i , possède l'inconvénient de noyer le bruit. Le second critère suit le principe des k plus proches voisins [CH67], et le critère n°3 est un cas particulier du second. Ces critères étant locaux pour chaque $i = 1 \dots n$, on peut en imaginer des plus sophistiqués faisant apparaître :

- 4 une combinaison de ces différents critères de choix, et surtout en cas de choix multiples, c'est à dire si plusieurs X_i donnent le même résultat.
- 5 les intervalles de $\bigcup_{i=1}^n X_i$ d'un même attribut j : $\bigcup_{i=1}^n X_i \cap I_j$ et $d(X_i \cap I_j)$ de manière à favoriser un attribut j , par exemple supposé non détérioré.
- 6 la notion de concepts discriminants (lorsque l'union des intervalles ne se retrouve que dans le concept maximal) et non discriminants (l'union des intervalles se retrouve dans un concept différent du concept maximal) pour favoriser les concepts non discriminants. Le choix dans ce cas n'exclurait pas le parcours de toute une partie du treillis de Galois et la prise de risque serait faible. A l'inverse, une décision entre des concepts discriminants est délicate car exclusive.
- 7 le taux d'intervalles de X_i inférés par rapport à l'ensemble $|X_i|$, c'est à dire le pourcentage d'intervalles validés par le choix de i alors qu'ils ne contiennent pas les valeurs de l'objet à reconnaître.

Enfin, nous envisageons d'évaluer la prise de risque à chaque étape élémentaire de classification par l'utilisation de la théorie du flou. Ainsi on aurait une estimation du critère de choix qui permettrait d'arrêter la progression dans le graphe lors d'une prise de décision comportant un risque d'erreur important. Ce serait une autre manière de repérer une prise de décision délicate, qui nécessite une meilleure signature de l'objet à reconnaître.

4 Résultats expérimentaux

Nous avons mis en évidence le fait que le treillis de Galois est approprié dans le cas d'une classification d'objets détériorés. En effet, le treillis de Galois propose plusieurs chemins pour reconnaître la même classe, qui sont autant de scénarii de classification correcte des objets bruités.

Cette étude expérimentale compare le treillis de Galois à un autre graphe mieux connu : l'arbre de décision. Nos données sont constituées des symboles de GREC 2003 [GRE03], où chaque classe de symbole, comporte 90 symboles détériorés par la méthode de Kanungo et al. [Ka94]. Nous avons réalisé l'ensemble des tests présentés dans cette expérimentation pour 2 groupes de 10 classes de symboles chacun, soit 2 ensembles de 900 symboles.

A partir des images de symboles nous avons extraits 3 signatures différentes : une signature de 33 valeurs numériques calculées avec la méthode de Fourier-Mellin [AOC⁺00], une signature de 50 valeurs basée sur la transformée de Radon [TW03] et une signature de 24 invariants de Zernike [Tea03]. A partir de ces 3 signatures, nous avons formé 3 combinaisons en les concaténant deux à deux : FMZern (Fourier-Mellin et Zernike), FMRad (Fourier-Mellin et Radon) et RadZern (Radon et Zernike).

Le tableau de la figure 6 présente une comparaison des taux de reconnaissance de l'arbre de décision et du treillis selon la signature utilisée. Nous avons utilisé le coefficient de Hotelling et le critère de choix n°4 défini par :

Choix : Choisir i tel que $|\{x \in X_i \text{ tel que } d(s, x) < 1\}|$ est maximal. Puis en cas de choix multiples, choisir i tel que $|\{x \in X_i \text{ tel que } d(s, x) < 1, 1\}|$ est maximal. Puis en cas de choix multiples, choisir i tel que $d(s, X_i)$ est minimal.

Les meilleurs résultats sont obtenus pour le treillis avec la signature de Radon. Les signatures combinées ne font pas mieux, cependant le second meilleur taux est obtenu avec la combinaison Radon/Zernike.

Signature	Reconnaissance (%)	
	Arbre de décision	Treillis de Galois
Fourier-Mellin (FM)	52.5	58.5
Radon (Rad)	61.5	75.5
Zernike (Zern)	56	60
FMZern	55.5	56
FMRad	52.5	58.5
RadZern	63	67.5

FIG. 6 – Comparaison des signatures

Pour comparer les taux de reconnaissance obtenus selon le critère de coupe (fig. 7), nous avons utilisé la signature

de Radon et le critère de choix n°4. Pour le treillis, la distance maximale et le coefficient de Hotelling donnent les meilleurs taux. Cependant, dans le tableau de la figure 8, on constate que le coefficient de Hotelling offre une structure (nombre de concepts du treillis) plus petite que celle obtenue avec la distance maximale. Ce critère supervisé est donc le meilleur compromis entre une faible taille de structure et un bon taux de reconnaissance.

Critère de coupe	Reconnaissance (%)	
	Arbre de décision	Treillis de Galois
Distance maximale	63.5	76.5
Entropie	51	61.5
Hotelling	61.5	75.5

FIG. 7 – Comparaison des critères de coupe

Critère de coupe	Taille des structures			
	Nombre d'étapes de discrétisation	Nombre d'intervalles discrétisés	Nombre de noeuds de l'arbre	Nombre de concepts du treillis
Distance maximale	56	96	18	127
Entropie	9	18	19	76
Hotelling	9	18	19	80

FIG. 8 – Taille des structures obtenues selon le critère de coupe utilisé pour la discrétisation

Nous avons ensuite comparé plusieurs critères de choix en utilisant la signature de Radon et le coefficient de Hotelling. On observe dans le tableau de la figure 9 que le treillis de Galois obtient les meilleurs taux de reconnaissance pour le critère n°4. Cependant, les résultats obtenus par les différents critères restent relativement proches.

Critère de choix	Reconnaissance (%)	
	Arbre de décision	Treillis de Galois
n°1	61.5	73
n°2	61.5	72
n°4	61.5	75.5

FIG. 9 – Comparaison des critères de choix

5 Conclusion et perspectives

Dans ce papier, nous décrivons génériquement l'utilisation du treillis de Galois en classification pour cibler et mettre en avant l'ensemble des points essentiels de cette

approche. Cette étape nous permet de nous focaliser sur les différents éléments ciblés pour une adaptation au cas spécifique de la reconnaissance de symboles détériorés. La figure 1 récapitule l'ensemble des étapes du traitement, les éléments paramétrables et les extensions possibles.

La discrétisation est essentielle car elle détermine l'organisation des données pour la construction du treillis de Galois. En effectuant un *découpage* des intervalles au-delà de l'étape permettant de caractériser chacune des classes, il est possible d'obtenir une description plus complète des classes. De plus, nous avons mis en évidence la possibilité d'intégrer des *données symboliques* au cours de l'apprentissage. Une telle extension peut être réalisée avant (pour affiner le critère de coupe) ou après la discrétisation.

Pour limiter la complexité exponentielle du treillis de Galois, il est possible de mettre en place un système de *règles d'implication* qui propose une représentation réduite, compréhensible et unique du graphe. De plus, ce système de règles permet de générer les concepts *à la demande*, c'est à dire de construire uniquement les concepts nécessaires à la phase de reconnaissance. Signalons également la possibilité d'évolution du treillis de Galois (ou des règles) sans remettre en cause le dispositif (aspect *incrémental*), aussi bien pour l'ajout de nouvelles classes que pour la reconnaissance trop risquée d'un objet qui nécessite une description plus précise. Cette prise de risque peut être évaluée soit par la notion d'intervalles discriminants, soit par l'utilisation de la théorie du flou.

Nous pensons instaurer un système de reconnaissance *itératif*, où la première itération serait un parcours du treillis jusqu'à un concept contenant plusieurs classes candidates et permettrait d'en éliminer certaines. Alors, un nouveau treillis serait construit à partir de cet ensemble de classes candidates. Cette approche paraît intéressante pour traiter de grands ensembles de classes, ce qui actuellement est assez coûteux et les premiers résultats semblent prometteurs. Nous envisageons aussi d'exploiter et de confronter les résultats obtenus avec plusieurs signatures ou plusieurs parties d'une même signature.

Nous envisageons également d'essayer de construire une structure que l'on pourrait situer entre l'arbre de décision et le treillis de Galois, c'est à dire une sorte de treillis simplifié où seuls certains concepts seraient conservés.

Enfin, nous prévoyons de fusionner plusieurs tables discrétisées pour construire un treillis de Galois comportant plus d'intervalles décrivant chacune des classes.

Références

[AOC⁺00] S. Adam, J.M. Ogier, C. Cariou, R. Mullot, J. Gardes, and Y. Lecourtier. Fourier-mellin based invariants for the recognition of multi-oriented and multi-scaled shapes : Application to engineering drawings analysis, in invariants for pattern recognition and classifi-

cation. *Series in Machine Perception and Artificial Intelligence*, 42 :132–147, 2000.

- [BN04] K. Bertet and M. Nebut. Efficient algorithms on the family associated to an implicational system. *DMTCS*, 6(2) :315–338, 2004.
- [Bor86] J.P. Bordat. Calcul pratique du treillis de Galois d'une correspondance. *Math. Sci. Hum.*, 96 :31–47, 1986.
- [CH67] T.M. Cover and P.E. Hart. Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, 13(1) :21–27, 1967.
- [Gan84] B. Ganter. Two basic algorithms in concept analysis. *Technische Hochschule Darmstadt (Preprint 831)*, 1984.
- [GMA91] R. Godin, R. Missaoui, and H. Alaoui. Learning algorithms using a Galois lattice structure. *Third International Conference on Tools for Artificial Intelligence, San Jose, Calif.*, pages 22–29, 1991.
- [GRE03] GREC. www.cvc.uab.es/grec2003/symrec-contest/index.htm, 2003.
- [Ka94] T. Kanungo and al. Document degradation models : parameter estimation and model validation. In *IAPR Workshop on machine vision applications, Kawasaki (Japan)*, pages 552–557, 1994.
- [MNN05] E. Mephu NGuifo and P. Njiwoua. Treillis des concepts et classification supervisée. In *Technique et Science Informatiques (à paraître), RSTI*. Hermès - Lavoisier, Paris, France, 2005.
- [NR99] L. Nourine and O. Raynaud. A fast algorithm for building lattices. In *Third International Conference on Orders, Algorithms and Applications*, Montpellier, France, august 1999.
- [TB01] R. Taouil and Y. Bastide. Computing proper implications. *Proceedings of ICCS-2001 International Workshop on Concept Lattices-Based Theory, Methods and tools for Knowledge Discovery in Databases*, pages 290–303, 2001.
- [Tea03] M. Teague. Image analysis via the general theory of moments. *Journal of Optical Society of America (JOSA)*, 70 :920–930, 2003.
- [TW03] S. Tabbone and L. Wendling. Recherche d'images par le contenu à l'aide de la transformée de radon. *Technique et Science Informatiques*, 2003.