

Description d'un objet graphique: proposition d'un modèle opérant dans l'espace image ou dans l'espace cosinus discret

A new model for graphical object description operating in the image space or in the Cosine Discrete space

Patrick Franco, Jean-Marc Ogier, Pierre Loonis, Rémy Mullot

Laboratoire Informatique, Image, Interaction (L3I),
Université de la Rochelle – Pôle Sciences et Technologies,
17042 La Rochelle Cedex 1 – France
patrick.franco@univ-lr.fr

Manuscrit reçu le 9 juin 2006

Résumé et mots clés

Cet article propose un nouveau modèle de description d'un objet dans une image. Ce modèle s'appuie sur la construction d'un arbre minimal, ses propriétés sont étudiées à travers le problème de la reconnaissance de symboles complexes. L'invariance de la reconnaissance – face aux translations et rotations de symboles dégradés – est vérifiée dans un contexte d'images binaires à faible résolution. Si les résultats sont concluants, le coût algorithmique peut être assez élevé. Une alternative consiste à exprimer l'objet cible dans l'espace Cosinus Discret (Transformation en Cosinus Discrète). La technique opère non plus dans l'espace image mais dans un espace compact où les données sont mieux décorrélatées. Certains de nos choix font référence à des concepts de compression d'images. Cette piste conduit à une diminution sensible du coût tout en conservant un niveau de discrimination significatif. Ces résultats sont d'abord observés lors d'une expérience élémentaire puis confirmés par un test à moyenne échelle, mettant en jeu 500 symboles issus de la base de données Graphics Recognition – GREC2003.

Analyse de documents, Reconnaissance de symboles, Extraction de caractéristiques, Arbres minimaux, Transformation en Cosinus Discrète, Compression d'Image.

Abstract and key words

In this article a new shape descriptor – based on minimal graphs – is proposed and its properties are checked through the problem of graphical symbols recognition. Recognition invariance in front shift and multi-oriented noisy object was studied in the context of small and low resolution binary images. The approach seems to have many interesting properties, even if the construction of graphs induces an expensive algorithmic cost. In order to reduce time computing an alternatively solution based on image compression concepts is provided. The recognition is realized in a compact space, namely the Cosine Discrete space. The use of blocks discrete cosine transform is discussed and justified. The experimental results led on the GREC2003 database show that the proposed method is characterized by a good discrimination power, a real robustness to noise with an acceptable time computing.

Document analysis, Graphical symbol recognition, Feature extraction, Minimal graphs, Discrete Cosine Transform, Image Compression.

Remerciements

Les auteurs remercient spécialement les Professeurs Olivier Michel (Université de Nice Sophia-Antipolis, Laboratoire d'Astrophysique) et Alfred Hero (University of Michigan, E.E.C.S. Department) pour le prêt gracieux du code de calcul des arbres.

1. Introduction

L'objectif de l'interprétation des documents techniques est de faciliter leur exploitation en proposant un ensemble d'étapes visant la transformation du document papier en un stockage numérique « intelligent » de l'information [EBI 94, MAR 97, OGI 95]. Les approches issues de l'interprétation, en soi un domaine de recherche vaste, semblent d'ailleurs constituer des alternatives intéressantes au besoin croissant d'indexation de documents (navigation sur un internet « sémantique »). De nombreux travaux ont été publiés dans le domaine des documents cartographiques, des plans cadastraux [AIR 95, HUA 99], ou des schémas mécaniques et électriques [MER 97]. Cette problématique soulève de nombreuses questions sur les différents niveaux d'une chaîne traditionnelle de traitement. Parmi les sujets difficiles n'ayant toujours pas trouvé de réponse générale et opérationnelle, celui de la reconnaissance de symboles est particulièrement délicat, compte tenu de la variabilité des représentations des symboliques et compte tenu du paradigme classique « segmentation/reconnaissance ». Sur un plan scientifique, la communauté travaillant sur ces sujets rencontre donc encore un certain nombre de verrous liés à la multi-représentation de symboles, aux contraintes de multi-orientation et multi-échelle, et à la très grande densité de certains documents, rendant très difficile la caractérisation des objets à reconnaître. L'intérêt scientifique de ce sujet est d'ailleurs attesté au plan national et international par de nombreux programmes de recherche en relation avec l'évaluation de performances.

1.1. Un panorama des techniques de caractérisation/reconnaissance de symboles

En matière de reconnaissance de symboles sur des documents graphiques [LLA 01], il est possible de distinguer deux grandes catégories d'approches, suivant que les formes à reconnaître sont originellement connectées ou non à d'autres objets sur l'image. D'une part, la littérature propose un ensemble très important d'approches fondées sur une description structurelle. D'autre part, on trouve de nombreuses contributions s'appuyant sur des signatures discrètes et/ou statistiques des formes, suivant des schémas intégrant des classifieurs plus ou moins élaborés. Nous proposons ci-après un panorama synthétique de ces différentes approches.

1.1.1. Approches à base de description structurelle

La représentation/reconnaissance structurelle est une voie fréquemment utilisée en reconnaissance de symboles, comme en attestent de nombreuses publications faisant référence dans le domaine [LLA 01, AHS 01, PAS 96]. De façon générale, ces approches s'appuient sur des étapes préalables de traitement et d'extraction de primitives sur les images, visant à construire progressivement une représentation structurelle des données, sous la forme de graphe ou de grammaire. Ce type de modélisation permet d'intégrer des relations topologiques entre les éléments de bas niveau composant les objets.

En matière de reconnaissance, les approches qualifiées de « structurelles » utilisent principalement deux grandes familles de méthodes : les méthodes reposant sur l'appariement de graphes et celles fondées sur des grammaires de graphes. De nombreux travaux mettent en œuvre ces approches [LEE 92, YU 94, BUN00, LLA 04]. Par ailleurs, certains combinent ces approches structurelles avec des mécanismes probabilistes pour les phases d'apprentissage [LIN 85]. D'autres encore procèdent à une pré-classification globale du symbole à l'aide de mesures statistiques, puis la valident par des méthodes structurelles [BHA 94]. Citons encore des approches hybrides [LEE 92, DEL 94] qui suggèrent l'identification des éléments formant la structure du symbole au moyen d'approches statistiques, puis la combinaison de ces symboles au travers d'approches structurelles.

1.1.2. Approches à base de description statistique

Les approches fondées sur une description statistique s'appuient sur un calcul d'une signature de la forme à reconnaître. Lorsque les formes sont connexes, une étape préalable de séparation est nécessaire : cela soulève de nombreuses interrogations sur l'influence des paramètres des opérateurs de segmentation sur la décision finale. Par ailleurs, il est nécessaire de mettre en place des méthodes invariantes aux similitudes pour apporter de la fiabilité au système global. Nous distinguerons [ADA 01] trois principales catégories. Une première suggère un calcul préliminaire de l'orientation de la forme, et essaie, par l'intermédiaire d'étapes de normalisation et de rotation, d'obtenir une forme dans une position de référence, qui peut ainsi être introduite dans un système de classification. Une seconde approche consiste à utiliser un classifieur neuronal prenant en entrée l'image de la forme et rend lui-même, lors de sa phase d'apprentissage, le problème invariant aux transformations

[FUK 92]. Les problèmes posés par ce type de stratégie reposent sur la constitution de bases d'apprentissage représentatives de toute la variabilité du problème. La dernière approche, la plus classique, consiste à extraire un ensemble de descripteurs invariants aux transformations, avant d'alimenter un système de classification. Un bon état de l'art sur la description de formes peut être trouvé dans [TRI 96]. S. Adam [ADA 01] propose une catégorisation des vecteurs utilisés pour décrire les formes indépendamment de leur position, de leur taille et de leur orientation en deux catégories, correspondant respectivement aux descriptions globales et aux descriptions locales. Parmi les travaux récents concernant les descriptions globales, des approches originales à base de transformées de Radon, ou de F-signatures, ont été proposées par l'équipe Qgar du LORIA à Nancy [TAB 04, TAB 01a]. Concernant les approches locale, à l'inverse des approches présentées ci-dessus, une description invariante peut aussi être effectuée en utilisant des primitives qui, localement, sont théoriquement très informatives. Par exemple, les contours sont fréquemment utilisés pour obtenir des descriptions invariantes des formes par utilisation des descripteurs de Fourier [PEI 92] ou des descripteurs elliptiques de Fourier [LIN 87]. T. Taxt [TAX 90] a mené une étude comparative entre ces descripteurs qui montre leur intérêt potentiel, surtout en ce qui concerne leur simplicité et leur robustesse.

1.2. Motivations

Dans le domaine de la reconnaissance de formes invariante à différentes transformations géométriques, la plupart des chercheurs [TRI 96] s'accordent pour dire que l'étape d'extraction de primitives est probablement le facteur le plus important. La littérature décrit un grand nombre de primitives issues est rapporté d'approches structurelles ou statistiques. Étant donné le grand nombre de méthodes existantes, on pourrait prétendre qu'il n'est pas nécessaire de développer de nouveaux jeux d'invariants. Pourtant, l'étude comparative menée dans [ADA 01] montre que la sensibilité des primitives au bruit a des conséquences directes sur les taux de reconnaissance. De plus, les travaux de synthèse de K. Tombe [TOM 03] et J. Lladós [LLA 03] soulèvent implicitement la question fondamentale du paradigme segmentation/reconnaissance, non résolue dans le contexte des documents denses, bruités, avec de fortes interconnexions entre les objets graphiques à reconnaître. Ces constatations nous amènent à envisager l'exploitation de techniques permettant de coupler la segmentation et la reconnaissance, notre exploitation des arbres de représentation minimaux, intègre complètement cette stratégie.

2. Organisation de l'article

Dans cet article, nous proposons une nouvelle mesure captant la topologie d'un objet. Dans le système de coordonnées de l'image, tout objet est défini par un ensemble de points – d dimensionnel – ayant une topologie spécifique. La connexion de tous les points à l'aide d'un arbre de longueur minimale (Minimum Spanning Tree, MST) caractérise la topologie de l'objet cible. Or la longueur du MST ne peut être utilisée directement dans la comparaison de deux objets sans risque d'ambiguïté. La définition «du mélange» des objets à comparer lève cette ambiguïté et permet même de détecter les rotations éventuelles. Ce principe est abordé dans le paragraphe 3 et illustré à travers un exemple. La technique est appliquée à la reconnaissance de symboles bruités et multi-orientés dans le cas d'images binaires à faibles résolutions (*cf.* paragraphe 3.4). Les résultats semblent concluants mais l'approche n'est pas dénuée d'un certain coût algorithmique. Dans le paragraphe 4, une estimation de ce coût est réalisée et les pistes visant à sa réduction sont évoquées. L'alternative retenue consiste à exprimer l'objet cible dans l'espace Cosinus Discret (Transformation en Cosinus Discrète). La technique opère non plus dans l'espace image mais dans un espace compact où les données sont mieux décorréelées. Les arguments justifiant ce choix sont avancés, certains ne sont pas étrangers aux principes inhérents à la compression d'images (*cf.* Paragraphe 5). La pertinence de l'approche est d'abord vérifiée dans une expérience élémentaire où sont évalués le gain algorithmique et le niveau de discrimination inter-objets. Un test à plus grande échelle (500 symboles) est ensuite entrepris pour analyser le comportement de l'approche dans des conditions réalistes (*cf.* Paragraphe 5.3).

3. Reconnaissance de symboles par arbres minimaux construits sur le mélange

3.1. Définition

Soit $O = \{p_1, p_2, \dots, p_i, \dots, p_N\}$ l'ensemble de N points évoluant en \mathbb{R}^d ($d = 2$ dans le cas binaires, $d = 3$ pour des images en niveau de gris) définissant l'objet dans le système de coordonnées de l'image. Un arbre recouvrant minimal (Minimum Spanning Tree : MST) construit sur $O = \{p_1, p_2, \dots, p_i, \dots, p_N\}$ est un arbre G :

- *totalemment connecté* : tous les points $p_i \in O$ sont connectés,
- *non orienté* : il n'y a pas de sens de parcours des arcs $e_{i,j}$ reliant les points p_i à p_j ,
- *pondéré* : les longueurs des arcs $e_{i,j}$ sont pondérées.

G est donc une liste de sommets (les points p_i) et de connexions $e_{i,j}$ entre ces sommets. La longueur totale d'ordre γ de l'arbre est la somme des longueurs (norme euclidienne) pondérées en loi de puissance d'ordre $\gamma \in]0, d[$:

$$L_\gamma = \sum_{e_{i,j} \in G} \| e_{i,j} \|^{\gamma} \tag{1}$$

Le terme *minimal* signifie que, parmi tous les arbres totalement connectés qu'il est possible de construire, le MST est celui dont la longueur est minimale :

$$G^* = \underset{G}{\text{Argmin}} L_\gamma \tag{2}$$

On doit les premiers algorithmes à Kruskal et Prim [COR 94] ; un historique du problème est retracé dans [GRA 85]. Si parfois plusieurs arcs ont la même longueur (points équidistants) plusieurs arbres peuvent théoriquement résulter mais tous auront la même longueur et de surcroît toujours minimale.



3.2. Description d'une forme par MST

Tout objet est représenté par une distribution de points, dont la topologie est significative de sa classe d'appartenance. Par construction, le MST réalise une mesure de cette topologie à l'aide de L_γ . Cependant, cette application n'est pas bijective : pour une même longueur peuvent correspondre deux objets de formes différentes. Le descripteur est donc incomplet et peut conduire à une confusion comme l'illustre l'exemple de la Figure 1. L'introduction du mélange permet de lever cette ambiguïté. Si deux objets sont topologiquement distincts alors le MST défini sur le mélange des deux objets aura une longueur différente des longueurs des MST calculées individuellement sur chacun des objets. On définit un critère d'erreur fondé sur cette propriété.

3.3. Comparaison

Dans les problèmes réels les objets peuvent être similaires et multi-orientés. Ainsi l'exemple décrit par les Figures 2 à 6, traite de symboles architecturaux identiques, translétés et ayant subi une rotation. Une transformation géométrique T peut être alors appliquée itérativement sur l'objet test et le mélange des objets défini à chaque itération (cf. Figures 5 et 6). Puisque les objets sont identiques, il existe une transformation T^* pour laquelle les objets coïncident totalement (cf. Figure 6). Dans ce cas, la topologie définie par l'ensemble résultant du mélange est identique à la topologie propre à l'objet référence (cf. Figures 6 et 3). Par voie de conséquence, les MST construits sur ces ensembles ont même longueur ($L_{12} = L_1$) et l'erreur inter-objet $E^T = |L_{12} - L_1|$ est minimale (elle est nulle dans cet exemple didactique).

L'algorithme consiste à la recherche de la transformation optimale T^* qui minimise l'erreur inter-objet. Finalement, la distance d'ordre γ^1 retenue entre deux objets à comparer est :

$$E^* = \min_T \{E^T\} \text{ with } T^* = \underset{T}{\text{Argmin}} \{E^T\}$$

Dans notre exemple didactique, les Figures 2 à 6 illustrent la construction des arbres sur les objets et leurs mélanges, au cours des diverses transformations géométriques. L'exemple traite d'objets d'épaisseur unité dans le but d'illustrer clairement la construction des arbres, mais l'approche fonctionne sur des symboles aux contours d'épaisseurs variées occasionnant une densité d'arcs plus importante.

L'algorithme proposé permet de reconnaître une forme inconnue par rapport à un prototype de forme, dans un ensemble de classes possibles et connues par avance par la minimisation de E^T . La prise en compte de la topologie du mélange permet de pallier l'incomplétude du descripteur. Notons qu'il existe un lien entre longueur d'arbre et mesure d'entropie. En effet, la longueur du MST construit sur une distribution de points multidimensionnelle participe à l'estimation de l'entropie de Rényi calculée sur cette distribution [RÉN 61]. Ces résultats sont issus

1. γ , est la métrique choisie pour l'expression des longueurs d'arbres.

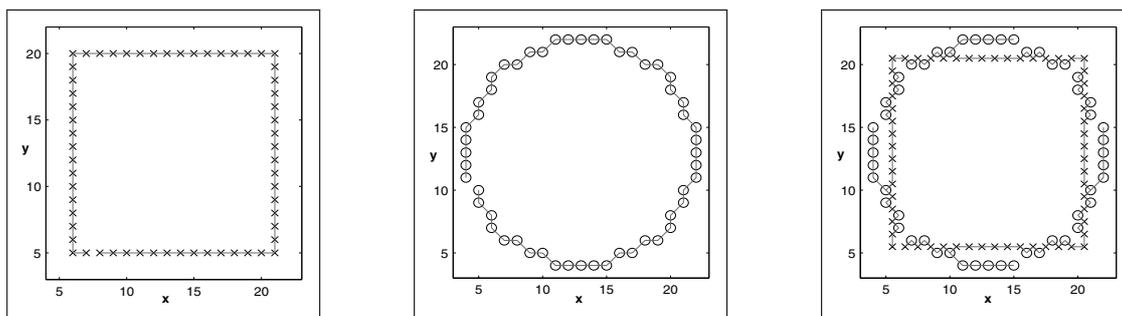


Figure 1. Cas où deux objets de formes différentes ont même longueur : un carré, la longueur du MST associée $L_1 = 59,0$ ($\gamma = 1$) ; un cercle avec $L_2 = 58,9$; le mélange permet de séparer le cercle du carré $L_{12} = 93,9 \neq L_1$.

des travaux de J.M. Steele [STE 88] qui montrent qu'une famille d'arbres (arbre de Steiner, MST, etc.) – vérifiant la propriété de quasi-additivité de Redmond [RED 94] – définie sur une distribution de points sont des estimateurs robustes de l'entropie de Rényi caractérisant cette distribution. Par ailleurs, B. Ma *et al.* [MA 00] ont développé une solution au problème de recalage d'images s'appuyant sur l'analyse suivante : quand deux images sont appariées, leurs distributions de pixels se chevauchent, la densité de probabilité conjointe est donc maximale, par conséquent l'entropie de Rényi estimée sur l'union des distributions est minimale. La démarche qui consiste à ne plus considérer les objets individuellement mais leur mélange, trouve donc une justification d'un point de vue statistique.

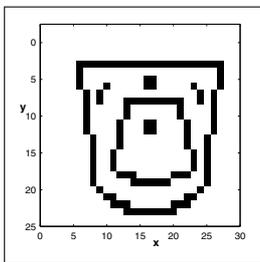


Figure 2. Prototype du symbole lavabo.

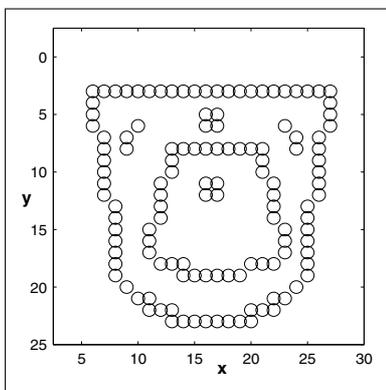


Figure 3. Prototype dans l'espace image
($\gamma = 1$, $L_1 = 138,213$).

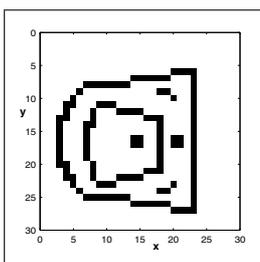


Figure 4. Objet inconnu (lavabo translaté et réorienté).

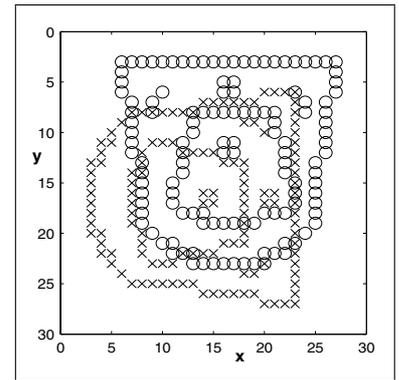


Figure 5. Le mélange des objets ($L_{12} = 242,556$).

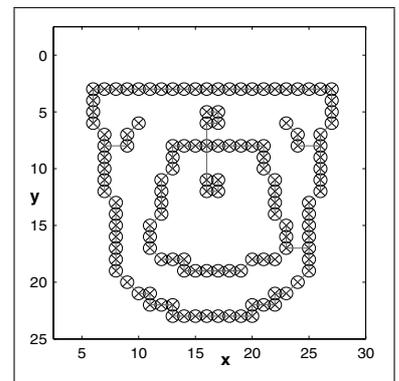


Figure 6. MST sur le mélange dans le cas d'une transformation optimale : les objets sont appariés ($L_{12} = L_1$).

3.4. Application à la reconnaissance de symboles

L'approche décrite précédemment est mise en œuvre dans le cadre de la reconnaissance de symboles graphiques. Les symboles architecturaux que nous traitons (*cf.* Figure 7) sont référencés dans [TAB 01b]. Ils font également partie de la base de test sélectionnée dans le concours de reconnaissance de symboles, organisé dans le cadre de GREC'2003 [GREG03]. Pour montrer les caractéristiques de notre méthode, les prototypes originaux ont subi des translations et des rotations puis des dégradations par l'ajout d'un bruit « poivre et sel » de niveau 30% (*cf.* Figure 8). Un filtrage médian (masque arbitraire 3×3) puis une opération de segmentation ont permis d'isoler les objets à classer (*cf.* Figure 9). Il résulte de ce processus artificiel que les objets à classer diffèrent alors sensiblement des prototypes. Par exemple, des détails, ou même des régions entières de la classe « évier » n'apparaissent plus dans l'objet 1 qui est censé lui appartenir (*cf.* Figure 9). De nombreuses coupures peuvent être également observées, elles perturberont la notion de voisinage essentielle pour la construction de l'arbre de distance. Nous verrons que notre algorithme est assez robuste face à ce type de détérioration et ce grâce au mélange. L'opération d'union des ensembles comble les absences de cer-

Tableau 1. Reconnaissance de symboles par arbres minimaux ($\gamma = 1$) opérant dans l'espace image. Matrice des distances E^* .

	Objets					
	obj.1	obj.2	obj.3	obj.4	obj.5	obj.6
évier	241,151	813,667	796,598	788,656	763,928	661,071
téléviseur	771,110	176,791	338,359	915,355	918,071	805,424
lavabo	769,000	345,120	60,968	919,571	890,757	773,585
table	796,019	848,280	851,381	616,001	759,019	815,000
canapé	762,252	906,763	893,877	775,057	63,083	805,583
lit	689,019	784,095	765,703	806,026	828,034	192,001
Réalité	évier	téléviseur	lavabo	table	canapé	lit

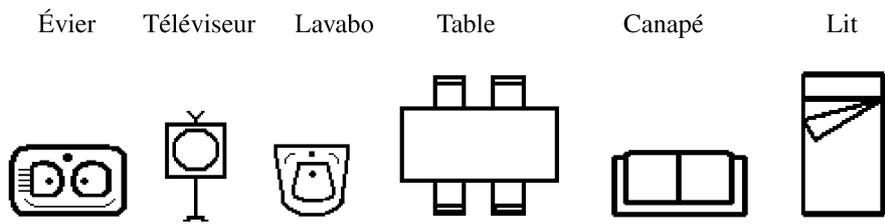


Figure 7. Symboles originaux : 6 classes.

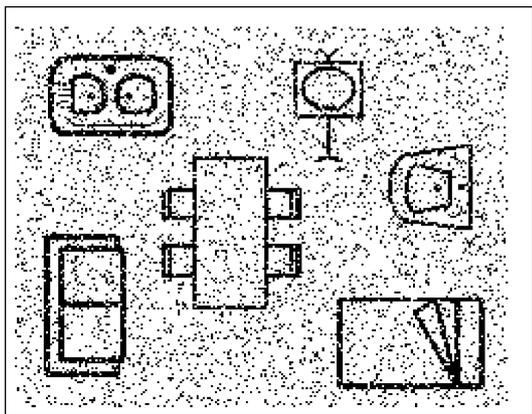


Figure 8. Symboles dégradés : bruit à 30 %.

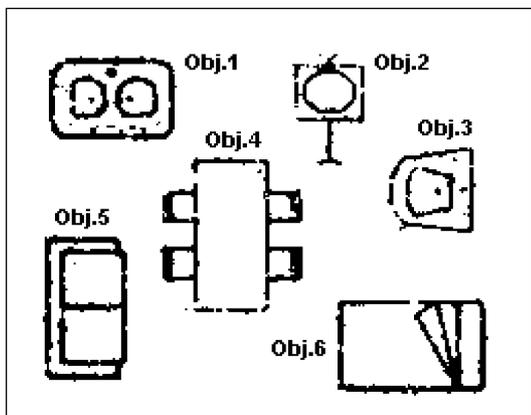


Figure 9. Après filtrage et segmentation : objets inconnus.

tains détails : l'arbre construit sur le mélange conserve une structure proche de celui défini sur le prototype.

Les résultats sont présentés dans le Tableau 1. L'identité de l'objet est celle du symbole dont le prototype minimise l'erreur. Dans cette expérience élémentaire tous les objets ont été reconnus (distances minimales situées sur la diagonale) malgré le niveau de dégradation de certains objets.

4. Réduction du Coût Algorithmique

4.1. Coût algorithmique et temps de calcul

Le coût² algorithmique de notre solution est étroitement lié à la construction d'un MST³. Pour un nombre de points à connecter N évoluant en dimension $\{2,3\} \in d$ le coût algorithmique est majoré à $N^2 \log N$. Si l'on se réfère à l'algorithme présenté dans le Paragraphe 3.3, $T + 1$ arbres sont nécessaires pour conclure sur la distance inter-objets, avec T , le nombre maximal d'itérations dans la recherche. Le premier arbre est calculé sur N

2. Pour lever toute ambiguïté, on parle de *coût* algorithmique comme le nombre d'opérations requises pour un nombre de points donné, par opposition à *complexité* qui fait plus référence à la classe d'un algorithme (quadratique pour les deux algorithmes proposés).

3. Le code retenu est celui développé récemment par A. Hero et O. Michel. Le problème étant NP complet, leur algorithme fournit une bonne approximation de la longueur du MST en un temps raisonnable [HER 99, HER 97].

points, alors que les T autres sont calculés sur l'union des objets, soit $2N$ points⁴. Le coût algorithmique s'élève donc à $C_1 \approx N^2 \log(N) + T \cdot \{(2N)^2 \log(2N)\}$.

N varie suivant la taille des symboles considérés, dans l'expérience présentée le nombre moyen de points par symbole est $N_{moy} = 1000$ ($N_{min} = 470$, $N_{max} = 1026$). Pour des symboles aux rotations aléatoires il est clair que T doit être conséquent d'autant plus que les symboles à identifier sont bruités. On perçoit ici le rôle de T dans le ratio reconnaissance – coût. Dans l'exemple proposé, T a été arbitrairement choisi à 24. Il s'agit donc d'une recherche itérative par pas de 15 degrés à l'aveugle. Dans ces conditions, la méthode discrimine deux objets avec un coût algorithmique moyen atteignant $319,898 * 10^6$ opérations. Pour donner un ordre d'idée, $2h33mn12s$ ont été nécessaires pour établir la matrice des distances (6,6) présentée dans le Tableau 1. Ce temps de calcul résulte de l'implémentation de notre approche sous Matlab version 6.5. Le PC utilisé est cadencé par un processeur Pentium 4 à 3,06 GHz et dispose de 512 Mo de RAM. Le temps d'exécution est quasiment divisé par 6 lorsque la recherche est guidée par l'identification des axes d'inertie des objets soit environ $26mn33s$. Ces chiffres sont donnés à titre indicatif et n'ont d'autre prétention que de fixer un ordre de grandeur. De plus, ils résultent d'un programme non optimisé n'utilisant pas les fonctions accélératrices désormais disponibles sur Matlab.

Quelques pistes visant à réduire le temps de calcul sont évoquées par la suite. Le coût algorithmique est un point critique mais, il n'est pas rédhibitoire et plusieurs leviers existent tant algorithmiques qu'issus du domaine du Traitement d'Image. Ainsi, l'emploi d'un algorithme de construction des arbres plus performant (D. Karger et P. Klein dans [KAR 95]) ou bien une approche de type incrémentale (M. Soss dans [SOS 98]) sont les pistes actuellement explorées par la communauté. En pratique, une autre solution consiste à diminuer le nombre de transformations géométriques appliquées (rotations), il est aussi possible d'entreprendre une analyse en composantes principales pour accélérer la recherche de la transformation optimale (identification des axes d'inertie). Des méthodes de traitement d'image peuvent aussi être utilisées pour réduire l'ensemble de pixels à prendre en compte, comme par exemple des opérations d'érosion morphologique [SER 88]. L'extraction du squelette minimal d'une image est au centre de certains travaux où la minimisation de la redondance est en filigrane [MAR 86, KRE 94]. Cependant, ces approches morphologiques conduisent à une perte d'information structurelle sur l'objet à classer; cela peut s'avérer pénalisant dans notre approche fondée sur une description spatiale. L'alternative que nous proposons consiste à opérer une transformation globale de l'image préservant la dimension structurelle des objets.

4. Sous l'hypothèse que les objets soient définis par le même nombre de points.

4.2. Réduction du coût par Transformation en Cosinus Discrète

L'idée centrale est de travailler dans un mode de représentation de l'image dans lequel les coefficients soient mieux décorrélés. La Transformée en Cosinus Discrète (TCD) est retenue. Introduite en 1974 par N. Ahmed, T. Natarajan et K. Rao [AHM 74], elle s'exprime de la façon suivante :

$$F(u,v) = \frac{1}{\sqrt{2N}} C(u)C(v) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} I(x,y) \cdot \cos \frac{\pi u(2x+1)}{2N} \cos \frac{\pi v(2y+1)}{2N}$$

avec: $I(x,y)$ l'image originale et $F(u,v)$ l'image transformée; x,y , les coordonnées dans le domaine spatial et u,v dans le domaine fréquentiel;

$$C(u) = C(v) = \frac{1}{\sqrt{2}} \quad \text{si } u,v = 0 \quad C(u) = C(v) = 1 \quad \text{sinon}$$

Le terme $\frac{1}{\sqrt{2}}$ permet de rendre la transformation orthogonale.

La TCD est très proche de la Transformée de Fourier (TF) et possède des propriétés intéressantes pour notre objectif de minimisation du nombre de points décrivant un symbole :

- *concentration* de l'énergie assez élevée (convergence relativement rapide);

- *corrélation* entre les coefficients transformés faible.

De plus, K. Rao montre que la TCD fournit une meilleure approximation d'une image avec moins de coefficients que la TF [RAO 90]. Par ailleurs, la capacité de la TCD à décorréler les données s'explique statistiquement. En effet, en considérant les pixels de l'image comme les réalisations d'un processus aléatoire stationnaire, certains travaux [GUI 94, RAO 90] montrent que la TCD est une bonne estimation de la transformation de Karhunen-Loeve; de plus, elle ne se heurte pas à la diagonalisation de la matrice de covariance toujours délicate. Ces propriétés ont été largement éprouvées puisque la TCD fut retenue dans le standard de compression JPEG⁵ en 1992 [CHE 84, WAL 91].

5. Reconnaissance de symboles graphiques dans l'espace « compressé »

5.1. Démarche et éléments de justification

Certaines notions de compression d'image [RAO 90, GER 99, PEN 93] sont utilisées dans ce Paragraphe. Les concepts entou-

5. Joint Photographic Expert Group.

rant les étapes de transformation, quantification et codage guident certains de nos choix. Par exemple, la TCD n'est pas appliquée sur l'image entière, mais par bloc. En effet, pour des raisons de temps de calcul il est préférable de décomposer l'image originale en K blocs de taille $M \times M$ et de calculer la TCD sur chacun des blocs. Cependant, il faut veiller à ce que la taille des blocs (liée directement à l'abandon de certaines fréquences) soit suffisante pour éviter l'effet mosaïque. Un autre argument concernant spécialement notre application renforce ce choix : construire un arbre dans l'espace transformé, sans partition de l'image originale, conduit inévitablement à la perte d'informations spatiales dans l'image. En effet, après intégration (sur toute l'image) il est difficile de relier une fréquence à une région de l'objet dans l'image. L'information spatiale est diluée dans l'ensemble des fréquences. En partitionnant l'image et en ordonnant les blocs, on conserve une part de la dimension spatiale, ce qui devrait faciliter la comparaison des objets par des arbres.

Les matrices transformées ne subiront pas de quantification. La quantification n'ayant pour but que d'atténuer l'amplitude des coefficients afin de minimiser le nombre de bits nécessaires à leur codage. Il semble au contraire intéressant de conserver la plus grande variabilité intercoefficients de manière à mieux discerner les objets.

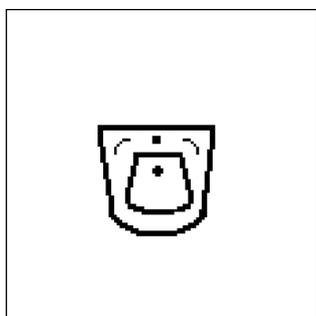


Figure 10. Image 100×100 (espace original).

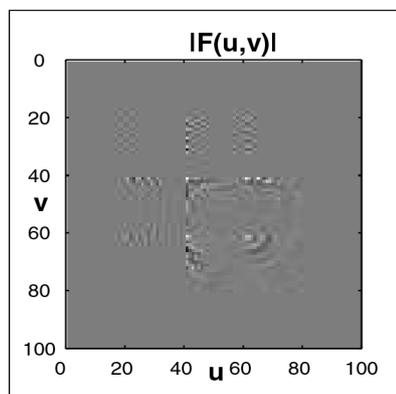


Figure 11. TCD par bloc 20×20 (espace transformé). L'amplitude des coefficients est codée en niveaux de gris.

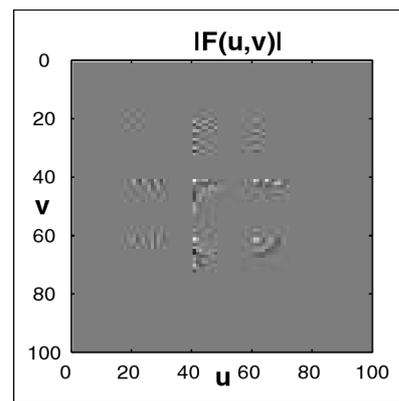


Figure 12. Application de l'algorithme Zig-Zag. Seul 19,5% des coefficients sont conservés (78 coefficients par blocs).

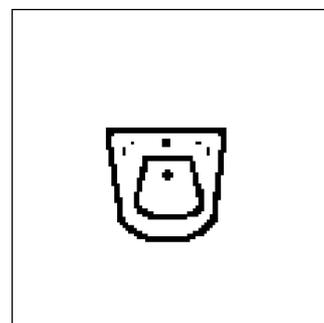


Figure 13. Image reconstruite à partir de la TCD tronquée.

La TCD est très compacte, autrement dit peu de coefficients sont nécessaires à son approximation. De plus, la majorité des coefficients non nuls sont concentrés autour de la composante continu. Plus on s'éloigne des basses fréquences (dans les deux directions) et plus les coefficients sont proches de zéro. Cette répartition peut être utilisée pour omettre certains coefficients sans risque d'altérer l'image originale (cf. Figures 10 à 13). Cette tendance est d'ailleurs utilisée par des techniques de balayage (théorie du codage) qui tentent de minimiser l'adresse du dernier coefficient non nul à transmettre. L'exemple de l'algorithme Zig-Zag est assez célèbre [GUI 94, GER 99]. Les arbres seront donc construits sur des versions tronquées de la TCD par bloc.

5.2. Application à notre exemple didactique

L'expérience d'identification menée au Paragraphe 3.4 est reproduite dans les mêmes conditions. L'algorithme présenté dans le Paragraphe 3.3 est appliqué en intégrant les étapes décrites au Paragraphe 5.1. À savoir, les symboles sont préalablement plongés dans l'espace CD (K blocs de taille $M \times M$). Une sélection des coefficients est opérée suivant l'algorithme Zig-Zag (cf. Figure 12). Seul N' points par bloc sont conservés.

Tableau 2. Reconnaissance de symboles par arbres minimaux ($\gamma = 1$) opérant dans l'espace Cosinus Discret. Matrice des distances E^* .

	Objets					
	obj.1	obj.2	obj.3	obj.4	obj.5	obj.6
évier	16,752	11031,616	11327,937	91,292	67,839	48,822
téléviseur	62,281	7,885	10561,515	110,458	82,309	78,956
lavabo	68,056	30,257	13,196	110,748	96,129	20298,654
table	19217,062	11955,169	11892,729	29,400	23828,005	54,041
canapé	19790,828	10893,603	11545,823	82,740	17,976	54,041
lit	19372,944	11006,482	00291,834	83,158	25060,909	29,064
Réalité	évier	téléviseur	lavabo	table	canapé	lit

Dans l'espace CD, un bloc est un sous-ensemble de N' points de dimension 3 ($u, v, F(u, v)$). Un MST est construit sur chaque sous-ensemble. Au final, K MSTs sont calculés et leurs longueurs sont les composantes du vecteur caractéristique⁶ du symbole traité. La distance inter-symbole est la norme Euclidienne entre vecteurs caractéristiques. Au final, la distance inter-symbole retenue est la distance minimale observée au cours des différentes transformations géométriques (T). Les résultats sont présentés dans le Tableau 2. Ils correspondent au jeu de paramètres $\{K = 25, M = 20, N' = 78\}$. Ce choix est guidé par les remarques formulées dans le Paragraphe 5.1.

Tous les objets sont reconnus (cf. Tableau 2). La qualité de la reconnaissance est mesurée par un critère Δ évoquant le caractère discriminant de la décision. Δ est évalué par comparaison des distances associées aux deux symboles émergents. Par exemple, le *téléviseur* est identifié avec un indice de discrimination s'élevant à $\Delta = \frac{30,257-7,885}{7,885} = 283,7\%$. Dans le cadre des hypothèses formulées dans cette expérience (modèle, niveau de bruit...), les deux approches ont un niveau de discrimination du même ordre de grandeur. Les niveaux médians sont de 244,3% vs. 271,7% respectivement lorsque les arbres opèrent dans l'espace image et dans l'espace CD. Par contre, le gain algorithmique est indéniable lorsque la comparaison a lieu dans l'espace transformé, on se propose de l'estimer.

Le coût de calcul de la TCD est négligé devant celui engendré par la construction des MSTs. Cette hypothèse nous paraît acceptable compte tenu de la classe d'algorithmes rapides proposée par C. Loeffler, A. Ligtenberg et G. Moschytz dans [LOE 89]. Les fonctions cosinus sont pré-calculées et la séparabilité des noyaux est exploitée. Le coût (approximatif) évolue donc comme $C_2 \approx K \cdot N^2 \log(N) + T \cdot \{K \cdot (N)^2 \log(N')\}$. Le premier terme fait référence au calcul du vecteur caractéristique de l'objet prototype. Le deuxième, à celui du vecteur caractéristique du mélange des objets à comparer. Pour $\{K = 25, T = 24,$

$N' = 78, N = 1000\}$ le gain algorithmique moyen s'élève à $\frac{C_1}{C_2} = 44,463$. Ce gain est significatif et provient en grande partie du deuxième terme de C_2 par rapport à celui de C_1 . La complexité de construction des arbres est contournée en construisant plus d'arbres (K fois plus) mais sur un volume de données moindre ($\frac{N'}{N} = 3,9\%$). On tire donc profit de l'évolution polynomiale du coût en nombre de points. Moins de $3mn$ sont alors nécessaires pour acquérir les résultats de la reconnaissance synthétisés dans le Tableau 2. À titre comparatif, $2h33mn12s$ sont requises lorsque la technique opère dans l'espace image. Des résultats complémentaires – relatif à un deuxième échantillon de symboles – sont fournis en Annexe.

5.3. Test à plus grande échelle

La méthode est évaluée sur 500 symboles issus de la base de données GREC'2003 [GRE03]. L'échantillon test est constitué de 25 classes comptant 20 candidats par classe. Chaque classe est représentée par un symbole prototype (cf. Figure 14). Il s'agit d'images binaires, de taille (100×100) pixels, et de faible résolution ($< 240 dpi$). Les objets à reconnaître sont des versions translattées, tournées et bruitées des prototypes (cf. Figure 15). Les rotations sont aléatoires par pas de 10 degrés sur le support $[0, 360^\circ]$. Le modèle de bruit utilisé est additif, de type « poivre et sel ». Deux expériences sont menées suivant le niveau de bruit considérés: $RBS = 20\%$ et $RBS = 30\%$. Les critères d'évaluations sont le taux de reconnaissance et le temps de calcul. Les résultats sont synthétisés dans le Tableau 3.

Dans le cadre des hypothèses formulées dans ces expériences (modèle et niveau de dégradation etc.) le taux de reconnaissance est encourageant. Il atteint 94,6% malgré un $RBS = 30\%$. De plus ce taux illustre une vérité terrain, puisque la base recense des symboles d'origines variées (architecturaux, électriques) et aux caractéristiques géométriques diverses. Par exemple, certains symboles sont définis par des régions disjointes, d'autres possèdent des axes de symétrie, des cas de symboles ouverts sont même représentés. Nous pensons que l'opération de mélange joue un rôle central dans le processus de comparaison. Le

6. La caractérisation par des arbres des coefficients transformés n'a de sens que si les symboles sont décomposés sur la même partition. Les symboles ayant des tailles variables, le rectangle englobant maximum est pris comme référence pour le calcul de la TCD.

Tableau 3. Reconnaissance de symboles par arbres minimaux opérant dans l'espace Cosinus Discret : résultats.

	Classes	Candidats	RBS	Reconnaissance	Temps moyen/candidat
Exp. 1	25	500	20 %	98,8 %	18,458 s
Exp. 2	25	500	30 %	94,6 %	18,565 s

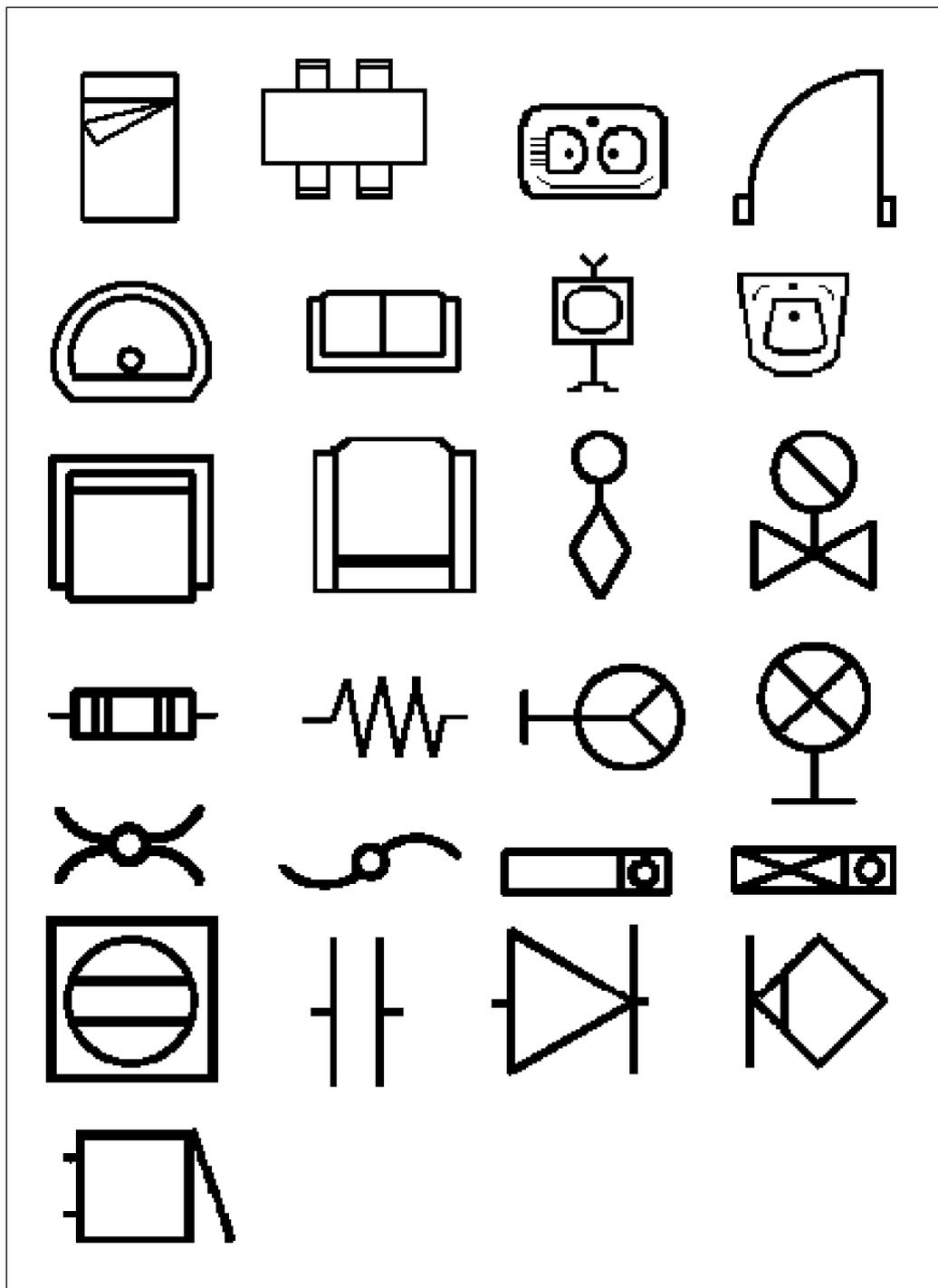


Figure 14. Les symboles originaux : 25 classes.

mélange des objets dans l'espace image est un moyen d'accroître les différences entre les objets à comparer. Cela se répercute dans l'espace CD. On peut le vérifier en confrontant les décompositions propres à chacun des objets et celles relatives aux mélanges des objets. De plus, dans l'espace CD, les coefficients les plus prépondérants des décompositions sont conservés. Les arbres sont construits sur les coefficients qui subsisteraient (non nuls) si l'on avait appliqué une table de quantification. C'est un critère de sélection qui permet de s'affranchir de la définition d'un seuil toujours difficile.

Cependant, des erreurs d'identification existent. Elles se produisent majoritairement lorsque les images contiennent des données aberrantes (cf. Figure 16). Il s'agit de pixels déconnectés des objets originaux qui ont résisté à l'opération de débruitage⁷ menée avant la comparaison. Le fait que ces artefacts gravitent dans un voisinage éloigné des objets, affecte leurs centres de gravités et par conséquent l'alignement des objets n'est plus

optimal. La tolérance à ces données aberrantes peut être améliorée en circonscrivant le calcul du centre de gravité aux points inclus dans le rectangle englobant l'objet.

Un objet est reconnu en 18,565s en moyenne. Cela représente un gain d'un facteur 45 par rapport à une reconnaissance directe dans l'espace image. Ce temps de reconnaissance moyen devient raisonnable même s'il n'est pas encore compatible avec les contraintes temporelles d'une application réelle. Si d'un point de vue résistance au bruit, l'expression des images dans l'espace CD, semble adaptée, il vient rapidement à l'esprit l'utilisation d'autres transformations pour améliorer le temps de calcul. Par exemple, la transformation de Haar [HAA 10], par définition adaptée aux images binaires. On peut également citer la transformée en ondelettes sélectionnée dans le nouveau standard JPEG-2000 (codage en sous-bandes [MAL 89, VET 95]). Toutefois, il convient de vérifier que le gain algorithmique final soit réel. Le coût estimé dans le Paragraphe 4.1 concerne des points de dimension 3. Dans l'espace position-échelles les points sont de dimension supérieure, ce qui rend caduque cette estimation. Il convient donc de réestimer à la hausse le coût algorithmique de construction des arbres. Par conséquent, une meilleure approximation grâce aux échelles doit compenser une majoration du coût.

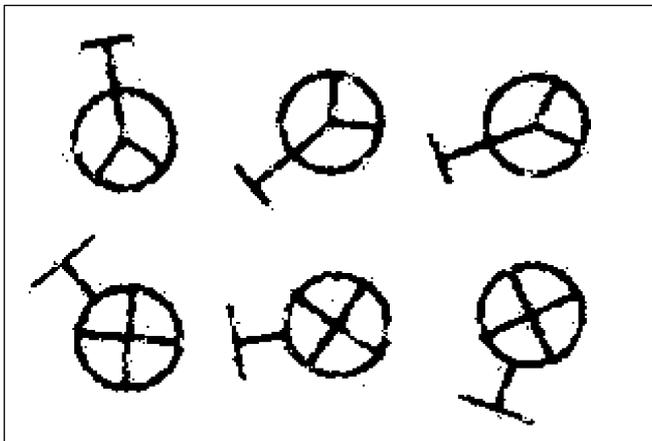


Figure 15. Exemples d'objets à classe appartenant à deux classes visuellement proches.

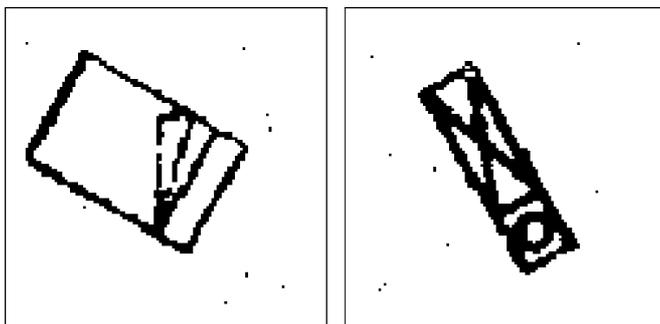


Figure 16. Exemples d'objets à classer : cas de données aberrantes.

7. Il s'agit d'un filtrage médian, dont la taille arbitraire 3×3 n'est pas forcément adaptée au niveau de bruit en action. Notre but est d'estimer rigoureusement les capacités intrinsèques de la méthode. Pour cela, nous ne faisons pas appel à des techniques d'estimations de bruit suivies de filtres adaptés qui auraient certainement pour effet d'adoucir artificiellement les conditions expérimentales.

6. Conclusion

Dans cet article, un nouveau descripteur de forme s'appuyant sur des arbres minimaux est proposé. La description et la comparaison, étapes clés de la méthode, sont détaillées et justifiées. En reliant tous les points caractérisant un objet, sous la contrainte de définir un chemin minimal, une mesure de la topologie de l'objet est définie. Il s'agit d'une technique fondée sur les régions.

Le temps de calcul – dépendant du coût algorithmique de construction des arbres – est un point critique mais non rédhibitoire. Plusieurs pistes visant à le réduire sont décrites, telle que l'expression de l'objet cible dans l'espace Cosinus Discret : norme JPEG-1992. On opère alors non plus dans l'espace image mais dans un espace compact où les données sont mieux décorréliées. Les arguments justifiant le choix de l'espace Cosinus Discret sont avancés. Cette piste conduit à une diminution sensible du coût tout en conservant un niveau de discrimination inter-objets significatif. On gagne un facteur 45 à niveau de discrimination constant. Ces résultats sont issus de tests à moyenne échelle mettant en jeu des symboles variés et réalistes issus d'une base de données référence [GRE03]. Rappelons que les expérimentations portent sur des symboles isolés, d'épaisseurs unitaires et ayant subis des translations et rotations auxquelles s'ajoutent des dégradations croissantes (bruit « poivre et sel »). Des tests plus approfondis – intégrant plusieurs types de symboles et modèles de bruit – confirment ces résultats. Le comportement de la méthode face à un nombre plus important de classes est en cours d'étude.

Cependant, la reconnaissance des objets n'est pas invariante aux échelles. Dans la version proposée de l'algorithme, le même objet représenté à deux échelles différentes n'est pas reconnu. Cette limite est due à l'intégration du mélange des objets dans le processus de comparaison. C'est le seul inconvénient de cette phase de mélange qui reste un moyen de rendre le descripteur complet et de pallier aux occultations partielles des objets. De plus, il joue un rôle important dans la discrimination. Si deux objets sont distincts, leur mélange accentue artificiellement leurs différences, ce qui se traduit dans l'espace CD. Les perspectives sont multiples ; à court terme, la généralisation du modèle à des images à niveaux de gris et couleurs élargirait le champ d'investigation de l'approche au problème de la reconnaissance d'objets dans des images naturelles. À moyen terme, l'intégration du modèle au niveau de la segmentation nous paraît

être une piste à explorer. Cette orientation va dans le sens des travaux de synthèses de K. Tombe [TOM 03] et J. Lladós [LLA 03] qui soulèvent implicitement la question fondamentale de l'influence de la segmentation sur les performances de la reconnaissance dans le contexte des documents denses et bruités.

Annexe : résultats complémentaires

Reconnaissance de symboles complexes (c'est-à-dire multi-orientés, dégradés et parfois définis par des régions disjointes). Support: image binaire de faible résolution (< 240 dpi).

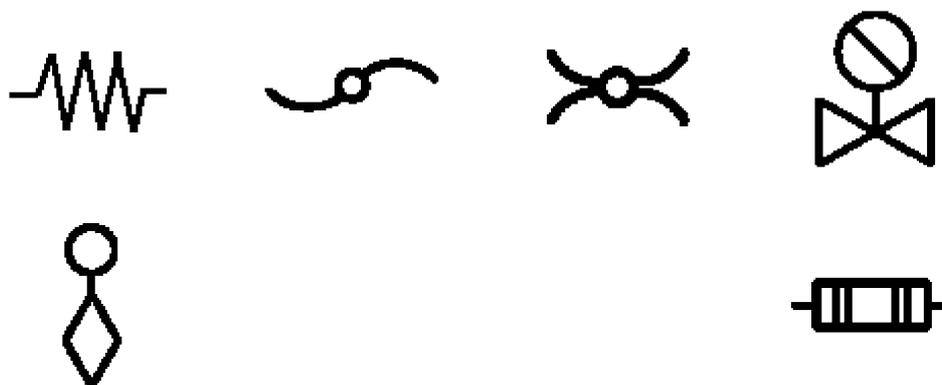


Figure 17. Symboles originaux (respectivement, de gauche à droite, symb. 1 à symb. 6).

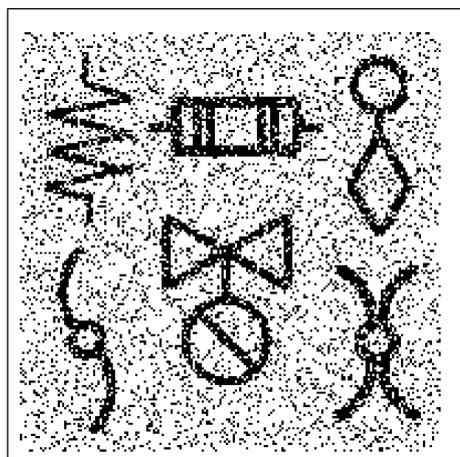


Figure 18. Symboles dégradés (bruit « poivre et sel », niveau NSR = 30 %).

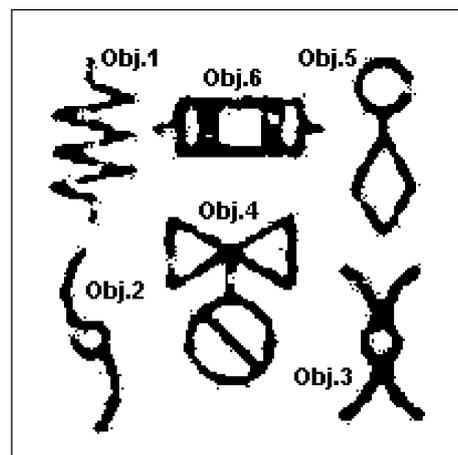


Figure 19. Après filtrage et segmentation : objets à classer.

Tableau 4. Reconnaissance de symboles par arbres minimaux ($\gamma = 1$) opérant dans l'espace image. Matrice des distances E^* .
Niveau de discrimination median: $\Delta = 278,1\%$, temps de calcul : 2h28mn34s.

	Objets					
	obj.1	obj.2	obj.3	obj.4	obj.5	obj.6
symb.1	134,656	588,306	651,173	1146,886	564,485	749,122
symb.2	580,235	138,171	597,887	1238,154	522,313	935,886
symb.3	589,093	598,551	90,892	1150,123	633,514	770,414
symb.4	1087,915	1154,539	1126,148	272,472	1112,076	910,770
symb.5	564,656	494,585	643,242	1175,009	91,656	767,236
symb.6	679,679	858,757	737,521	991,402	703,514	198,122
Réalité	symb.1	symb.2	symb.3	symb.4	symb.5	symb.6

Tableau 5. Reconnaissance de symboles par arbres minimaux ($\gamma = 1$) opérant dans l'espace Cosinus Discret.
Matrice des distances E^* .
Niveau de discrimination median: $\Delta = 367,2\%$, temps de calcul : 2mn3s.

	Objets					
	obj.1	obj.2	obj.3	obj.4	obj.5	obj.6
symb.1	8,762	13564,972	20644,405	87,801	18334,453	24845,423
symb.2	54,728	14,357	38,627	107,445	47,365	63,031
symb.3	50,695	13593,552	6,243	100,729	18347,104	24884,850
symb.4	18280	13541,843	20588,324	16,341	18262,776	24870,032
symb.5	45,373	17,158	33,005	98,618	7,271	24876,376
symb.6	40,946	24,934	39,714	92,548	32,276	8,507
Réalité	symb.1	symb.2	symb.3	symb.4	symb.5	symb.6

Références

- [ADA 01] ADAM S., OGIER J.-M., CARIOU C., MULLOT R., GARDES J., LECOURTIER, «Utilisation de la transformée de Fourier-Mellin pour la reconnaissance de formes multi-orientées et multi-échelles: application à l'analyse automatique de documents techniques», *Traitement du Signal*, vol. 18, n° 1, 2001, p. 17-33.
- [AHM 74] AHMED N., NATARAJAN T., RAO K., «On image processing and a discrete cosine transform», *IEEE Transactions on Computers*, vol. C-23, n° 1, 1974, p. 90-93.
- [AHS 01] AH-SOON C., TOMBRE K., «Architectural symbol recognition using a network of constraints», *Pattern Recognition Letters*, vol. 22, n° 2, 2001, p. 231-248.
- [AIR 95] AIRAULT S., JAMET O., «Détection et restitution automatique du réseau routier sur des images aériennes», *Traitement du Signal*, vol. 12, n° 2, 1995, p. 189-200.
- [BHA 94] BHATTACHARJEE S., MONAGAN G., «Recognition of cartographic symbols», *IAPR Workshop on Machine Vision Applications*, Japan 1994, p. 226-229.
- [BUN 00] BUNKE H., «Recent developments in graph matching», *In the 15th International Conference on Pattern Recognition*, vol. 2, Spain 2000, p. 117-125.
- [CHE 84] CHEN W., PRATT K., «Scene adaptative coder», *IEEE Transactions on Communications*, vol. COM-32, 1984, p. 225-232.
- [COR 94] CORMEN T., LEISERSON C., RIVEST R., *Introduction to algorithms*, The MIT Press, 1994.
- [DEL 94] DELLA VENTURA A., SCETTINI R., «Graphic symbol recognition using a signature technique», *In the 12th International Conference on Pattern Recognition*, Israel 1994, p. 533-535.
- [EBI 94] EBI N., LAUTERBACH B., ANHEIER W., «An image analysis system for automatic data acquisition from colored scanned maps», *Machine Vision and Application*, 1994, p. 148-164.
- [FUK 92] FUKUMI M., OMATU S., TAKEDAN T., «Rotation invariant neural pattern recognition system with application to coin recognition», *IEEE Transactions on Neural Networks*, vol. 3, 1992, p. 272-279.
- [GER 99] GERSHO A., GRAY R. M., *Vector Quantization and Signal Compression*, Kluwer Academic Publishers, Boston, 1992 (6^e édition, 1999).
- [GRA 85] GRAHAM R., HELL P., «On the history of minimum spanning tree problem», *IEEE Annals of the History of Computing*, vol. 7, n° 1, 1985, p. 43-57.
- [GRE03] *Fifth IAPR International Workshop on Graphics Recognition*, Computer Vision Center, Barcelona, Catalonia, Spain, July 30-31, 2003.

- [GUI 94] GUICHARD J., NASSE D., «Traitement des images numériques pour la réduction du débit binaire», *Le Traitement du Signal – Actes des Forums de France Télécom Recherche*, n° 2, 1994, p. 1-15.
- [HAA 10] HAAR A., «Zur Theorie der Orthogonalen Funktionensysteme», *Math. Ann.*, n° 69, 1910, p. 331-371.
- [HER 97] HERO A., MICHEL O., «Robust estimation of point process intensity features using k -minimal spanning trees», *IEEE International Symposium on Information Theory*, Germany June 1997, page 74.
- [HER 99] HERO A., MICHEL O., «Asymptotic theory of greedy approximations to minimal K -point random graphs», *IEEE Transactions on Information Theory*, vol. IT 45, 1999, p. 1921-1939.
- [HUA 99] HUANG C. L., LIN W., «Automatic Taiwanese municipal color reading system», *Vision Interface '99*, Canada, 19-21 May 1999.
- [KAR 95] KARGER D., KLEIN P., TARJAN R., «A randomized linear-time algorithm to find Minimum Spanning Trees», *Journal of the Association for Computing Machinery (ACM)*, vol. 42, n° 2, 1995, p. 321-328.
- [KRE 94] KRESCH R., MALAH D., «Morphological reduction of skeleton redundancy», *Signal Processing*, vol. 38, 1994, p. 143-151.
- [LEE 92] LEE S.-W., «Recognizing hand-drawn electrical circuit symbols with attributed graph matching», *H.S. Baird, H. Bunke, K. Yamamoto (eds.), Structured Document Image Analysis*, Springer-Verlag, 1992, p. 340-358.
- [LIN 85] LIN X., SHIMOTSUJI S., MINOH M., SAKAI T., «Efficient diagram understanding with characteristic pattern detection», *Computer Vision, Graphics and Image Processing*, vol. 30, n° 1, 1985, p. 84-106.
- [LIN 87] LIN C., «New forms of shape invariants from elliptic Fourier Descriptors», *Pattern Recognition*, n° 20, 1987, p. 535-545.
- [LLA 01] LLADÓS J., VALVENY E., SANCHEZ G., MARTÍ E., «Symbol Recognition: current advances and perspectives», *GREC 2001*, 2001, p. 104-127.
- [LLA 03] LLADÓS J., VALVENY E., SÁNCHEZ G., «A Case Study of Pattern Recognition: Symbol Recognition in Graphic Documents», *Pattern Recognition in Information Systems*, 2003, p. 1-13.
- [LLA 04] LLADÓS J., SÁNCHEZ G., «Graph matching versus graph parsing in graphics recognition – a combined approach», *International Journal of Pattern Recognition and Artificial Intelligence (IJPRAI)*, vol. 18, n° 3, 2004, p. 455-473.
- [LOE 89] LOEFFLER C., LIGTENBERG A., MOSCHYTZ G., «Practical Fast 1-D DCT Algorithms with 11 Multiplications», *International Conference on Acoustics, Speech, and Signal Processing (ICASSP '89)*, 1989, p. 988-991.
- [MA 00] MA B., HERO A., GORMAN J., MICHEL O., «Image registration with minimal spanning tree algorithm», *IEEE International Conference on Image Processing*, Vancouver October 2000.
- [MAL 89] MALLAT S. G., «Theory for Multiresolution Signal Decomposition: The Wavelet Representation», *IEEE Transactions on PAMI*, vol. 11, n° 7, 1989, p. 674-693.
- [MAR 86] MARAGOS P., SHAFER R., «Morphological skeleton representations and coding of binary images», *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 34, n° 5, 1986, p. 1228-1244.
- [MAR 97] MARIANI R., LECOURT F., DESEILLIGNY M., LABICHE J., LECOURTIER Y., «Interprétation de cartes géographiques. Algorithme de reconstruction des réseaux hydrographiques et routiers», *Traitement du Signal*, vol. 14, n° 3, 1997, p. 317-335.
- [MER 97] MERCY P., TACONET B., MAINI J.-L., «Une méthode de vectorisation de dessin techniques adaptée aux plans mécaniques», *16^{ième} colloque GRETSI*, Grenoble, 15-19 septembre 1997, p. 869-872.
- [OGI 95] OGIER J.-M., MULLOT R., LABICHE J., LECOURTIER Y., «Interprétation de document par cycles perceptifs de construction d'objets cohérents. Application aux données cadastrales», *Traitement du Signal*, vol. 12, n° 6, 1995, p. 627-637.
- [PAS 96] PASTERNAK B., «Adaptierbares kernsystem zur interpretation von zeichnungen», *Dissertation zur erlangung des akademischen grades eines doktors der naturwissenschaften (Dr. rer. nat.)*, Universität Hamburg, April 1996.
- [PEI 92] PEI S., LIN C., «Normalisation of rotationally symmetric shapes for pattern recognition», *Pattern Recognition*, n° 25, 1992, p. 913-920.
- [PEN 93] PENNEBAKER W., MITCHELL J., *The JPEG Still Image Data Compression Standard*, Van Nostrand Reinhold, New York, 1993.
- [RAO 90] RAO K., YIP P., *Discrete Cosine Transforms – Algorithms, Advantages, Applications*, Academic Press, Boston, 1990.
- [RED 94] REDMOND C., YUKICH J., «Limit theorems and rates of convergence for Euclidean functionals», *Annals of Applied Probability*, vol. 4, n° 4, 1994, p. 1057-1073.
- [RÉN 61] RÉNY A., «On measures of entropy and information», *Symposium on Mathematics Statistics and Probabilities*, Berkeley 1961, p. 547-561.
- [SER 88] SERRA J., *Image analysis and mathematical morphology Volume 2: Theoretical Advances*, Academic Press, London, 1988.
- [SOS 98] SOSS M., «On the size of the sphere on influence graph», PhD thesis, Mc Gill University Scholl of Computer Science Montreal, 1998.
- [STE 88] STEELE J., «Growth rate of euclidean minimal spanning trees with power weiggthed edges», *Annals of Probability*, vol. 16, 1988, p. 1767-1787.
- [TAB 01a] TABBONE S., WENDLING L., TOMBRE K., «Indexing of technical line drawings based on F-signatures», *In the 6th International Conference on Document Analysis and Recognition, Seattle (Washington, USA)*, 2001, p. 1220-1224.
- [TAB 01b] TABBONE S., WENDLING L., TOMBRE K., «Indexing of technical line drawings based on F-Signatures», *6th International Conference on Document Analysis and Recognition (ICDAR)*, Seattle, Washington, USA Sept 2001, p. 1220-1224.
- [TAB 04] TABBONE S., WENDLING L., «Recognition of symbols in grey level line drawings from an adaptation of the Radon transform», *In the 17th International Conference on Pattern Recognition*, Cambridge (UK), 2004, p. 570-573.
- [TAX 90] TAXT T., OLAFSDOTTIR J., DAEHLEN M., «Recognition of handwritten symbols», *Pattern Recognition*, vol. 23, 1990, p. 1155-1166.
- [TOM 03] TOMBRE K., LAMIROY B., «Graphics Recognition – from Re-engineering to Retrieval», *In 7th International Conference on Document Analysis and Recognition (ICDAR 2003)*, IEEE Computer Society, 2003, p. 148-156.
- [TRI 96] TRIER D., JAIN A., TAXT T., «Features extraction methods for character recognition – a survey», *Pattern Recognition*, vol. 29, 1996, p. 641-662.
- [VET 95] VETTERLI M., KOVACEVIC J., *Wavelets and Subband Coding*, Englewood Cliffs, Prentice Hall, 1995.
- [WAL 91] WALLACE G., «The JPEG still picture compression standard», *Communications of the Association for Computing Machinery*, vol. 34, n° 4, 1991, p. 30-44.
- [YU 94] YU Y., SAMAL A., SETH S., «Isolating symbols from connected lines in a class of engineering drawings», *Pattern Recognition*, vol. 27, n° 3, 1994, p. 391-404.



Patrick **Franco**

Patrick Franco est né à Sète (Hérault) en 1971. Il vient d'être récemment nommé Maître de Conférence à l'Université de la Rochelle après avoir obtenu un D.E.A. en Automatique et Traitement du Signal au LIRMM (Université Montpellier II, 1997) et soutenue sa thèse (Génie Informatique, Automatique et Traitement du Signal) au L3I (Université de la Rochelle, 2001). Ses recherches actuelles portent sur le développement de signatures pour l'image et leurs applications dans des problèmes de reconnaissances de formes.



Jean-Marc **Ogier**

Jean-Marc Ogier est Professeur à l'Université de la Rochelle depuis Sept. 2001. Il mène ses recherches sur l'analyse de documents numériques dans le laboratoire L3i où il est co-responsable de l'équipe projet Système d'Informations Documentaires (SID). Il est Membre du Conseil d'Administration du G.R.C.E (Groupe de Recherche en Communication Écrite) depuis 2000. Il est co-animateur du GT5.2. (Écrit) du CNRS-GRDI3, dont il assure la direction adjointe depuis 2006. Il est directeur scientifique de l'ACI Masses de données Madonne, sur le thème de la numérisation du patrimoine 2003-2006 ; il est également membre du programme RNTL EPEIRES, sur l'évaluation de performances concernant la reconnaissance de symboles depuis Janvier 2005.



Pierre **Loonis**

Pierre Loonis est actuellement Maître de Conférence – HDR à l'Université de la Rochelle. Ses recherches entrent dans le cadre de la reconnaissance des formes pour des applications complexes et évolutives. Au sein du L3I, il est responsable du projet e-FISH (Évaluation de la Fraîcheur par l'Instrumentation des Sens Humains), visant à la conception de méthodes pour la reconnaissance du phénomène du vieillissement des produits de la mer post-mortem. L'adaptativité d'un système à son environnement y est étudié à travers les approches de fusion d'information multi-sources en tenant compte des aspects temporels des observations et de la notion de contexte. eFISH est lauréat du concours Tremplin-recherche 2006 organisé par le Sénat pour promouvoir le transfert de la Recherche dans la Société.



Rémy **Mullot**

Rémy Mullot est Professeur à l'Université de la Rochelle depuis Sept. 2001 après avoir été Maître de Conférence à l'Université de Rouen (1991). Ses recherches – traitent de l'analyse de documents numériques – et sont menées au sein du I3I dans l'équipe Systèmes d'Informations Documentaires (SID) dont il est co-responsable depuis 2001. Il est membre du Groupe de Recherche en Communication Écrite depuis 1998, responsable du secrétariat permanent de la SDN depuis 2004 et également animateur du groupe de recherche SCATI du GDR ISIS depuis 2003.



