

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

Using the Fourier Mellin transform for multi-oriented and multi-scaled patterns recognition: application to automatic analysis of technical documents

par S. ADAM ^{*,***}, J.-M. OGIER ^{*}, C. CARIOU ^{**}, R. MULLOT ^{*}, J. GARDES ^{***}, Y. LECOURTIER ^{*}

^{*} Laboratoire PSI, Ia3I, Université de Rouen, 76821 Mont Saint Aignan, France.
Tél. 02 35 14 68 75 – Fax 02 35 14 66 18
{Sebastien.Adam, Jean-Marc.Ogier}@univ-rouen.fr

^{**} LASTI/Groupe Image, ENSSAT Lannion, Université de Rennes I, 6, rue de Kerampont, B.P. 447, 22305 Lannion, France.
Tél. 02 96 46 66 39 – Fax 02 96 46 66 75
{Claude.Cariou@enssat.fr}

^{***} France Télécom Recherche et Développement, DMI/GRI, Technopôle Anticipa, 2, avenue Pierre Marzin, 22307 Lannion cedex, France.
Tél. 02 96 05 33 23 – Fax 02 96 05 32 86

résumé et mots clés

Dans cet article, nous proposons une méthodologie originale permettant la détection et la reconnaissance de caractères multi-orientés et multi-échelles. Les supports sur lesquels la méthode est appliquée sont des documents techniques représentant le réseau de l'opérateur de télécommunication français France Télécom. La technique adoptée, basée sur la transformation de Fourier-Mellin (TFM), est intégrée dans une stratégie globale permettant la résolution de situations ambiguës, par intégration d'informations contextuelles. La stratégie appliquée pour résoudre ce problème de reconnaissance de caractères et symboles multi-orientés et multi-échelles peut être divisée en deux étapes. La première réside dans l'extraction d'un ensemble de descripteurs invariants pour chacune des formes isolées de la couche « caractères » identifiée à partir d'un extracteur de composantes connexes. La seconde étape, basée sur un processus de filtrage, consiste à détecter et reconnaître les formes connectées au réseau ou à d'autres formes. Les résultats de l'application de cette technique sont très encourageants puisque le taux de classification atteint d'excellents niveaux en comparaison avec les techniques classiques de la littérature.

Reconnaissance de caractères, extraction de primitives et classification, analyse de document, transformée de Fourier-Mellin.

abstract and key words

In this paper, we propose an original methodology which allows the detection and the recognition of multi-oriented and multi-scaled patterns. The supports on which the method is applied are technical documents representing the network of the French Telephone operator *France Telecom*. The adopted technique, based on the Fourier-Mellin Transform (FMT) is integrated in a global strategy that solves ambiguous situations, through the providing of contextual information. The strategy which is applied to solve the character and symbol classification problem can be divided into two stages. The first one consists in computing a set of invariant descriptors for each isolated pattern belonging to a characters layer detected thanks to a connected components extractor. The second stage, based on a filtering scheme, consists in detecting and recognising the shapes which are either interconnected or connected to any other object. The results of the application of this technique are very encouraging since the classification rate reaches excellent scores in comparison with classical techniques.

Character recognition, primitives extraction and classification, document analysis, Fourier-Mellin transform

1. introduction

Les systèmes d'information représentent aujourd'hui un enjeu important pour la maîtrise de la mémoire et de l'expertise d'une entreprise, et ce, quel que soit son domaine d'activité. Dans le cadre de l'alimentation de ces systèmes, les processus de rétro-conversion de la documentation papier tiennent une place prépondérante, permettant l'intégration de données sémantiquement exploitables à partir de supports papier multi-sources et multi-rédacteurs.

Cette constatation explique le développement actuel de nombreuses études concernant l'analyse automatique de documents d'ingénierie, dont le but est de faciliter un stockage numérique des informations présentes sur la version papier [BOATTO][VAXIVIERE][JOSEPH][OGIER-2].

Dans cet article, nous concentrons notre attention sur une technique originale permettant la reconnaissance de formes multi-orientées et multi-échelles. Le problème est illustré dans le cadre d'une application particulière, développée par France Télécom sous le nom de projet NATALI (Numérisation et Acquisition pour le Traitement Automatique et la Lecture de l'Image de l'Itinéraire), dont le but est d'interpréter les plans techniques décrivant l'itinéraire des câbles téléphoniques et la répartition des diverses infrastructures sur le territoire français. Les concepts présentés dans ce papier sont les résultats d'une collaboration entre France Télécom, à travers sa branche Recherche et Développement (FT R&D) et les laboratoires PSI et LASTI.

Le papier sera organisé comme suit ; dans la seconde partie, nous décrirons les problématiques associées au traitement des données textuelles dans un document technique, les stratégies

généralement adoptées dans la littérature et nous proposerons un état de l'art sur les techniques de reconnaissance de formes basées sur des descripteurs géométriques invariants. A partir des conclusions de cette partie, nous présenterons alors, dans la troisième partie, la stratégie que nous utilisons pour reconnaître les caractères et symboles multi-orientés et multi-échelles présents sur la documentation papier France Télécom. Cette partie sera alors suivie par une présentation de l'implémentation de notre approche, par une discussion des résultats obtenus et des améliorations possibles. Enfin, dans la dernière partie, une conclusion du papier sera donnée, et les différentes perspectives qu'ouvre ce travail seront proposées.

2. traitement des informations textuelles

L'analyse et la gestion des informations textuelles est fondamentale dans le cadre de la lecture d'un document, cette couche d'informations étant sémantiquement très riche. Chacune des informations textuelles représente une part d'information exploitable par le lecteur du plan (cotations sur les schémas mécaniques, toponyme sur une carte routière, etc), généralement en adéquation avec la symbologie adoptée sur le schéma. La connaissance précise de cette information représente donc un enjeu important pour les systèmes d'interprétation de documents.

La littérature et les produits commerciaux font aujourd'hui une large place aux systèmes de reconnaissance optique des caractères (*Optical Character Recognition*) dans un cadre classique :

caractères et alignements horizontaux, fontes connues. Cependant, la généralisation à d'autres orientations et d'autres alignements des caractères réduit d'autant le nombre de propositions. Il est même possible de dire qu'aujourd'hui, il n'existe pas de produits industriellement fiables permettant de reconnaître les caractères acceptant des variabilités intrinsèques. Or, le contexte de la reconnaissance des éléments textuels présent sur des documents techniques fait référence à ce type d'outils. Dans la plupart des cas, les caractères sont en effet écrits manuellement, avec contraintes, et avec une orientation relative à celle de l'objet décrit. Ainsi, le nom d'une route suivra l'orientation de cette dernière.

Concernant cette problématique du traitement automatique des informations textuelles, lorsque celles-ci sont facilement « séparables » sur le document (i.e. lorsque les caractères, symboles ou objets linéaires sont sur des couches bien déconnectées les unes des autres), différentes techniques sont généralement utilisées de manière séquentielle. Celles-ci concernent la segmentation des caractères, leur reconnaissance individuelle, puis leur regroupement en mots ou toponymes. Ces différents points sont synthétisés dans la partie 2.1 (approche classique), en regard avec la littérature actuelle.

Cependant, un « verrou méthodologique » important demeure au sujet de la reconnaissance des informations textuelles qui sont inter-connectées ou connectées à d'autres objets. Sur ce point difficile, une synthèse bibliographique est également proposée dans la partie 2.2 (les données textuelles connectées).

2.1. approche classique

D'un point de vue système, le traitement des éléments textuels est réalisé classiquement en trois phases successives, que nous décrivons ici.

2.1.a. la segmentation des caractères

Cette étape vise à isoler les caractères des autres éléments du document. La plupart des auteurs proposent de détecter les caractères en isolant les composantes connexes de taille prédéfinie [TRIER-3], [FLETCHER], [LU], [SHIMOSUTSHI-1], [LAI], [LANGRANA]. Notons ici le travail original de Dori [DORI-1], [WENYIN] qui propose une segmentation permettant d'effectuer ce filtrage sur la base de parallélogrammes inclinés (*Charbox*) prenant en compte l'orientation du caractère. Ce processus permet de filtrer les candidats en fonction des dimensions précises en s'affranchissant de l'orientation du caractère. Ce principe est de plus un indicateur intéressant pour rechercher les groupes de caractères (*Textbox*). Cependant, ces traitements n'acceptent que peu de variabilité et sont applicables dans un cadre strict de normalisation de l'écriture (principalement des schémas mécaniques). A partir de cette segmentation des élé-

ments textuels, des algorithmes de regroupement sans reconnaissance de caractères sont mis en œuvre afin de construire les chaînes de caractères. Une référence dans le domaine est l'article de Fletcher [FLETCHER] qui propose de déterminer ces chaînes suivant un ensemble d'heuristiques et l'utilisation de la transformation de Hough pour déterminer les alignements. L'algorithme requiert trois composantes connexes colinéaires pour valider une zone de « texte » et peut, dans certaines conditions, valider des lignes pointillées en « texte ». Plus récemment, Lu [LU] propose une méthodologie de regroupement de « nouvelles composantes connexes » (NCC) basée sur un ensemble d'heuristiques (4 seuils prédéfinis). L'orientation du texte est estimée par minimisation du rectangle englobant par rotations successives de l'image des contours de 10° .

2.1.b. la reconnaissance des caractères

La reconnaissance de formes répondant aux contraintes d'invariance vis à vis de transformations telles que les similitudes est un point essentiel dans la conception de systèmes fiables pour l'interprétation de documents techniques. Dans ce champ de recherche, on peut distinguer trois principales catégories d'approches :

- La 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 d'OCR. Cependant, les méthodes basées sur une telle stratégie ne sont pas fréquemment utilisées à cause du manque de méthodes permettant un calcul fiable de l'orientation d'une forme. De plus, des distorsions dues à des erreurs d'échantillonnage apparaissent lors des transformations géométriques nécessaires à la normalisation de la forme.
- Une seconde approche classique consiste à utiliser un classifieur neuronal prenant en entrée l'image de la forme et qui rend lui-même, lors de sa phase d'apprentissage, le problème invariant aux transformations [FUKUMI]. Les problèmes posés par ce type de stratégie reposent sur la constitution de bases d'apprentissage représentative des différentes configurations susceptibles d'être rencontrées en phase de tests.
- La dernière approche, qui est probablement la plus utilisée, consiste à extraire de la forme un ensemble de descripteurs invariants aux transformations. Une excellente description de l'état de l'art dans ce domaine peut être trouvée dans un article de Trier [TRIER-1]. Généralement, on peut noter que les vecteurs utilisés pour décrire les formes indépendamment de leur position, de leur taille et de leur orientation peuvent être séparés en deux catégories:

a) Descripteurs basés sur l'aspect global de la forme

De nombreuses caractéristiques peuvent être utilisées pour décrire l'aspect global d'une forme. Depuis les travaux de Hu en

1961, les moments invariants [HU], qui sont basés sur des combinaisons de moments réguliers, ont été très fréquemment utilisés [ROTHE]. Parmi ceux-ci, on peut citer les moments de Zernike [TEAGUE] [KHOTANZAD-1] [KHOTANZAD-2] [LIAO] qui restent une référence dans le domaine, les pseudo-moments de Zernike [TEAGUE], les moments de Bamieh [BAMIEH], ou les moments de Legendre [CHEN].

Ces moments invariants, qui peuvent être extraits d'une image binaire ou d'une image en niveaux de gris, offrent généralement des propriétés de reconstructibilité, ce qui permet d'assurer que les primitives extraites contiennent la plus grande partie de l'information incluse dans la forme étudiée. De bonnes études comparatives concernant ces moments invariants peuvent être trouvées dans [TEH] et [BELKASIM], chacune d'elles soulignant la supériorité de moments de Zernike quant aux performances de reconnaissance.

Parmi ces résultats, Belkasim [BELKASIM] conclut que les moments invariants normalisés de Zernike d'ordre 6 donnent les meilleures performances pour la reconnaissance : 94,9 % de reconnaissance, 5 % de confusion en utilisant un classifieur de type K Plus Proches Voisins (KPPV) sur une base de 320 chiffres répartis dans 10 classes. Notons qu'une transformation mathématique est indispensable afin de rendre les moments de Zernike invariants à la rotation. Une autre comparaison intéressante entre les moments de Legendre, Zernike et pseudo-Zernike est proposée dans [BAILEY]. Cette étude montre que les moments de Zernike d'ordre 7 associés à un classifieur KPPV obtient les meilleurs résultats de reconnaissance, de l'ordre de 91,7% sur une base de caractères

manuscrits. Ces taux, même s'ils présentent un intérêt certain, ne peuvent cependant être validés que sur une base commune, base qui n'existe pas aujourd'hui dans le cadre de l'interprétation de documents techniques.

Néanmoins, ces études prouvent également que les approches à base de moments sont sensibles au bruit et qu'elles sont coûteuses en terme de temps de calcul, et ce, malgré les nombreuses méthodes d'optimisation et/ou de réduction de complexité qui existent dans la littérature [DAI] [LIN-1].

b) Descripteurs basés sur une approche locale

A l'inverse des approches présentées ci-dessus, une description invariante géométrique 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] ou des descripteurs elliptiques de Fourier [LIN-2]. Taxt [TAXT] a mené une étude comparative entre ces descripteurs qui montre leur intérêt potentiel, surtout en ce qui concerne leur simplicité et leur robustesse.

En particulier, Taxt [TAXT] propose de retenir les moments elliptiques de Kuhl [KUHLE] lorsque l'orientation du caractère est connue. Ce type de technique est utilisé également par Trier [TRIER-3] pour la reconnaissance des caractères appliquée à

des cartes hydrographiques. Ce dernier annonce un score de 78 % de reconnaissance pour 2,3 % de confusion sur une base de 1760 caractères manuscrits avec contraintes. Par ailleurs, des primitives invariantes structurelles peuvent aussi être extraites de caractères ou de leur squelette [BOATTO] [SHIMOTSUJI-1] [SHIMOTSUJI-2]. On peut par exemple prendre en compte le nombre d'occlusions, le nombre de jonctions en T ou en X, le nombre de coudes, etc. Cependant, il a été montré que de telles caractéristiques utilisées seules ne conduisaient pas à des systèmes de reconnaissance robustes [MORI]. Les primitives circulaires qui sont, par définition, bien adaptées à la reconnaissance invariante à la rotation, ont été utilisées dans [KITA]. Celles-ci sont basées sur l'analyse de la forme par l'intermédiaire d'un ensemble de cercles. En particulier, Lefrère [LEFRERE] propose un ensemble de sondes circulaires permettant de décrire la forme à partir de son centre de gravité. Ce type d'approche reste néanmoins très sensible aux variabilités et à la présence de bruit. Une étude comparative, disponible dans [DUDANI], montre que ce type de primitives circulaires permet l'obtention de meilleurs résultats que les moments de Hu.

2.1.c. constitution des mots – vérification

La dernière étape concernant la reconnaissance des éléments textuels porte sur la constitution des mots au sens des algorithmes proposés par Fletcher, Lu ou Dori. Des heuristiques sont parfois nécessaires afin de détecter des ruptures dans les chaînes de caractères. Notons ici le travail Deseilligny [DESEILLIGNY-1], [DESEILLIGNY-2] qui propose une méthodologie de construction de chaînes basée sur un algorithme de programmation dynamique permettant de construire le chaînage optimal en intégrant des contraintes de cohérence d'orientation des chaînes. A la suite de cette construction des chaînes (mots), des vérifications syntaxiques (contenu de la chaîne conforme à une grammaire interne) ou externe (utilisation d'une base de référence), améliore largement la confiance pouvant être donnée à la chaîne reconnue.

Le premier outil présenté dans cet article porte sur le point 2 de ce système, c'est-à-dire la reconnaissance de caractères isolés (Fig. 1.b) en référence à un ensemble de caractères présents dans une base et constituant des prototypes connus.

2.2. les données textuelles connectées

Une seconde problématique associée au traitement des données textuelles porte sur la localisation et la reconnaissance des caractères connectés entre eux ou connectés à d'autres éléments du document (lignes, symboles, ...). Les méthodes de séparation du texte s'appuient principalement sur les composantes connexes extraites. Ainsi, un caractère connecté à une ligne ne pourra être localisé simplement par une analyse des tailles des composantes connexes. De même, deux ou trois caractères

connectés entre eux constitueront une composante connexe de taille supérieure à celle des caractères isolés. Sur ce point, la littérature est relativement peu fournie, illustrant bien la difficulté du problème.

Le traitement des caractères interconnectés est certainement la problématique la moins complexe à traiter. En effet, il est possible de localiser relativement simplement ces regroupements sur la base d'une analyse des dimensions des composantes connexes, puis d'appliquer des techniques classiques de segmentation de caractères avec ou sans redressement de la forme. Casey [CASEY] propose une synthèse des méthodes de segmentation de caractères lorsque l'orientation du texte est horizontale. Ces techniques de segmentation ne sont pas applicables lorsque les données textuelles sont connectées aux autres éléments du document : lignes, textures, etc. La localisation de ces caractères dans une grande composante connexe (une route par exemple) est beaucoup plus complexe, les éléments caractérisant le ou les caractères étant difficilement identifiables parmi

les autres éléments de la route. Pour séparer ces éléments textuels, certaines équipes ont tenté de mettre en œuvre des algorithmes de traitement d'images basés sur des traitements morphologiques des formes [LU] ou l'adaptation locale du seuil de binarisation sur l'image en niveau de gris [TRIER-2], [TRIER-3]. Des techniques plus structurales consistent à traiter directement l'image vecteur afin d'extraire des regroupements de segments pouvant former des caractères [SHIMOTSUJI-2] ou ne correspondant pas au modèle de lignes et comportant des éléments textuels en proximité [NAGY]. Boatto [BOATTO-2] propose de détecter des caractères connectés sur la base d'une analyse des longueurs de plage (*run length*). À chaque nœud, une analyse locale est amorcée pour vérifier l'hypothèse caractère ou l'hypothèse ligne. Cette analyse est dirigée par un ensemble de règles intégrant des heuristiques. Si le caractère est détecté, de nouvelles règles de reconstitution sont mises en œuvre sur des bases topologiques et géométriques. Ces règles sont adaptées aux documents du cadastre italien comme le sou-

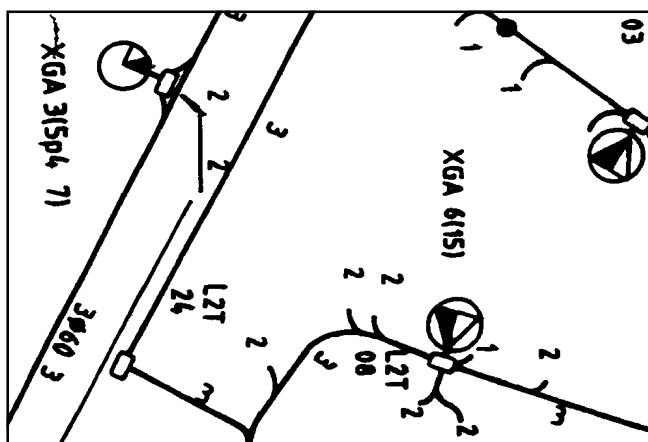


Figure 1a. – Exemple d'image originale.

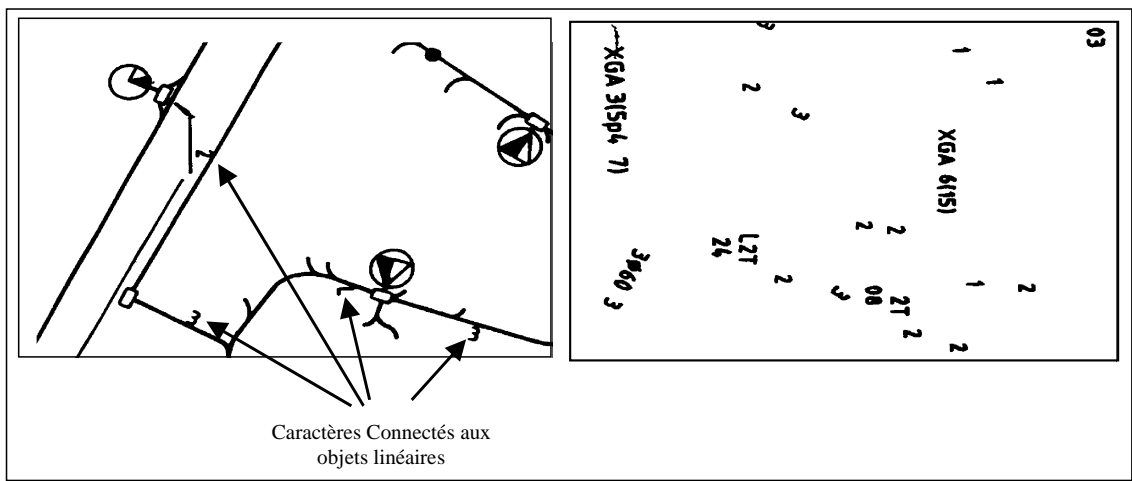


Figure 1b. – Caractères connectés au réseau.

Figure 1c. – Caractères isolés.

ligne les auteurs, et sont difficilement généralisables à d'autres types de documents.

Une approche intéressante est proposée par Dori [DORI-2], [LIU], [DORI-1]. A partir de l'image vectorisée, un processus récursif de constitution des caractères est amorcé sur les segments pouvant potentiellement constituer un caractère. Ce processus tente de regrouper les segments courts connexes de façon itérative jusqu'à la constitution d'une forme rentrant dans le parallélogramme de taille connue (*Charbox*), caractérisant la taille la plus précise possible des caractères. L'association de *charboxes* permet de valider la présence d'une chaîne de caractères (*Textbox*). Ce processus intéressant s'affranchit des connexités entre les caractères et avec les lignes, mais n'autorise que peu de variabilité sur la taille des caractères.

Enfin, Deseilligny [DESEILLIGNY-1] amorce une procédure de recherche de caractères connectés, lors du chaînage des caractères, de façon systématique de part et d'autre des extrémités des chaînes, et entre un couple d'hypothèses consécutives si l'espace entre deux caractères est suffisant. Dans ce cas, une recherche d'un caractère de la même police, à la même échelle, dans le voisinage et suivant la même orientation que la chaîne traitée est réalisée. Cette recherche s'effectue par appariements de primitives calculées sur les masques des matrices des caractères de références, et la zone de l'image où les caractères sont recherchés après avoir normalisé la zone de recherche. Afin de simplifier cette mise en correspondance, seul un ensemble de points caractéristiques du contour constitue les primitives à apparier. Cette technique relativement robuste s'adapte bien au traitement des chaînes régulières comportant plusieurs caractères sans variabilité excessive. Elle ne peut s'appliquer aux caractères isolés connectés aux lignes.

Le second outil présenté dans cet article propose l'utilisation de la transformée de Fourier-Mellin dans un mode permettant une segmentation et une reconnaissance des caractères connectés à d'autres éléments du plan (Fig. 1.b).

2.3. bilan concernant les problèmes de reconnaissance

Dans le domaine de la reconnaissance de formes invariante à différentes transformations géométriques, la plupart des chercheurs [TRIER-1] s'accordent pour dire que l'étape d'extraction de primitives est probablement le facteur le plus important. Un grand nombre de types de primitives est rapporté dans la littérature. Comme nous l'avons vu précédemment, certaines sont basées sur l'aspect global de la forme (les moments par exemple) alors que d'autres cherchent à sélectionner des points particuliers sur celle-ci pour la décrire (les primitives circulaires par exemple).

Etant donné le grand nombre de méthodes existantes, on pourrait prétendre qu'il n'est pas nécessaire de développer de nou-

veaux jeux d'invariants. Néanmoins, l'étude comparative menée dans [BELKASIM] montre que les primitives présentées ci-dessus ne sont pas parfaites quant aux performances de reconnaissance, surtout lorsque les images sont bruitées. De plus, cette analyse de l'état de l'art met également en évidence la difficulté de la mise en place d'une stratégie générique pour l'analyse de caractères/symboles connectés, ce qui constitue une contrainte majeure pour notre application.

3. reconnaissance de formes multi-orientées et multi-échelles

La stratégie de reconnaissance que nous proposons ici couvre les deux contraintes énoncées plus haut, c'est-à-dire l'invariance par rapport à la composition d'une rotation et d'une homothétie (c'est-à-dire une *similitude*), et la robustesse vis-à-vis des formes interconnectées. Plus précisément, nous utilisons la Transformée de Fourier-Mellin, dont les propriétés sont très intéressantes au regard de notre application. La technique ici développée est à l'intersection de nombreux travaux, dont ceux de Ghorbel [GHORBEL] et Derrode [DERRODE-1] [DERRODE-2], de Ravichandran et Trivedi [RAVICHANDRAN], et d'autres chercheurs travaillant dans le domaine de la reconnaissance de formes par calcul optique [ARSENAULT] [SHENG].

Dans un premier temps, nous rappellerons la définition de la transformée de Fourier-Mellin (TFM). Nous présenterons ensuite un ensemble complet et stable d'invariants aux similitudes positives, issu des travaux de Ghorbel. Puis nous discuterons les aspects numériques du calcul de la TFM sur des images discrètes, et montrerons quelques résultats expérimentaux de cette technique.

3.1. la transformée de Fourier-Mellin (TFM)

Soit $f(r, \theta)$ une fonction réelle et positive (la forme) définie en coordonnées polaires. La TFM de cette fonction est définie par la relation suivante :

$$M_f(q, s) = \int_{r=0}^{\infty} \int_{\theta=0}^{2\pi} r^{s-1} \exp(-iq\theta) f(r, \theta) dr d\theta \quad (1)$$

avec $q \in \mathbb{Z}$, $s = \sigma + iv \in \mathbb{C}$

Dans cette expression, i désigne l'unité imaginaire. La TFM est donc la combinaison de la représentation par coefficients de Fourier d'une fonction 2π -périodique (pour la variable angulaire θ) et de la transformée de Mellin (pour la variable radiale r).

La TFM existe pour les valeurs de (q, s) pour lesquelles la double intégrale (1) est définie. En pratique, pour une fonction $f(r, \theta)$ positive donnée, la convergence est obtenue pour tout q , mais pas pour tout $s = \sigma + iv$: en effet, si $f(r, \theta)$ se comporte au voisinage de l'origine comme Kr^λ , alors la TFM ne sera définie qu'à condition que $\sigma > -\lambda$. En général, le comportement de la forme au voisinage de l'origine correspond à des valeurs de λ proches de 0, ce qui implique souvent l'emploi d'une valeur de σ constante et comprise dans l'intervalle [0.5, 1.5]. La TFM est alors calculée pour s décrivant le contour de Bromwich $]\sigma - i\infty, \sigma + i\infty[$.

Notons qu'il existe une définition plus restreinte de la TFM, correspondant à $\sigma = 0$, décrite dans [GHORBEL][DERRODE-1]. Cette autre définition trouve son origine dans la théorie des groupes et l'analyse harmonique. En effet, en élargissant la notion de transformée de Fourier aux groupes de transformation topologiques, compacts ou localement compact, et commutatifs, Ghorbel [GHORBEL] a montré que la transformée de Fourier sur le groupe des similitudes conduisait à cette autre définition de la TFM, que l'on retrouve par ailleurs dans [WOOD]. Afin de pallier aux problèmes de divergence de (1) pour $\sigma = 0$, l'auteur a proposé la définition d'un prolongement analytique de la TFM (TFMA), ce qui revient tout simplement à l'écriture de l'éq. (1) avec $\sigma > 0$, et assure ainsi la convergence de la TFM, en particulier pour les images numériques. Pour résumer, on peut dire que la TFM avec $\sigma = 0$ est à la TFM de l'éq. (1) ce qu'est la transformée de Fourier à la transformée de Laplace d'un signal continu.

La fonction $f(r, \theta)$ peut être reconstruite à partir de la TFM, en utilisant les relations suivantes :

$$f(r, \theta) = \frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} \sum_{q \in \mathbb{Z}} r^{-s} \exp(iq\theta) M_f(q, s) ds$$

$$= \frac{1}{2\pi} \int_{-\infty}^{+\infty} \sum_{q \in \mathbb{Z}} r^{-\sigma-iv} \exp(iq\theta) M_f(q, s = \sigma + iv) dv$$
(2)

Cette relation définit la TFM inverse.

3.2. invariants issus de la TFM

La construction de descripteurs de formes invariants par rotation et homothétie s'appuie sur l'application du théorème du retard, que l'on retrouve dans toute l'analyse de Fourier. Soit $g(r, \theta) = f(\alpha r, \theta + \beta)$ une version modifiée en orientation et en échelle de $f(r, \theta)$, on peut montrer la relation suivante :

$$M_g(q, s) = \alpha^{-s} \exp(iq\beta) M_f(q, s) \quad (3)$$

En prenant le module des termes de l'équation (3), il est possible de déduire des descripteurs qui sont invariants aux changements de rotation de la forme mais pas aux changements d'échelle.

Pour obtenir l'invariance au changement d'échelle, il est néanmoins possible d'utiliser l'ensemble de descripteurs suivants :

$$I_f(q, s) = |M_f(q, s)| [M_f(0, \sigma)]^{-1} \quad (4)$$

Cet ensemble d'invariants fournit une représentation simple des formes. Cependant, il ne vérifie pas la propriété de complétude, c'est-à-dire qu'il n'existe pas de bijection entre les représentations duales d'une forme, à cause de la perte de l'information de phase.

Dans [GHORBEL], l'auteur propose un ensemble de primitives invariants aux changements d'échelle et d'orientation ; celui-ci repose sur l'expression suivante, généralisée ici à une valeur de σ quelconque :

$$I_f(q, s) = M_f(q, s) [M_f(0, \sigma)]^{-\frac{s}{\sigma}} [M_f(1, \sigma)]^{-q} |M_f(1, \sigma)|^q \quad (5)$$

Partant de ces invariants, si $g(r, \theta) = f(\alpha r, \theta + \beta)$, il peut être montré facilement que $I_g(q, s) = I_f(q, s)$.

Les deux propriétés importantes de ces primitives sont les suivantes :

(i) *Complétude* : à partir des valeurs de $\{I_f(q, s)\}$, $M_f(0, \sigma)$ et $M_f(1, \sigma)$ il est possible de revenir à l'ensemble des coefficients de FM $\{M_f(q, s)\}$, et donc reconstruire $f(r, \theta)$, en utilisant la TFM inverse donnée par la relation (2).

(ii) *Convergence* : dans [GHORBEL], il est prouvé, en supposant que $\{M_f(q, s)\}$ est un ensemble convergent, qu'il existe

$$\ll x \in \mathbb{R}, x > 1 \gg \text{ tel que } \left(\int_{\mathbb{R}} \sum_{\mathbb{Z}} |I_f(q, s)|^x dv \right)^{1/x} < +\infty.$$

Une conséquence importante des propriétés de complétude et de convergence de cet ensemble d'invariants réside dans l'existence d'une métrique dans l'espace de représentation des formes.

3.3. application aux images numériques

De nombreuses difficultés surviennent lorsque l'on essaie d'appliquer la TFM à des images numériques. La première vient de la nécessité d'échantillonner, suivant la variable (continue) v , les fonctions $\{M_f(q, s)\}$. Ce problème – ouvert – peut admettre une solution simple en échantillonnant régulièrement cette variable, et en calculant $\{M_f(q, s)\}$ et $\{I_f(q, s)\}$ pour tout $v = p \in \mathbb{Z}$ [DERRODE-1].

Le second problème réside dans le choix nécessaire d'un point invariant et représentatif de la forme comme Centre de Développement (CdD) de la TFM. Un choix *a priori* raisonnable consiste à appliquer la TFM au centre de gravité de la forme. Pour les formes bien conditionnées (c'est-à-dire déconnectées de toute autre forme), le calcul des coordonnées du centre de gravité est facile et permet d'assurer que la forme est incluse dans un rayon donné, argument qui sera d'un grand intérêt dans la suite.

La troisième et principale difficulté est due à la structure bidimensionnelle, cartésienne et discrète de l'image numérique, sur laquelle on veut appliquer une transformation définie dans un repère polaire. Pour résoudre ce problème, deux options sont disponibles : l'interpolation de l'image en coordonnées polaires [DERRODE-1][GHORBEL] ; ou bien la convolution de l'image en coordonnées cartésiennes, au CdD d'une forme, avec un banc de filtres appropriés [RAVICHANDRAN].

Avec la première option, la TFM est obtenue grâce à la relation (1) appliquée au CdD de la forme considérée.

Cela suppose donc que l'image soit d'abord interpolée en coordonnées polaires pour couvrir la gamme entière de $r \in [0, r_{\max}]$ et $\theta \in]0, 2\pi[$ (r_{\max} étant le rayon maximal de la forme). Cette technique entraîne un coût de calcul qui peut être important, et peut aussi introduire de l'information redondante dans les données originales.

Avec la deuxième option, il n'est pas nécessaire d'interpoler la forme à partir de l'image. En effet, la version discrète de la TFM, calculée au CdD (k_0, l_0) peut être approchée de la façon suivante:

$$M_f(q, \sigma + ip) \approx \sum_{1 \leq (k^2 + l^2) \leq r_{\max}^2}^k \sum_l h_{p,q}(k, l) f(k_0 - k, l_0 - l) \quad (6)$$

avec

$$h_{q,p}(k, l) = \frac{\exp \left\{ i \left[\frac{p}{2} \ln(k^2 + l^2) - q \arctan \left(\frac{l}{k} \right) \right] \right\}}{(k^2 + l^2)^{1 - \frac{\sigma}{2}}} \quad (7)$$

Cela signifie que la convolution de la forme avec un filtre 2-D de réponse impulsionnelle finie $h_{q,p}(\cdot, \cdot)$ au CdD de la forme donne le coefficient de la TFM d'ordre $(q, \sigma + ip)$. Notons ici que tous les filtres, mis à part $h_{0,0}(\cdot, \cdot)$ sont à valeurs complexes. Il peut être facilement démontré que l'ensemble des filtres $h_{q,p}(\cdot, \cdot)$ est orthonormal en fonction du coefficient circulaire q mais pas en fonction du coefficient radial p , à cause du noyau de la transformée de Mellin, défini originellement pour une variable v continue. Une orthogonalisation complète des fonctions de base est néanmoins possible si des conditions spéciales sur la discrétisation de la variable v , et sur les valeurs de σ et r_{\max} sont respectées ; néanmoins, cette orthogonalisation n'a pas été effectuée dans ce travail.

Dans le cadre de notre application, du fait de sa simplicité, nous avons choisi la seconde option qui, de plus, offre l'avantage d'une implémentation parallélisable pour le calcul des coefficients de la TFM. Cet aspect pratique du calcul de la TFM fait d'ailleurs l'objet de recherches actuelles dans le laboratoire PSI. Dès lors que les coefficients de la TFM sont disponibles au CdD d'une forme donnée, l'Eq. (5) est utilisée pour générer des primitives invariantes aux changements d'échelle et d'orientation pour cette forme. Ces primitives sont toutes complexes, excep-

tion faite de $I_f(0, \sigma)$ qui vaut 1 et de $I_f(1, \sigma)$ qui est un réel pur. De plus, l'ensemble d'invariants vérifie la propriété de symétrie suivante:

$$I_f^*(q, s) = I_f(-q, s^*) \quad (8)$$

où * dénote le complexe conjugué. Cette propriété de symétrie hermitienne vis-à-vis des variables p et q est intéressante au regard du temps de calcul, car elle permet de restreindre le nombre de primitives non redondantes à l'ensemble :

$$A = \{(q, p) | (q=0; 0 \leq p \leq P) \cup (1 \leq q \leq Q; -P \leq p \leq P)\}$$

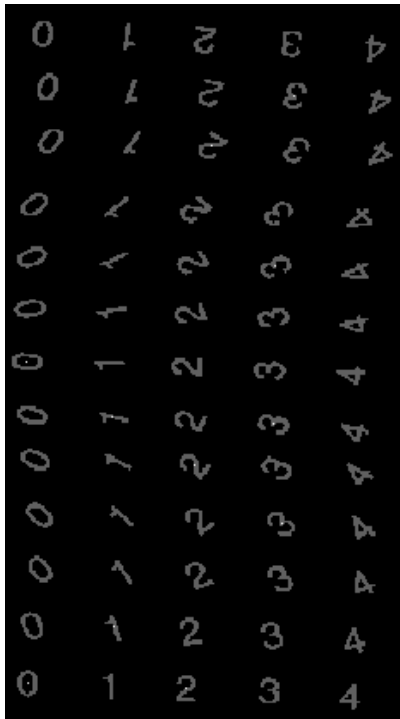
Partant de ces considérations, nous allons donner quelques illustrations des calculs de ces primitives invariantes, afin de mettre en évidence leur stabilité, ainsi que les paramètres importants dans leur estimation.

En terme d'illustration, un exemple de description invariante pour 5 caractères différents est donné en Figure 2. Elle correspond aux 13 chiffres multi-orientés (de 0 à 4) présentés en Figure 2-(a). Ces formes ont été analysées par le banc de filtres $\{h_{q,p}(\cdot, \cdot)\}$, avec $\sigma = 1$, $r_{\max} = 20$, $Q = 3$ et $P = 2$. Compte tenu des remarques précédentes, ces valeurs de Q et P conduisent à l'utilisation de 18 filtres, et engendrent un vecteur de 33 paramètres invariants réels et non redondants. Les mêmes formes ont aussi été analysées à un facteur d'échelle de 0.5, simplement obtenues par sous-échantillonnage de l'image originale de la Figure 2-(a), afin de mettre en évidence la stabilité des invariants vis-à-vis du changement d'échelle.

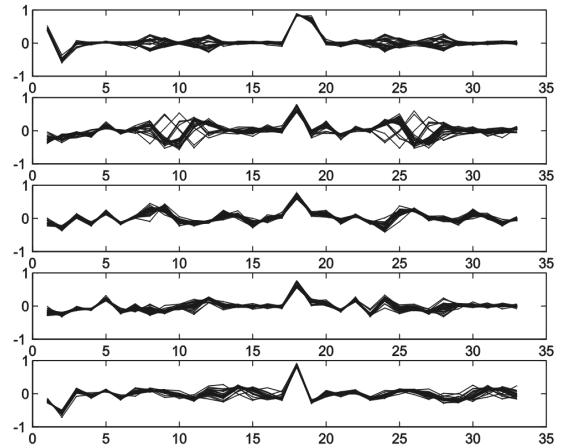
Certes, cette stratégie de sous-échantillonnage n'apparaît pas comme parfaitement représentative de la multi-résolution réellement rencontrée dans le cadre de l'analyse documentaire. Néanmoins, nos expérimentations mettent en évidence le fait que cette démarche d'évaluation met le dispositif dans des conditions plus difficiles que les conditions réelles puisque le sous-échantillonnage induit une perturbation des formes supérieure à celle que nous rencontrons dans la réalité, lors du traitement de formes de tailles variables.

Pour cette analyse de formes multi-échelles, nous avons donc délibérément opté pour la démarche du « qui peut le plus peut le moins ». Toutefois cette démarche, visant à démontrer la faisabilité du procédé, n'est pas la seule stratégie d'étude de la robustesse des invariants de Fourier Mellin au changement d'échelle. La réalité industrielle et la base de 20 000 caractères issus de réels documents France Télécom viendront étayer ce discours autour de la stabilité des invariants.

La Figure 2-(b) montre les primitives invariantes obtenues, chaque ligne correspondant aux invariants pour une forme donnée. Par exemple, la première ligne représente la superposition des 33 invariants pour chacune des 26 formes « 0 » disponibles en 13 orientations et 2 tailles différentes. Notons que sur cette figure, l'ordonnement des invariants n'a aucun sens particulier, et correspond à un balayage quelconque des valeurs de q et p . Toutefois, le pic en position 18 correspond au coefficient



(a) : un ensemble de 5 formes multi-orientées



(b) : Les invariants correspondants

Figure 2. – Résultats expérimentaux préliminaires.

(réel) $M_f(q = 0, s = \sigma)$. Sur la figure 3, nous illustrons la propriété de complétude des invariants, en montrant la reconstruction par TFM discrète inverse de quelques formes à partir d'une analyse basée sur un ensemble de 61 filtres ($Q = 5$ et $P = 5$), donnant 119 invariants non redondants par forme. Notons que la TFM discrète inverse, appliquée en coordonnées cartésiennes peut être approchée par la relation suivante :

$$f(k, l) \approx \frac{1}{2\pi} \sum_{q=-Q}^Q \sum_{p=-P}^P M_f(q, s = \sigma + ip) g_{q,p}(k, l) \quad (9)$$

avec

$$g_{q,p}(k, l) = (k^2 + l^2)^{1-\sigma} h_{q,p}^*(h, l) \quad (10)$$

$$= (k^2 + l^2)^{-\frac{\sigma}{2}} \exp \left(-i \left[\frac{p}{2} \ln(k^2 + l^2) - q \arctan \left(\frac{l}{k} \right) \right] \right)$$

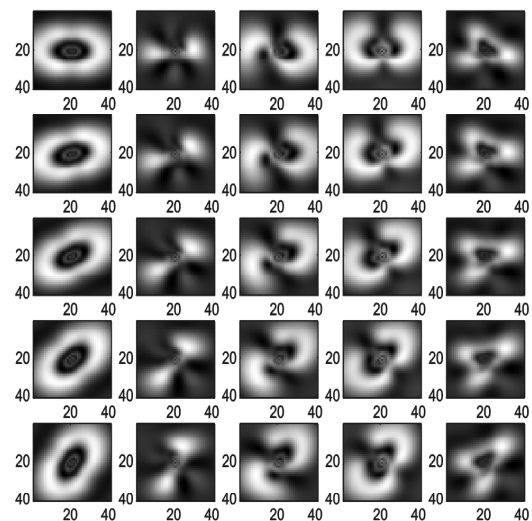


Figure 3. – Reconstruction des formes à partir des invariants.

4. application de la TFM dans le cas de l'interprétation de document

4.1. contexte

Comme nous avons pu le mettre en évidence dans l'analyse précédente, les propriétés de la TFM présentent des caractéristiques intéressantes pour notre problème. Toutes ces propriétés nous ont encouragés à utiliser cette approche dans le contexte du projet NATALI.

Afin d'illustrer l'implémentation de la TFM dans le cadre de notre problème d'analyse de document, nous proposons de présenter de manière séparée les deux modes de fonctionnement de cet outil de détection/reconnaissance de formes. Nous présentons donc dans une première partie l'implémentation de la TFM pour la reconnaissance des formes isolées, parfaitement segmentées, car non connectées à toute autre structure. Dans un second temps, nous développerons l'implémentation du « mode filtrage » pour la détection et la reconnaissance de formes connectées entre elles ou connectées aux parties linéaires du réseau. Nous présentons dans chacune de ces parties les tests significatifs que nous avons réalisés pour obtenir une évaluation fiable de la méthodologie présentée ci-dessus. La stratégie de la classification utilisée est basée sur le test de deux classifieurs différents : le *Learning Vector Quantization* (LVQ) et les K-Plus-Proches-Voisins (KPPV). Le choix de ces classifieurs a été motivé par deux arguments : d'abord, la technique des KPPV représente une bonne référence sur laquelle une évaluation peut être effectuée objectivement; ensuite, l'approche du LVQ est une technique intéressante du point de vue de la mise en œuvre, puisque dans notre contexte industriel les réseaux de neurones présentent la particularité d'être rapides en décision. Nous rappelons que les techniques de classification ici utilisées consistent en la détermination d'une distance entre le vecteur qui représente une forme inconnue et un ensemble de vecteurs de référence qui décrivent des formes déjà étiquetées. Dans toutes les expériences, le nombre de classes à identifier est de 51 (c'est-à-dire majuscules et minuscules de l'alphabet et chiffres de 0 à 9, en intégrant, par exemple, le fait que les lettres « n » et « u » appartiennent à la même classe, ces deux formes étant identiques, à une rotation de 180° près).

4.2. analyse des formes déconnectées

4.2.a. méthodologie

Cette première partie vise donc à présenter l'analyse et la reconnaissance des formes isolées sur le document, ne posant ainsi aucun problème de segmentation. La stratégie consiste à extraire

de l'image toutes les composantes connexes (c'est à dire les formes non connectées), opération réalisée grâce à un extracteur de composantes connexes. Ensuite, un calcul des descripteurs invariants est réalisé au centre de gravité de chacune des formes et le vecteur résultant de cette opération est introduit en entrée d'un classifieur qui permet l'obtention de la classe de la forme.

Afin d'avoir un point de vue objectif au sujet des primitives invariantes issues de la TFM, nous les avons comparées à d'autres techniques considérées comme étant excellentes dans la littérature. Tout d'abord, nous avons comparé les invariants de Fourier-Mellin avec les moments de Zernike, méthode de caractérisation hautement considérée si l'on en croit la littérature, également dans le domaine de la reconnaissance de caractères multi-orientés et multi-échelles [TEAGUE][KHOTANZAD]. Par ailleurs, notre comparaison s'est également appuyée sur une représentation à base de primitives circulaires, technique également fréquemment exploitée dans le cadre de la reconnaissance de forme multi-échelles et multi-orientées [LEFRERE]. Les détails concernant les techniques retenues et conditions de comparaison avec ces méthodes de caractérisation furent les suivantes :

- considérant le vecteur issu de la TFM, nous avons utilisé un jeu de 33 invariants, basé sur l'utilisation des filtres $\{h_{q,p}(\cdot, \cdot)\}$, avec $\sigma = 1$, $r_{\max} = 20$, $Q = 3$ et $P = 2$;
- considérant les moments de Zernike, nous avons appuyé notre caractérisation sur un jeu de 24 moments, issus des polynômes de Zernike pour des ordres allant de 3 à 6. Ces moments ont subi la technique de normalisation proposée par Belkasim [BELKASIM]. Le vecteur obtenu est alors de dimension 24, suivant les recommandations de l'auteur ;
- concernant les primitives circulaires, de nombreuses techniques sont également proposées dans la littérature. Nous appuyons notre analyse sur un jeu de 100 primitives circulaires, issu de la méthodologie très prometteuse proposée par [LEFRERE]. Cette méthode, similaire à celle proposée par Kita sous le nom de « CCD » (*Concentric Circular Description*) dans un article paru en 1992 [KITA], repose sur un ensemble de cercles concentriques de rayons proportionnels à la forme afin d'assurer l'invariance au changement d'échelle. Elle utilise comme vecteur caractéristique de la forme à reconnaître, un jeu de valeurs angulaires obtenues en analysant les angles d'intersection entre l'ensemble de cercles concentriques et la forme. Cette liste est obtenue en parcourant un cercle dans un sens donné. Lors de ce parcours, les écarts angulaires existants entre les points d'entrées et de sorties dans la forme sont mesurés et stockés dans la liste des valeurs angulaires. En fin de parcours, afin de rendre le problème invariant à la rotation, une permutation circulaire de la liste est effectuée en positionnant le secteur angulaire le plus important en tête de liste (cf. figure 4). Dans le cadre de cette étude, le jeu d'invariants est issu d'un ensemble de 10 cercles concentriques.

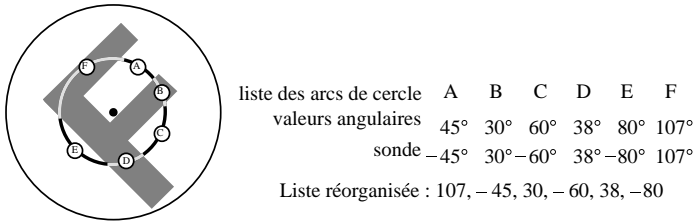


Figure 4. – Construction des listes caractéristiques selon [LEFRERE]

4.2.b. résultats

La stratégie que nous avons adoptée pour nos tests concernant les formes isolées peut être divisée en deux parties. Dans un premier temps, pour valider la méthodologie, nous avons appliqué notre technique sur une base de caractères « synthétiques », spécialement conçue à cet effet. Puis, dans une deuxième partie, nous avons considéré un ensemble de caractères issus de plans de France Télécom.

a) la base de caractères « synthétiques »

Nous présentons sur la figure 5 un échantillon d'image qui montre un ensemble de caractères issus d'une base de données conçue pour cette application. Cette base a été constituée à partir d'un logiciel de dessin permettant d'appliquer un pas de rotation programmable à des formes isolées. En ce qui concerne les tests qui suivent, nous avons utilisé un ensemble de 394 échantillons pour l'apprentissage et une base de test composée de 412 échantillons.

Les résultats obtenus en utilisant cette méthode, résumés dans le tableau 1 suivant, sont encourageants puisqu'ils entraînent un taux de classification correcte de 97,5 % avec la stratégie KPPV, (K = 1) et de 97,0 % avec la procédure LVQ. La matrice de confusion issue de ce procédé de classification est assez difficile à mettre en évidence dans le cadre de cet article, du fait du nombre important de classes analysées. Néanmoins, une analyse fine de la matrice de confusion montre que celles-ci sont

cohérentes puisqu'elles concernent des formes semblables telles que « m » et « E » ou « N » et « Z », lesquelles sont assez similaires, à une rotation de 90° près, pour les polices de caractères utilisées.

b) la base de données réelle

Comme nous l'avons expliqué ci-dessus, la deuxième expérience a consisté à tester la méthode sur un ensemble de caractères issus de la documentation France Télécom. Le problème est bien entendu beaucoup plus compliqué que dans le cas de caractères « synthétiques », puisque la variabilité des formes est beaucoup plus grande, et que les contours de la forme sont très bruités après l'étape de binarisation. Pour illustrer cet aspect « variabilité », nous présentons sur la figure 6 un exemple d'image traitée.

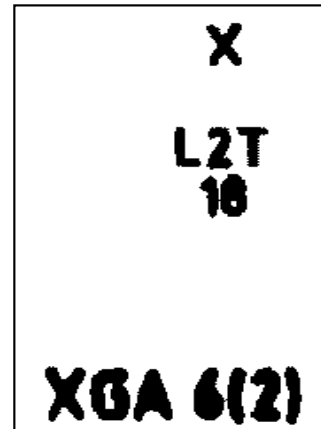


Figure 6. – un ensemble de caractères issus de documents techniques

La base d'apprentissage était composée de 8192 échantillons et la base de test de 11118 échantillons. Le taux de classification obtenu est assez bon, au regard de nos contraintes industrielles, puisque le taux de bonne reconnaissance atteint 95.8 % avec l'approche KPPV et 95.1 % avec le classifieur LVQ. Les confusions rencontrées sont de nouveau, pour la plupart, dues aux formes semblables telles que « B » et « 8 » (lesquelles sont très difficiles à distinguer à cause du bruit sur l'image), ou « 5 » et « S. » Ces cas particuliers ne remettent pas en cause la méthodologie générale puisque ce genre de problème peut être résolu grâce à l'introduction d'informations contextuelles, telles que l'orientation du mot dans lequel se trouve le caractère analysé, par exemple. Dans notre stratégie d'interprétation, cette information contextuelle sera fournie par l'approche système permettant, au travers de traitements complémentaires, de mesurer la pertinence des résultats obtenus et de valider ou rejeter des propositions faites par ce procédé de classification.

Les résultats de tous ces tests, qui mettent en évidence la supériorité des invariants de Fourier-Mellin sur les méthodes utilisées dans le cadre de notre comparaison sont présentés dans le tableau 1.

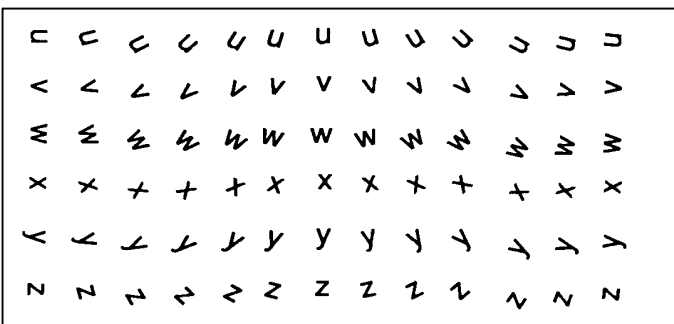


Figure 5. – un ensemble de caractères synthétiques

Tableau 1. – Résultats de Classification et Comparaison sur les bases de données Synthétique et Réelle

Classifieur	Base de données synthétiques		Base de données réelles	
LVQ	Fourier-Mellin	97,0 %	Fourier-Mellin	95,1 %
	Moments de Zernike	75,5 %	Moments de Zernike	83,9 %
	Primitives circulaires	73,6 %	Primitives circulaires	82,6 %
1-PPV	Fourier-Mellin	97,5 %	Fourier-Mellin	95,8 %
	Moments de Zernike	80,9 %	Moments de Zernike	83,5 %
	Primitives circulaires	77,5 %	Primitives circulaires	83,9 %

Ces résultats sont tout à fait intéressants dans le cadre de notre application industrielle vis-à-vis des seuils de rentabilité. Néanmoins, nos études actuelles portent sur l'analyse de la corrélation et de la redondance des invariants de Fourier-Mellin dans le but d'analyser l'opportunité de la réduction du vecteur de paramètres actuellement intégré dans la chaîne d'analyse. Une analyse en composantes principales à partir des invariants de Fourier-Mellin met en évidence la corrélation entre certains paramètres. Les études actuelles portent sur l'analyse de cette corrélation et sur l'éventualité de la réduction de certains paramètres. D'autres études portent par ailleurs sur la recherche de la combinaison optimale de descripteurs, basée sur l'utilisation des invariants de Fourier-Mellin et les descripteurs utilisés dans le cadre de notre analyse comparative. Une première étude, proposée dans [ADAM], montre l'intérêt de ces travaux.

4.3. analyse des formes connectées

4.3.a. méthodologie

La seconde stratégie a pour objectif de détecter et de reconnaître les caractères et symboles qui sont connectés les uns aux autres ou au tracé du réseau. Cette opération est exécutée grâce à la possibilité d'utiliser la TFM en mode filtrage. Plus précisément, l'image est d'abord convoluée par l'ensemble de filtres $h_{q,p}(\cdot, \cdot)$. Une procédure de localisation des pixels pour lesquels on obtient une réponse pré-spécifiée est ensuite déclenchée. Les pixels correspondants peuvent donc être assimilés au CdD d'une forme connue. Pour ce faire, un ensemble de descripteurs invariants est calculé pour chaque pixel de l'imagette analysée et les résultats obtenus sont introduits à l'entrée du classifieur. Une illustration de cette stratégie permettant de détecter/reconnaître les formes inter-connectées est donnée en figure 7. Rappelons que cette analyse par convolution est donc déclenchée unique-

ment sur des imagettes provenant de l'analyse des ambiguïtés issues de l'approche système. Dans le cadre de notre application, deux critères sont actuellement utilisés pour le déclenchement de la recherche de caractères connectés. Ainsi, une zone peut être considérée comme ambiguë lorsque la syntaxe de la chaîne de caractères reconstruite à l'issue de la phase de reconnaissance des caractères isolés ne correspond au modèle « métier », c'est-à-dire lorsque le mot reconstitué ne correspond pas à un mot du dictionnaire de la base de données accompagnant le document papier. Par ailleurs, au travers d'une analyse structurelle des informations extraites, par recherche d'isomorphisme exact ou inexact dans le graphe modèle, il est également possible de détecter des ambiguïtés correspondant potentiellement à des caractères connectés : à titre d'illustration, il est impossible de rencontrer un élément de réseau linéaire se terminant suivant des règles non prévues dans le cahier des charges du dessinateur du réseau. Dans le cadre de l'analyse des objets linéaires, la détection de ce type de structure se traduit dans notre processus d'analyse par un vecteur, sans nœud, incohérent vis-à-vis du modèle. Ce type de zone correspond à une ambiguïté au sein de laquelle la procédure de filtrage est déclenchée. Grâce à cette procédure de filtrage, la reconnaissance de formes collées devient dès lors possible. En fait, la bonne faculté de détection de cet opérateur est principalement due à la décroissance du module de la réponse impulsionnelle $h_{q,p}(\cdot, \cdot)$ lorsque la distance au centre de gravité augmente : dans la plupart des configurations où les caractères et symboles sont connectés les uns aux autres, ce problème se produit au voisinage de l'enveloppe externe de la forme uniquement. Dans de tels cas, la contribution d'une autre forme connectée sur le calcul appliqué au centre de gravité de la forme courante est minimale, et n'affecte pas trop le processus de la classification. Bien entendu, pour réduire le coût calculatoire, et prendre en considération la grande dimension des filtres, la convolution du banc de filtres avec l'imagette est exécutée par l'intermédiaire d'une transformée de Fourier 2-D rapide sur des régions d'intérêt émises par l'approche système.

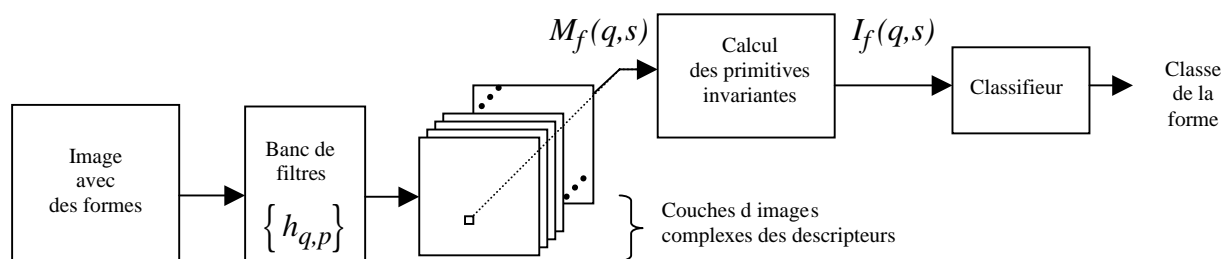


Figure 7. – Schéma pour l'analyse d'images avec des formes connectées.

4.3.b. résultats

Pour les formes connectées, la stratégie que nous avons appliquée était la suivante. Tout d'abord, un ensemble de 257 sous-images comprenant des caractères connectés a été sélectionné manuellement. Sur ces images, les formes collées pouvaient être des caractères connectés les uns aux autres ou des caractères connectés au réseau (exemple figure 8.a). Un nombre total de 408 caractères collés était présent sur l'ensemble des images.

Deux points ont alors été considérés dans ce test, à savoir détection et classification. Dans un premier temps, la détection est évaluée en fonction du fait que la technique a détecté ou n'a pas



Figure 8.a. – Caractères connectés au réseau

détecté une forme en un pixel particulier. Le taux de non-détection est encourageant puisque seules 7 des 408 échantillons n'ont pas été détectés comme caractères. Dans un second temps, nous avons considéré le taux de fausse détection en considérant le nombre de pixels pour lesquels une forme a été détectée alors qu'il n'y en avait pas sur l'image. Le résultat obtenu est bon puisque seule une forme est faussement détectée parmi les 257 images.

Le deuxième point de notre évaluation concerne le taux de classification lorsqu'une forme a été détectée pour un pixel donné. Les résultats obtenus sont là encore excellents puisque le taux de classification correcte atteint 83%. Comme dans le cas des formes synthétiques, les principales erreurs de classification sont essentiellement dues aux formes semblables. Cependant, dans certains cas particuliers, malgré la faible influence des formes connectées sur la description des formes à reconnaître, certaines erreurs semblent dues à une légère perturbation du calcul des invariants. Sur la figure 8.b, nous présentons une illustration de l'application du mode filtrage permettant la détection (figure 8.b, image centrale) et la reconnaissance d'un caractère « 3 » (figure 8.b, image de gauche) initialement connecté au réseau.

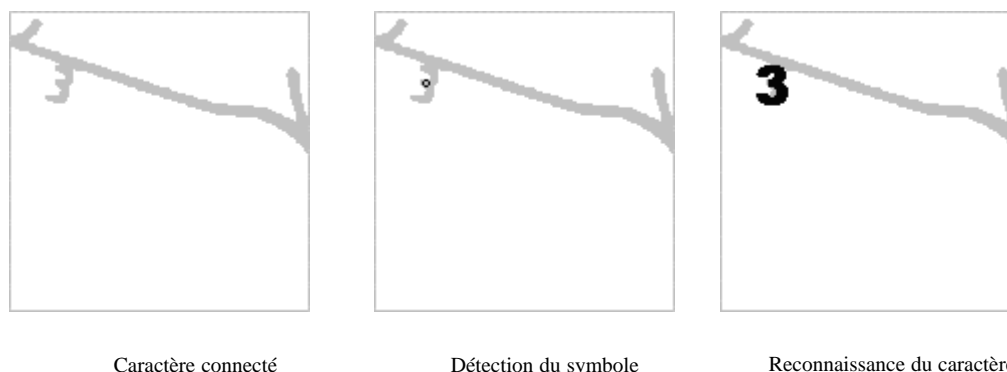


Figure 8. – Schéma pour l'analyse d'images avec des formes connectées.

4.4. mise en concurrence des deux stratégies de classification

Afin de mettre en évidence la fiabilité de ce procédé de filtrage, au regard de la connexion de différents objets, nous avons essayé de tester les deux stratégies sur un jeu de formes simples, constituées d'un « 3 » incliné, progressivement perturbé par l'introduction d'une barre verticale.

Sur les figures 9.a et 9.b, nous illustrons cette stratégie et mettons en évidence le résultat de la classification obtenu suivant chacune des deux méthodologies précédemment présentées. La figure 9.a illustre le résultat de la classification de la forme principale, en utilisant le procédé de reconnaissance supposant que la forme est parfaitement segmentée. La figure 9.b utilise le mode filtrage et met en évidence la supériorité de cet outil ainsi que sa stabilité au regard du déplacement du centre de gravité de la composante connexe.

3[0.83]	3[0.57]	3[0.61]	3[0.62]	3[0.62]	3[0.67]	3[0.70]	3[0.78]
3[0.80]	3[0.75]	3[0.74]	3[0.74]	3[0.69]	3[0.70]	3[0.67]	3[0.65]
3[0.69]	3[0.67]	3[0.59]	8[0.58]	3[0.61]	3[0.61]	8[0.65]	8[0.69]
3[0.51]	3[0.55]	3[0.57]	3[0.58]	0[0.54]	0[0.55]	0[0.55]	0[0.55]
0[0.57]	0[0.52]	0[0.50]	0[0.50]	8[0.33]			

Figure 9 (a). – Résultat de l'application de la TFM en mode « normal ».

3[0.92]	3[0.75]	3[0.76]	3[0.78]	3[0.80]	3[0.82]	3[0.83]	3[0.83]
3[0.83]	3[0.82]	3[0.79]	3[0.79]	3[0.77]	3[0.77]	3[0.76]	3[0.75]
3[0.71]	3[0.71]	3[0.71]	3[0.70]	3[0.67]	3[0.67]	8[0.78]	8[0.79]
3[0.58]	3[0.59]	3[0.64]	3[0.66]	3[0.69]	3[0.70]	3[0.73]	3[0.73]
3[0.69]	3[0.76]	3[0.80]	3[0.80]	3[0.92]			

Figure 9 (b). – Résultat de l'application de la TFM en mode « filtrage ».

L'ensemble de ces résultats de classification, tant du point de vue de l'analyse des formes isolées que du point de vue de la reconnaissance des caractères/symboles connectés, est très encourageant au regard des contraintes industrielles de notre projet.

5. conclusion et perspectives

Dans ce papier, nous avons proposé une méthodologie originale permettant la détection et la reconnaissance de formes multi-orientées et multi-échelles. Les supports sur lesquels la méthode a été appliquée sont des documents techniques qui représentent le réseau de l'opérateur téléphonique français France Télécom en surimpression sur des cartes urbaines. La technique adoptée, basée sur le transformée de Fourier-Mellin, est intégrée dans une stratégie globale d'interprétation qui résout les situations ambiguës en fournissant des informations contextuelles.

La stratégie qui est appliquée pour résoudre la reconnaissance des caractères et symboles peut être divisée en deux étapes. La première consiste à construire un vecteur de caractéristiques géométriques invariants pour chaque forme de la couche caractères issue de l'approche système. La deuxième étape, qui constitue probablement la plus grosse contribution de ce travail, consiste à détecter et reconnaître des formes connectées.

La technique est appliquée sur de petites images, appelées « zones d'intérêt », issues de l'approche système. Ces zones correspondent à des ambiguës détectées par le système en analysant la cohérences des objets reconstruits. Pour ces régions, une technique de filtrage détecte et reconnaît des formes présentes.

Les résultats de l'application de cette technique sont très encourageants puisque le taux de bonne classification atteint d'excellents scores, et puisqu'aucune information contextuelle, telle que l'orientation des toponymes ou les bases de données contenant des informations numériques, n'est pour l'instant intégrée dans le processus de reconnaissance.

Les perspectives de ce travail sont nombreuses et concernent différents points. En premier lieu, à propos de l'approche basée sur la transformation de Fourier-Mellin, une étude est actuellement menée pour optimiser la construction des vecteurs d'invariants. Plus précisément, nous pensons qu'il est nécessaire de tester différents choix de vecteurs d'invariants et de chercher à constituer un vecteur optimal, en essayant d'éviter la « malédiction de la dimensionnalité » qui survient dans beaucoup de problèmes de reconnaissance de formes statistiques. Cette optimisation pourra être effectuée par utilisation d'algorithmes génétiques, par exemple. Nos réflexions s'orientent aussi vers l'amélioration de la méthodologie par ajout de descripteurs complémentaires.

D'un autre point de vue, nous travaillons actuellement à l'intégration d'informations contextuelles dans l'approche système.

Par exemple, un ensemble de caractères alignés fournit de l'information à propos des toponymes, et par conséquent, pourrait guider le processus de la reconnaissance avec des informations d'orientation. De ce fait, le processus de reconnaissance pourrait être associé avec d'autres techniques plus habituelles.

Une autre ouverture très importante liée à l'analyse du mode filtrage est la possibilité d'utiliser cet outil puissant pour répondre à des requêtes spécifiques, dans un contexte d'indexation de documents, par exemple. Les besoins autour de cette nouvelle thématique se faisant de plus en plus marquants ces derniers temps, cette perspective de développement est actuellement considérée de manière très sérieuse au sein du laboratoire PSI. Toutes ces considérations et les résultats courants nous rendent très optimistes à propos du futur du projet NATALI, puisque les résultats et les améliorations possibles de la méthodologie paraissent indiquer que le taux de classification devrait augmenter encore de plusieurs points.

6. remerciements

Les auteurs aimeraient remercier MM. Joël Iviglia et Mickaël Barbey pour leur contribution à ce travail.

BIBLIOGRAPHIE

- [ADAM] S. Adam, R. Mullet, J-M. Ogier, C. Cariou, J. Gardes, Y. Lecourtier, « Combination of Invariant Pattern Recognition primitives on Technical Documents », *Proceedings of Graphic RECOgnition, GREC'99*, Jaipur, India, pp. 203-210.
- [ANTOINE] D. Antoine, S. Collin, K. Tombre, « Analysis of technical documents : The REDRAW system », *Structured Document Analysis*, Baird H.S., Bunke H., Yamamoto K. Editors, Springer-Verlag, 1992, pp. 385-402.
- [ARSENAULT] H. Arsenaault and Y. Sheng, « Properties of the circular harmonic expansion for rotation-invariant pattern recognition », *Applied Optics*, **25**, 1986, pp. 3225-3229.
- [BAILEY] R.R. Bailey, M. Srinath, « Orthogonal moment features for use with parametric and non parametric classifiers », *IEEE Trans. on PAMI*, **18** (4), 1996, pp. 389-399.
- [BAMIEH] R. Bamieth, R. De Figueiredo, « A general moments invariants/attributed graph method for the three dimensional object recognition from a single image », *IEEE Journal of Robotics Automation*, **2**, 1986, pp. 240-242.
- [BELKASIM] S. O. Belkasim, M. Shridar, M. Ahmadi, « Pattern recognition with moment invariants : a comparative study and new results », *Pattern Recognition*, **24**, 1991, pp. 1117-1138.
- [BOATTO-1] L. Boatto *et al.*, « An interpretation system for land register maps », *IEEE Computer Magazine*, vol. **25**, 1992, pp. 25-33.
- [BOATTO-2] L. Boatto *et al.*, « Detection and separation of symbols connected to graphics in line drawings », *Proc. of 11th International Conference on Pattern Recognition*, The Hague, vol. 2, 1992, pp. 545-548.
- [CASEY] R.G. Casey, E. Lecolinet, « A Survey of methods and strategies in character segmentation », *IEEE Trans. on PAMI*, **18** (7), 1996, pp. 690-706.
- [CHEN] Y. Chen, N.A. Langrana, A.K. Das, « Perfecting vectorized mechanical drawings », *Computer Vision and Image Understanding*, **63** (2), 1996, pp. 273-286.
- [DAI] M. Dai, P. Baylou and M. Najim, « An efficient algorithm for computation of shape moments from run-length codes or chain codes », *Pattern Recognition*, **25**, 1992, pp. 1119-1128.
- [DENHARTOG] J.E. Den Hartog, T. K. Ten Kate, J.J. Gerbrands, « Knowledge-based interpretation of utility maps », *Computer Vision and Image Understanding*, **63**, 1996, pp. 105-117.
- [DERRODE-1] S. Derrode, *Représentation de formes planes à niveaux de gris par différentes approximations de Fourier-Mellin analytique en vue d'indexation de bases d'images*, Thèse de Doctorat, Université de Rennes 1, 1999.
- [DERRODE-2] S. Derrode, F. Ghorbel, « Transformée de Fourier-Mellin numérique – Reconstruction et estimation de mouvements d'objets à niveaux de gris », *Actes du Colloque GRETSI*, Grenoble, France, 1997, pp. 655-658.
- [DESSEILLIGNY-1] M. P. Deseilligny, H. Le Men, G. Stamon, « Character string recognition on maps, a rotation-invariant recognition method », *Pattern Recognition Letters*, **16**, 1995, pp. 1297-1310.
- [DESSEILLIGNY-2] M.P. Deseilligny, H.L. Men, G. Stamon, « Lecture Automatique des Écritures sur Cartes Scannées », *Traitement du Signal*, **12** (6), 1995, pp. 639-652.
- [DORI-1] D. Dori, L. Wenyin, « Automated CAD conversion with the machine drawing understanding system: concepts, algorithms, and performances », *IEEE Trans. on SMC – Part A*, **29** (4), 1999, pp. 411-416.
- [DORI-2] D. Dori, W. Liu, « Vector-based segmentation of text connected to graphics in engineering drawings », P. Perner, P. Wang and A. Rosenfeld, eds., *Lecture Notes in Computer Science*, **1121**, 1996, pp. 322-331.
- [DUDANI] S. A. Dudani, K. J. Bredding, R. M. McGhee, « Aircraft identification by moment invariants », *IEEE Trans. on Computers*, **26**, 1977, pp. 39-45.
- [FLETCHER] L.A. Fletcher, R. Kasturi, « A robust algorithm for text string separation from mixed text/graphics images », *IEEE Trans. on PAMI*, **10** (6), 1988, pp. 910-918.
- [FUKUMI] M. Fukumi, S. Omatu, T. Takeda, T. Kosaka, « Rotation invariant neural pattern recognition system with application to coin recognition », *IEEE Trans. on Neural Networks*, **3**, 1992, pp. 272-279.
- [GHORBEL] F. Ghorbel, « A complete invariant description for gray level images by the harmonic analysis approach », *Pattern Recognition Letters*, **15**, 1994, pp. 1043-1051.
- [GRENIER] V. Grenier, R. Mullet, J-M. Ogier, S. Adam, J. Gardes, Y. Lecourtier, « Une architecture distribuée pour l'interprétation de documents techniques », *Actes de RFIA 2000*, Paris, Janvier 2000.
- [HU] M. K. Hu, « Visual pattern recognition by moment invariants », *IRE Trans. on Information Theory*, **8**, 1962, pp. 179-187.
- [JOSEPH] S. H. Joseph, P. Pridmore, « Knowledge-directed interpretation of mechanical engineering drawings », *IEEE Trans. on Pattern Analysis and Machine Intelligence*, **14**, 1992, pp. 928-940.
- [KASTURI] R. Kasturi, J. Alemany, « Information extraction from images of paper-based maps », *IEEE Trans. on Software Engineering*, **14** (5), 1988, pp. 671-675.
- [KHOTANZAD-a] A. Khotanzad, Y. H. Hong, « Rotation invariant image recognition using features selected via a systematic method », *Pattern Recognition*, **23**, 1990, pp. 1089-1101.
- [KHOTANZAD-b] A. Khotanzad, Y. H. Hong, « Invariant image recognition by Zernike moments », *IEEE Trans. on PAMI*, **12** (5), 1990, pp. 489-497.

- [KITA] N. Kita, « Object locating based on concentric circular description », *Proc. 11th IEEE International Conference of Pattern Recognition*, The Hague, vol. 1, 1992, pp. 637-641.
- [KUHL] F.P. Kuhl, C.R. Giardana, « Elliptic Fourier features of closed contour », *Computer Vision, Graphics and Image Processing*, **18**, 1982, pp. 236-258.
- [LAI] C.P. Lai, R. Kasturi, « Detection of dimension sets in engineering drawings », *IEEE Trans. on PAMI*, **16** (8), 1994, pp. 848-855.
- [LANGRANA] N.A. Langrana, Y. Chen, A.K. Das, « Feature identification from vectorized mechanical drawings », *Computer Vision and Image Understanding*, **68** (2), 1997, pp. 127-145.
- [LEFRERE] L. Lefrere, *Contribution au Développement d'Outils pour l'Analyse Automatique de Documents Cartographiques*, Thèse de Doctorat, Université de Rouen, 1993.
- [LIAO] S.X. Liao, M. Pawlak, « On the accuracy of Zernike moments for image analysis », *IEEE Trans. on PAMI*, **20** (12), 1998, pp. 1358-1364.
- [LIN-1] B. C. Lin, J. Shen, « Fast computation of moment invariants », *Pattern Recognition*, **24**, 1991, pp. 807-813.
- [LIN-2] C. H. Lin, « New forms of shape invariants from elliptic Fourier descriptors », *Pattern Recognition*, **20**, 1987, pp. 535-545.
- [LIU] W. Liu, D. Dori, « A protocol for performance evaluation of line detection algorithms », *Machine Vision and Applications*, **9** (5), 1997, pp. 240-250.
- [LU] Z. Lu, « Detection of text regions from digital engineering drawings », *IEEE Trans. on PAMI*, **20** (4), 1998, pp. 431-439.
- [MORI] S. Mori, C. Y. Suen, K. Yamamoto, « Historical review of OCR research and development », *Proceedings of IEEE*, **80**, 1992, pp. 1029-1098.
- [NAGY] G. Nagy, A. Samal, S. Seth, T. Fisher, E. Guthmann, K. Kalafala, L. Li, P. Sarkar, S. Sivasubramaniam, Y. Xu, « A prototype for adaptative association with streets on maps », in K. Tombre and A.K. Chhabra, eds., *Lecture Notes in Computer Science*, **1389**, 1998.
- [OGIER-1] J.M. Ogier, C. Cariou, R. Mullot, J. Gardes, Y. Lecourtier, « Interpretation of technical document: Application to French telephonic network », *Proc. ISAS*, 1998, Orlando, Florida, pp. 457-463.
- [OGIER-2] J. M. Ogier, R. Mullot, J. Labiche and Y. Lecourtier, «Multilevel approach and distributed consistency for technical map interpretation: application to cadastral maps», *Computer Vision and Image Understanding*, **70**, 1998, pp. 438-451.
- [PEI] S. C. Pei and C. N. Lin, « Normalisation of rotationally symmetric shapes for pattern recognition », *Pattern Recognition*, **25**, 1992, pp. 913-920.
- [RAVICHANDRAN] G. Ravichandran, M. Trivedi, « Circular-Mellin features for texture segmentation », *IEEE Trans. Image Processing*, **4**, 1995, pp. 1629-1640.
- [REISS] T. H. Reiss, « Recognizing planar objects using invariants image features », in *Lecture Notes in Computer Science*, Springer-Verlag, Berlin, 1993.
- [ROTHER] I. Rothe, H. Susse, K. Voss, « The Method of normalization to determine invariants », *IEEE Trans. on PAMI*, **18** (4), 1996, pp. 366-379.
- [SHENG] Y. Sheng, L. Shen, « Orthogonal Fourier-Mellin moments for invariant pattern recognition », *J. Opt. Soc. Am.*, **11**, 1994, pp. 1748-1757.
- [SHIMOSUTJI-1] S. Shimotsuji, O. Hori, M. Asano, K. Suzuki, F. Hoshino, T. Ishii, « A robust recognition system for a drawing superimposed on a map », *IEEE Computer magazine*, **25** (7), 1992, pp. 56-64.
- [SHIMOSUTJI-2] S. Shimotsuji, O. Hori, M. Asano, « Robust drawing recognition based on based-guided segmentation », *Proc. of IAPR Workshop on Document Analysis Systems*, Kaiserslautern, Germany, 1994, pp. 337-348.
- [SUZUKI] S. Suzuki, T. Yamada, « MARIS : MAp Recognition Input System », *Pattern Recognition*, **23**, 1990, pp. 919-933.
- [TAN] T. N. Tan, « Rotation invariant texture features and their use in automatic script identification », *IEEE Trans. on PAMI*, **20**, 1998, pp. 751-756.
- [TAXT] T. Taxt, J. B. Olafsdottir, M. Daehlen, « Recognition of handwritten symbols », *Pattern Recognition*, **23**, 1990, pp. 1155-1166.
- [TEAGUE] M. Teague, « Image analysis via the general theory of moments », *Journal of Optical Society of America*, **70**, 1980, pp. 920-930.
- [TEH] C. Teh, R. Chin, « On image analysis by the method of moments », *IEEE Trans. on PAMI*, **10**, 1988, pp. 496-512.
- [TRIER-1] O. D. Trier, A. K. Jain, T. Taxt, « Features extraction methods for character recognition – a survey », *Pattern Recognition*, **29**, 1996, pp. 641-662.
- [TRIER-2] O. D. Trier, T. Taxt, A.K. Jain, « Data capture from maps based on gray scale topographic analysis », *Proc. of International Conference on Documents Analysis and Recognition*, 1995, Montreal, Canada, pp. 923-926.
- [TRIER-3] O.D. Trier, T. Taxt, A.K. Jain, « Recognition of digits in hydrographic maps : binary versus topographic analysis », *IEEE Trans. on PAMI*, **19** (4), 1997, pp. 399-404.
- [VAXIVIERE] P. Vaxiviere, K. Tombre, « CELESTIN : CAD conversion of mechanical drawings », *IEEE Computer Magazine*, **25**, 1992, pp. 46-54.
- [WENYIN] L. Wenyin, D. Dori, « A proposed scheme for performance evaluation of graphics/text separation algorithms », in K. Tombre and A.K. Chhabra, *Lecture Note in Computer Science*, **1389**, 1998, pp. 359-371.
- [WOOD] J. Wood, « Invariant pattern recognition : a review », *Pattern recognition*, **29**, 1996, pp. 1-17.

Manuscrit reçu le 7 décembre 1999

LES AUTEURS

Sébastien ADAM



Sébastien Adam est né en 1975 à Dieppe. Il est docteur dans le cadre d'une convention CIFRE à France Telecom R&D Lannion et au laboratoire Perception

Système et Information (PSI) de l'université de Rouen depuis octobre 1998. Sa problématique de recherche concerne l'interprétation de documents techniques, et plus particulièrement la documentation réseau France Telecom. Ce travail l'a amené à focaliser sur le problème de la reconnaissance de caractères et de symboles, lorsque ceux-ci sont multi-orientés et multi-échelle, et ce tant du point de vue outils de traitement que du point de vue stratégie de reconnaissance.

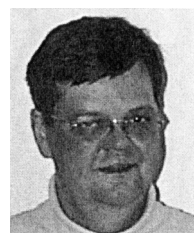
Claude CARIOU



Claude Cariou, 38 ans, est Maître de Conférences à l'École Nationale Supérieure de Sciences Appliquées et Technologie (ENSSAT) à Lannion, depuis 1993. Après une thèse de Doctorat en Electronique de l'Université de Bretagne Occidentale (Brest), soutenue en 1991 sur des travaux effectués à l'ENST de Bretagne en analyse spectrale d'images de télédétection, il rejoint le laboratoire LASTI de l'ENSSAT en 1992 en qualité d'ATER, et travaille sur la segmentation d'images texturées. Ses domaines de recherche actuels concernent la modélisation et l'extraction d'information pour la segmentation et la reconnaissance de formes, en particulier sous contrainte d'invariance géométrique.

Claude Cariou, 38 ans, est Maître de Conférences à l'École Nationale Supérieure de Sciences Appliquées et Technologie (ENSSAT) à Lannion, depuis 1993. Après une thèse de Doctorat en Electronique de l'Université de Bretagne Occidentale (Brest), soutenue en 1991 sur des travaux effectués à l'ENST de Bretagne en analyse spectrale d'images de télédétection, il rejoint le laboratoire LASTI de l'ENSSAT en 1992 en qualité d'ATER, et travaille sur la segmentation d'images texturées. Ses domaines de recherche actuels concernent la modélisation et l'extraction d'information pour la segmentation et la reconnaissance de formes, en particulier sous contrainte d'invariance géométrique.

Joël GARDES



Joël Gardes travaille à France Telecom R&D depuis 1995 dans le pôle de Recherche « Traitement des Données et des Connaissances ». Il y anime un axe de recherche sur la géomatique et, en particulier, l'utilisation de documents techniques faisant référence à des données cartographiques. Cette activité l'a conduit à établir des partenariats, en particulier avec le laboratoire PSI de l'université de Rouen, pour l'étude et le développement de systèmes d'interprétation de plans ou de cartes numérisés, pilotés par des bases de connaissances.

Jean-Marc OGIER



Jean-Marc Ogier est né en 1967. Après avoir passé son doctorat à l'Université de Rouen en 1994, il fut nommé Maître de conférences à l'I.U.T. de Lannion et exerça ses activités de recherche au L.A.S.T.I. de l'E.N.S.S.A.T.. Depuis 1998, il a rejoint l'équipe « Document et Mémoire » du laboratoire P.S.I. de l'Université de Rouen. Ses activités de recherche concernent l'interprétation des images, avec une très forte connotation vers l'analyse automatique de documents techniques. Ses intérêts portent sur le développement d'opérateurs bas niveau (segmentation, reconnaissance des formes,...), mais également sur l'implémentation de plates-formes génériques d'interprétation d'images. Sur ce point, différents secteurs sont explorés, depuis l'étude de formalisme de représentation et d'acquisition de connaissances, jusqu'au développement de plate-forme multi-agents. Jean-Marc Ogier est membre du G.R.C.E. (Groupe de Recherche en Communication Ecrite) et est co-animateur du GT5.2 (document) du GRD-PRC I³.

Jean-Marc Ogier est né en 1967. Après avoir passé son doctorat à l'Université de Rouen en 1994, il fut nommé Maître de conférences à l'I.U.T. de Lannion et exerça ses activités de recherche au L.A.S.T.I. de l'E.N.S.S.A.T.. Depuis 1998, il a rejoint l'équipe « Document et Mémoire » du laboratoire P.S.I. de l'Université de Rouen. Ses activités de recherche concernent l'interprétation des images, avec une très forte connotation vers l'analyse automatique de documents techniques. Ses intérêts portent sur le développement d'opérateurs bas niveau (segmentation, reconnaissance des formes,...), mais également sur l'implémentation de plates-formes génériques d'interprétation d'images. Sur ce point, différents secteurs sont explorés, depuis l'étude de formalisme de représentation et d'acquisition de connaissances, jusqu'au développement de plate-forme multi-agents. Jean-Marc Ogier est membre du G.R.C.E. (Groupe de Recherche en Communication Ecrite) et est co-animateur du GT5.2 (document) du GRD-PRC I³.

Rémy MULLOT



Rémy Mullot est né en 1963 à Rouen. Il reçoit la Maîtrise de Sciences et techniques en Micro-informatique à l'université de Savoie en 1986. Il obtient son doctorat de l'université de Rouen en 1991. Entre 1991, il devient Maître de Conférences à l'université de Rouen, dans le département de Physique. Ses activités de recherche portent sur l'interprétation de documents techniques au sein du laboratoire PSI. Ses centres d'intérêt portent sur les systèmes à base de connaissance, les systèmes distribués et les systèmes perceptifs pour l'interprétation de documents. Il est depuis 2000, professeur à l'université de La Rochelle et membre du laboratoire L3i. Il est par ailleurs membre du GRCE (Groupe de Recherche en Communication Ecrite).

Rémy Mullot est né en 1963 à Rouen. Il reçoit la Maîtrise de Sciences et techniques en Micro-informatique à l'université de Savoie en 1986. Il obtient son doctorat de l'université de Rouen en 1991. Entre 1991, il devient Maître de Conférences à l'université de Rouen, dans le département de Physique. Ses activités de recherche portent sur l'interprétation de documents techniques au sein du laboratoire PSI. Ses centres d'intérêt portent sur les systèmes à base de connaissance, les systèmes distribués et les systèmes perceptifs pour l'interprétation de documents. Il est depuis 2000, professeur à l'université de La Rochelle et membre du laboratoire L3i. Il est par ailleurs membre du GRCE (Groupe de Recherche en Communication Ecrite).

Yves LECOURTIER



Yves Lecourtier est né en 1950 à Marseille. Ancien élève de l'ENSET, il a soutenu une thèse de 3^e cycle en traitement du signal en 1978, et une thèse d'état en automatique en 1985 à l'université Paris Sud (Orsay). Depuis 1987 il est professeur à l'université de Rouen. Son domaine de recherche s'est orienté vers la reconnaissance de forme, les réseaux de neurones et les architectures multi-agent dans le cadre de l'analyse de documents et de la reconnaissance de texte. Y. Lecourtier est le coauteur de plus d'une centaine d'articles et de communications à des conférences sur ces différents thèmes. Il a créé et dirigé le laboratoire Perception Système et Information commun à l'université et l'INSA de Rouen. Il est membre des CA de l'AFRIF, et de l'ASTI et préside le conseil des associations fondatrices de l'ASTI. Il est actuellement président du GRCE, association qui rassemble les chercheurs francophones dans le domaine de la reconnaissance de l'écrit et de l'analyse de documents.

Yves Lecourtier est né en 1950 à Marseille. Ancien élève de l'ENSET, il a soutenu une thèse de 3^e cycle en traitement du signal en 1978, et une thèse d'état en automatique en 1985 à l'université Paris Sud (Orsay). Depuis 1987 il est professeur à l'université de Rouen. Son domaine de recherche s'est orienté vers la reconnaissance de forme, les réseaux de neurones et les architectures multi-agent dans le cadre de l'analyse de documents et de la reconnaissance de texte. Y. Lecourtier est le coauteur de plus d'une centaine d'articles et de communications à des conférences sur ces différents thèmes. Il a créé et dirigé le laboratoire Perception Système et Information commun à l'université et l'INSA de Rouen. Il est membre des CA de l'AFRIF, et de l'ASTI et préside le conseil des associations fondatrices de l'ASTI. Il est actuellement président du GRCE, association qui rassemble les chercheurs francophones dans le domaine de la reconnaissance de l'écrit et de l'analyse de documents.