



IRISA • Campus universitaire de Beaulieu • 35042 Rennes Cedex France • Tél.: +33 2 99 84 71 00 • Télécopie : +33 2 99 84 71 71 • Internet : www.irisa.fr

INSTITUT DE RECHERCHE EN INFORMATIQUE ET SYSTÈMES ALÉATOIRES

Rapport d'activité 2001

Projet IMADOC

Interprétation et Reconnaissance d'Images et de Documents

Rennes

————— THÈME 3B —————



Table des matières

1	Composition de l'équipe	3
2	Présentation et objectifs généraux	3
3	Fondements scientifiques	5
3.1	Vision précoce	5
3.2	DMOS, une méthode générique de reconnaissance de documents	8
3.3	Reconnaissance des formes par systèmes d'inférence floue	10
3.4	Modèle et système de perception et d'interprétation (SPI)	12
4	Domaines d'applications	14
4.1	Analyse de documents papier	14
4.2	Annotations et méta-données	15
4.2.1	Annotation de documents numériques	15
4.2.2	Bibliothèque virtuelle de typographie	15
4.3	Reconnaisseurs d'écriture manuscrite	16
4.4	Interfaces orientées stylo	16
5	Logiciels	17
5.1	RESIF, Reconnaissance d'Écriture Manuscrite en-ligne par Systèmes d'Inférence Floue	17
5.2	DocRead : générateur automatique de systèmes de reconnaissance de documents structurés	17
5.3	Bibliothèque d'extraction d'indices visuels	18
6	Résultats nouveaux	18
6.1	Application de la méthode DMOS à la reconnaissance de structures tabulaires	18
6.2	Préparation de l'insertion de la reconnaissance de l'écriture manuscrite dans la méthode DMOS	19
6.3	Annotations	20
6.4	Bibliothèque virtuelle de typographie	21
6.5	Modélisation automatique des connaissances par systèmes d'inférence floue hiérarchisés	21
6.6	Intégration du système de reconnaissance de caractères manuscrits ResifCar dans des téléphones mobiles	24
6.7	Expérimentation sur l'analyse en temps-réel de la structure d'une formule mathématique, utilisation de l'approche perceptive SPI	25
7	Contrats industriels (nationaux, européens et internationaux)	27
7.1	Portage du logiciel ResifCar de reconnaissance de caractères manuscrits	27
7.2	Annotation coopérative des documents d'archives	27
8	Actions régionales, nationales et internationales	27
8.1	Actions régionales	27
8.2	Actions nationales	28

8.2.1	ICL	28
8.2.2	GdR	28
8.2.3	Sociétés savantes	28
8.2.4	Listes de discussion	28
8.3	Actions internationales	28
9	Diffusion de résultats	29
9.1	Animation de la Communauté scientifique	29
9.1.1	Comité éditorial de journaux	29
9.1.2	Comités de programme de conférences	29
9.2	Enseignement	29
9.2.1	Enseignement universitaire	29
9.2.2	Autres enseignements	30
9.3	Participation à des colloques, séminaires, invitations	30
10	Bibliographie	30

1 Composition de l'équipe

Responsable scientifique

Jean Camillerapp [professeur Insa]

Assistante de projet

Édith Blin [TR Inria]

Personnel Inria

Jacques André [DR]

Personnel Insa

Éric Anquetil [maître de conférences]

Bertrand Coüasnon [maître de conférences]

Ivan Leplumey [maître de conférences]

Personnel Université de Rennes 1

Guy Lorette [professeur]

Charles Quéguiner [maître de conférences]

Hélène Richy [maître de conférences]

Laurent Pasquer [ATER]

Chercheur doctorant

Nicolas Ragot [boursier université]

Ingénieur Expert

Nicolas Briec [sur contrat région Mègalis depuis le 1/10/2001]

2 Présentation et objectifs généraux

De manière générale, les thèmes de recherche du projet touchent à la *communication écrite* et à *l'ingénierie du document* sous différents aspects : *analyse, reconnaissance, interprétation de documents numérisés, interaction homme-document*. Ces recherches concernent l'écrit et le document sous toutes leurs formes (manuscrit, imprimé, graphique, images, documents composites, etc.) ainsi que les activités qui y sont liées.

Pour faire face à la multiplicité et à la diversité des applications, la solution usuelle consiste à apporter des réponses dédiées à chaque type d'application. Le projet privilégie, au contraire, le développement d'approches *génériques*. Pour ce faire, la recherche s'effectue selon cinq thèmes interdépendants, deux de nature plus méthodologique et trois de nature plus applicative.

Axes de recherche

– Traitement d'image, vision précoce

Les traitements de bas niveau, également appelé traitements de vision précoce, ont une influence certaine sur les performances globales des chaînes de traitement. La numérisation des images de documents doit réaliser un compromis entre la taille totale des images et la taille apparente des caractères. Dans ce compromis, la largeur du trait peut se réduire à deux ou trois pixels. Le traitement de l'image de document nécessite donc une très bonne précision dans l'affectation d'un pixel au fond ou au tracé et dans l'extraction des caractéristiques des indices visuels.

– **Modélisation et gestion de connaissances**

Afin de développer des systèmes automatiques d'interprétation de signaux et d'images de documents, le projet étudie des méthodes de reconnaissance dynamique et multi-contextuelle basées sur des indices visuels. À cette fin, nous avons trois approches complémentaires de modélisation des connaissances contextuelles : l'une à base de règles (grammaire) pour les connaissances symboliques, une autre fondée sur les Systèmes d'Inférence Floue (SIF) pour les connaissances imprécises, et la troisième à base d'une colonne neuronale dynamique et récurrente pour les données numériques.

La modélisation des connaissances se fait, soit à l'aide de modèles de connaissances *a priori*, soit par apprentissage automatique à partir de données expérimentales.

Afin d'aboutir à l'interprétation finale, la gestion dynamique des connaissances en contexte est faite, selon les types de modélisation respectifs, soit par analyse syntaxique modifiée, soit par une méthode de « *soft computing* » (fusion de données floues ou cycles perceptifs).

Ces deux approches présentent l'avantage d'offrir plusieurs niveaux de lecture possibles, de repérer où subsistent des erreurs, d'aboutir à des systèmes transparents dont le fonctionnement peut être analysé et dont les performances peuvent être ainsi optimisées.

– **Écriture manuscrite**

Nous étudions des systèmes de reconnaissance de documents *manuscrits*, soit pendant leur phase de création (reconnaissance d'écriture *en-ligne*), soit après coup (reconnaissance d'écriture *hors-ligne*).

Le fonctionnement de ces systèmes repose sur des principes fondamentaux de l'écriture, des connaissances *a priori* sur les styles d'écriture et des indices visuels issus du tracé.

Ceci nous a permis de surmonter le problème complexe de la segmentation des mots en lettres afin de pouvoir reconnaître les mots d'un vocabulaire de grande taille (approche *analytique*). Ceci fait également que l'utilisation de tels systèmes peut se généraliser auprès d'un public très large (systèmes de *reconnaissance omni-scripteurs et multilingue*). Néanmoins, il reste toujours possible de spécialiser un système pour en faire une application dédiée.

– **Document numérique et web sémantique**

Les récentes évolutions technologiques (développement de l'internet, communications à haut débit, importantes capacités de stockage, algorithmes de compression d'images efficaces, politique de numérisation à grande échelle des archives patrimoniales ou industrielles) permettent de mettre à disposition d'énormes corpus de documents numérisés. Toutefois, un accès intelligent aux images (*web sémantique*) suppose l'existence d'annotations (méta-données [12]) décrivant avec suffisamment de précisions le contenu de ces images.

Les études que nous menons actuellement concernent la mise au point de modèles et de techniques visant à faciliter l'annotation automatique d'images numérisées. Nos recherches portent sur la transformation d'images de documents-papier existants (rétroconversion de documents anciens) en documents numériques structurés [13]. L'approche générique porte non seulement sur la description des formes (géométrie), mais aussi sur la définition de modèles d'annotation et d'ontologies sur les contenus. Cette approche permet de contrôler les différentes étapes de la transformation des documents :

- localisation : repérage des zones dans les images,
- rétroconversion : extraction d'information et structuration,
- annotation : analyse et interprétation des contenus,

- indexation : recherche d'informations (outils de recomposition d'image, filtrage., etc.)

– **Interaction homme-machine**

Le projet étudie de nouvelles interfaces homme-machine orientées stylo, car celui-ci possède l'avantage d'être d'un mode d'utilisation naturel, immédiat et connu de tout un chacun.

L'utilisation d'un *papier électronique* avec retour visuel par encre électronique est encore plus ergonomique et l'utilisateur retrouve alors sensiblement les habitudes qu'il avait avec un stylo et une feuille de papier.

Grâce au caractère générique de nos approches et à la maîtrise des axes de recherche cités ci-dessus, il est possible de faire face à des applications très diverses et d'effectuer des transferts industriels en étudiant et en réalisant, assez rapidement, des prototypes.

3 Fondements scientifiques

3.1 Vision précoce

Mots clés : vision précoce, binarisation, squelettisation, filtrage de Kalman,.

Résumé :

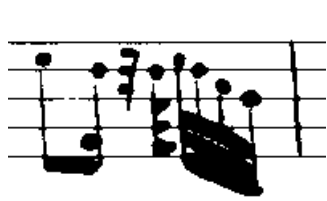
Afin de ne pas trop alourdir les volumes de mémoire et les temps de traitement, les documents sont généralement numérisés avec la densité la plus faible possible (souvent 200 dpi). Dans ces conditions, les éléments du tracé n'ont que quelques pixels d'épaisseur. Pour répondre aux exigences des procédures de reconnaissance, les traitements d'image doivent être capables de fournir une binarisation fiable des petits éléments, une extraction du squelette à partir d'images en niveaux de gris, et une détection des segments de droite même en présence de recouvrement et d'intersection.

Binarisation adaptative

Les documents sont généralement numérisés en limite de résolution et certains éléments très importants pour la reconnaissance peuvent n'occuper que quelques pixels. Ainsi, par exemple, la non détection d'une toute petite zone blanche peut masquer la présence, pourtant très informante, de la boucle d'un *e*.

Une numérisation en niveaux de gris des images de document permet de mettre en œuvre des algorithmes de binarisation adaptative, souvent basés sur des coopérations entre la détection des contours et le grossissement des régions. Ceux-ci donnent une interprétation locale des valeurs des niveaux de gris qui séparent bien les différents objets dans le document (cf figures 1 et 2).

Dans ces conditions, en limite de résolution, les mesures des caractéristiques des objets dans des images binaires sont sensibles au bruit provenant de l'échantillonnage spatial. L'interprétation locale des niveaux de gris permet de positionner la transition entre les objets et le fond avec une précision meilleure que le pixel. En augmentant ainsi la résolution spatiale, on améliore la pertinence des paramètres liés aux indices visuels.



(a) Seuil de 210 permettant la détection des lignes fines.

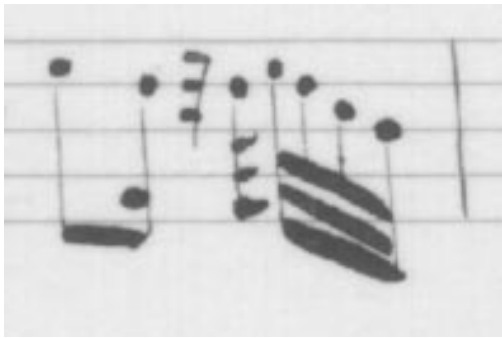


(b) Seuil de 195, gardant encore des éléments verticaux.



(c) Seuil de 110, permettant la séparation des barres de croche.

FIG. 1 – Binarisation avec un seuil fixe



(a) Image initiale en niveaux de gris.



(b) Image binarisé avec un seuil adaptatif.

FIG. 2 – Binarisation adaptative

Squelettisation d'images en niveaux de gris

Destinée au départ à l'affinage d'images de ligne, les techniques de squelettisation se sont par la suite développées en oubliant un peu cet objectif. Dans les images de documents, il nous semble important de revenir à la notion primitive d'axe médian pour modéliser certaines parties du tracé. Il faut alors distinguer les zones réellement squelettisables ou zones régulières qui correspondent à un tracé simple et les zones singulières localisées dans les portions du tracé qui comportent des recouvrements et dans lesquelles la notion théorique d'axe médian n'a pas d'intérêt pratique [9].

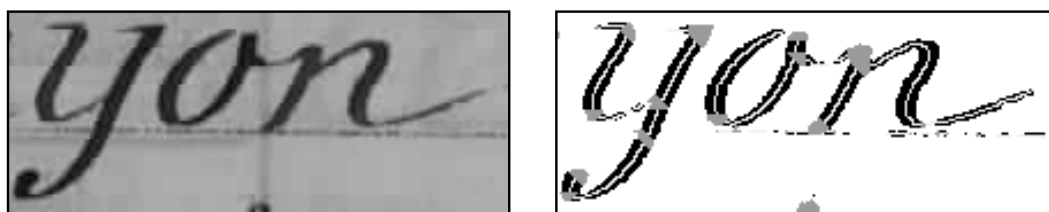


FIG. 3 – Traitement d'une image issue d'un document ancien, le squelette est en gris clair et les régions singulières en gris foncé.

Afin de limiter le bruit de binarisation évoqué ci-dessus, on peut obtenir une localisation *sub-pixelle* des points de cet axe en travaillant à partir d'images en niveaux de gris.

Segmentation par filtrage de Kalman

Les structures linéaires tiennent une grande place dans les documents mais elles sont susceptibles d'interférer entre elles ou avec le tracé lui-même. Cependant la simplicité de leur structure permet d'envisager de conduire simultanément la segmentation et la reconnaissance.

Le filtrage de Kalman est une technique d'identification des paramètres d'un modèle à partir d'une suite ordonnée de mesures. Dans le cas des tracés rectilignes, le modèle se réduit à l'épaisseur du trait, sa pente et l'équivalent de l'ordonnée à l'origine. La mesure se déduit de la position et de l'épaisseur des empan noirs ¹ dans une direction approximativement orthogonale au tracé.

Le filtrage de Kalman calcule également la matrice de covariance de son estimation des paramètres du modèle. Cette matrice permet d'évaluer d'une part la vraisemblance de l'affectation d'une mesure à un segment et d'autre part la vraisemblance de la poursuite d'une hypothèse perturbée par la présence d'un autre objet [4].

La véritable difficulté de cette approche ne réside pas dans les équations de Kalman, mais dans le processus d'association des empan extraits de l'image avec les empan prédits lorsque les structures linéaires se coupent ou se chevauchent. Un même empan de l'image peut être le résultat du recouvrement de plusieurs tracés. C'est à ce niveau que l'information apportée par les matrices de covariance est d'une grande importance.

Dans cette approche nous pouvons également utiliser une localisation *sub-pixelle* des extrémités des empan fournie par la binarisation adaptative.

¹Ensemble de pixels noirs consécutifs selon une des quatre directions : horizontale, verticale ou diagonales

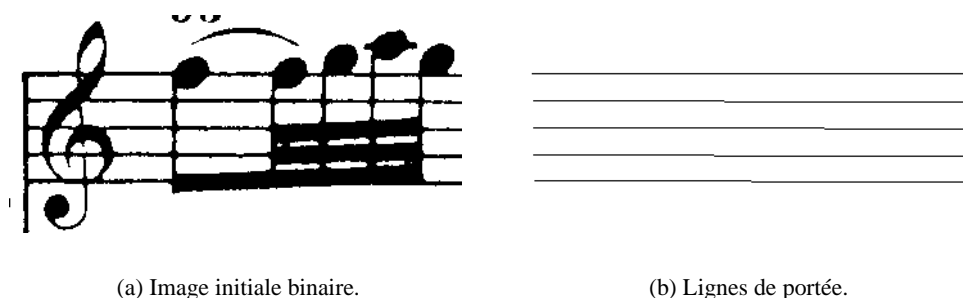


FIG. 4 – Exemple d'utilisation du filtrage de Kalman pour l'extraction des lignes de portées traversant des barres de croche.

3.2 DMOS, une méthode générique de reconnaissance de documents

Mots clés : analyse structurelle, grammaire, Definite Clause Grammar (DCG), segmentation, gestion des connaissances a priori, reconnaissance de documents, partitions musicales, formules mathématiques, tableaux, formulaires.

Résumé : Afin d'éviter de réécrire complètement un nouveau système de reconnaissance à chaque nouveau type de documents, nous proposons une méthode générique de reconnaissance de document, baptisée DMOS. Cette méthode est constituée du langage EPF permettant de modéliser grammaticalement la connaissance a priori, et d'un analyseur associé capable de remettre en cause, au cours de l'analyse, la structure analysée tout en gérant correctement le bruit très présent en reconnaissance de documents.

La conception d'un système de reconnaissance de documents est longue et difficile. Or nous constatons dans l'état de l'art que l'adaptation d'un système de reconnaissance à un nouveau type de documents est impossible. Elle nécessite en fait à chaque fois une réécriture complète du système, ce qui correspond à une perte d'énergie très importante. La conception d'un système générique de reconnaissance de documents constitue un défi important, car il permettrait d'éviter ce travail coûteux de réécriture.

En outre, l'augmentation de la qualité de la reconnaissance, nécessaire dans un contexte de traitement industriel de documents, peut se faire notamment en résolvant les problèmes de segmentation que l'on rencontre souvent dans les documents à traiter.

Nous pouvons finalement constater que ces problèmes de reconnaissance de documents se situent autour de la gestion des connaissances *a priori*. En effet :

- un système générique doit permettre de ne modifier que ce qui est nécessaire lors de l'adaptation à un nouveau type de document, la connaissance *a priori* sur le document ;
- pour résoudre les problèmes de segmentation, il faut introduire de la connaissance *a priori* dans le système.

Nous avons donc proposé, pour des documents à forte structure dans lesquels des règles d'écriture peuvent être connues, une méthode baptisée DMOS (Description avec MODification de la Segmentation), constituée de :

- un langage grammatical de description de documents, EPF (*Enhanced Position Formalism*), permettant de modéliser la connaissance *a priori* ;
- l’analyseur associé autorisant une modification en cours d’analyse de la structure analysée. Cette modification permet d’introduire le contexte (niveau symbolique) dans la phase de segmentation (niveau numérique), afin d’améliorer la reconnaissance.

Langage EPF

Un certain nombre de formalismes grammaticaux permettant de décrire des objets bidimensionnels ont été proposés. Cependant, soit ils offrent une expressivité trop faible (grammaires d’arbres ou grammaires Web), soit ils ont une syntaxe trop compliquée (grammaires Plex ou grammaires de graphes) qui rend très difficile la mise en oeuvre de connaissances complexes. En outre, aucun ne permet d’introduire la connaissance formalisée dans la phase de segmentation.

Nous avons donc défini et développé le langage EPF[16], permettant de décrire un document structuré aussi bien au niveau graphique que syntaxique. Ce langage peut être vu comme une extension bidimensionnelle des grammaires classiques, où au lieu d’être des caractères, les terminaux sont des segments ou des matrices de pixels (composantes qui représentent un symbole). Cette extension est également faite à l’aide d’un certain nombre d’opérateurs spécifiques dont voici quelques exemples :

Opérateur de position (encadré par AT):

`A && AT(pos) && B`

signifie A, et à la position `pos` relativement à A, on trouve B,

où `&&` est la concaténation dans la grammaire, A et B représentent un terminal ou un non-terminal.

Opérateur de factorisation (##, en association avec les opérateurs de position):

`A && (AT(pos1) && B ##
AT(pos2) && C)`

signifie `A && AT(pos1) && B` et `A && AT(pos2) && C`

Grâce à cette syntaxe, il est possible, par exemple de décrire un groupe de notes simplifié (des croches reliées par une seule barre de groupe) :

```
groupeDeNote ::= barreDeGroupe &&
               (AT(extremiteGauche) && noteGr ##
                notesAuMilieu ##
                AT(extremiteDroite) && noteGr)
```

où `::=` est le constructeur d’une règle grammaticale. Le concepteur de la grammaire peut définir à la demande des opérateurs de position, de la même manière qu’il peut le faire pour des non-terminaux.

Opérateurs de référence multiple (`--->` et `<---`): Afin de pouvoir faire référence plusieurs fois à une même instance du terminal ou du non-terminal A, nous proposons de sauvegarder cette instance à l’aide de `--->`. Cette sauvegarde permet ensuite de faire référence (à l’aide de `<---`) à A autant de fois que nécessaire. Il est ainsi possible de décrire ce qu’est un rectangle :

```

rectangle ::=
  (segV ---> segCoteGauche) &&
  AT(extremiteHaute) && segH &&
  AT(extremiteDroite) && segV &&
  AT(extremiteBasse) && segH &&
  AT(extremiteGauche) &&
  (segV <--- segCoteGauche).

```

Analyseur associé

Le langage EPF décrit ci-dessus permet de définir grammaticalement le document à reconnaître. De cette grammaire nous produisons automatiquement un analyseur qui possède des caractéristiques spécifiques à l'analyse de documents bidimensionnels. Ainsi, nous pouvons souligner les trois principales caractéristiques de l'analyseur que nous avons développé (à deux dimensions), par rapport à un analyseur classique (à une dimension) pour les langages formels :

- remise en cause de la structure analysée en cours d'analyse (pour effectuer des segmentations contextuelles) ;
- détection de l'élément suivant à analyser. En effet pour les analyseurs classiques, l'élément suivant est simplement celui qui est en tête de la chaîne analysée alors qu'en deux dimensions, l'élément suivant peut être n'importe où dans l'image, donc n'importe où dans la structure analysée ;
- gestion correcte du bruit. Contrairement aux analyseurs classiques où la structure analysée est peu bruitée, en reconnaissance de documents il est nécessaire que l'analyseur soit capable de reconnaître le maximum d'informations dans un flux très bruité. Nous pouvons considérer que la gestion du bruit correspond à trouver l'élément suivant, malgré le bruit.

En conclusion, nous sommes arrivés, grâce à la définition et la mise en oeuvre du formalisme EPF et de l'analyseur associé, à concevoir un système générique de reconnaissance de documents structurés.

La création d'un nouveau système de reconnaissance adapté à un nouveau type de document s'obtient alors par simple compilation d'une description du document réalisée avec le langage EPF. Éventuellement, un nouvel apprentissage automatique du classifieur servant à reconnaître les symboles (terminaux représentés par des matrices de pixels), peut être nécessaire pour de nouveaux symboles.

La méthode DMOS a ainsi permis de développer des systèmes de reconnaissance de partitions musicales, de tableaux, de formulaires et de formules mathématiques, en définissant uniquement une description du document.

3.3 Reconnaissance des formes par systèmes d'inférence floue

Mots clés : modélisation explicite des connaissances, systèmes d'inférence floue, apprentissage, classification.

Résumé : *Dans le domaine de la reconnaissance de forme, la modélisation de systèmes complexes engendre généralement l'interaction de plusieurs processus interdépendants. Face à la complexité d'interaction des différents maillons qui composent un tel système, il*

est important de pouvoir maîtriser chacun des concepts mis en jeu. Dans cet objectif, les systèmes d'inférence floue (SIF) permettent la conception de systèmes de reconnaissance interprétables basés notamment sur une modélisation explicite des connaissances.

Pour la plupart des problèmes réels de reconnaissance de forme, la conception automatique d'une modélisation explicite constitue un problème particulièrement complexe par suite de la grande variabilité observée tant à l'intérieur de chaque classe de formes (variabilité intra-classes) qu'entre classes de formes différentes (variabilité inter-classes). La plupart des approches s'appuient sur la capacité d'apprentissage automatique des méthodes de type *stochastique* ou *connexionniste* en se basant respectivement, soit sur une modélisation de l'étendue de la variabilité des formes, soit sur une discrimination directe des classes. Ces approches conduisent souvent à des systèmes de reconnaissance *opaques* (type boîte noire) pour lesquels il est très difficile d'avoir une interprétation des processus de décision mis en œuvre. C'est pourquoi, même si ces approches permettent la réalisation de systèmes affichant assez rapidement des performances intéressantes, ces systèmes se révèlent peu évolutifs, il est donc difficile d'améliorer leurs performances.

Dans l'optique de pouvoir réaliser des systèmes interprétables, nous avons mis au point une méthodologie de modélisation basée sur la génération automatique de *systèmes d'inférence floue* [2]. Plus précisément, les règles floues utilisées répondent aux hypothèses suivantes :

- les entrées $x = (x_1, \dots, x_n)$ sont non floues et définies dans l'espace à n dimensions appelé espace des attributs ou encore espace des entrées ;
- les sous-ensembles flous de sortie B_{ik} sont des singletons $\{b_{ik}\}$;
- les conclusions des règles ($Y_k = b_{ik}$) sont donc précises : Y_k représente le degré d'appartenance de la forme présentée en entrée, vis-à-vis de la règle R_i et de la classe k .

Considérons maintenant N_r règles floues formant un SIF permettant de caractériser ou de classer les formes considérées dans C classes. L'expression linguistique d'une règle floue associée à la caractérisation d'une classe k se formalise d'une manière générale de la façon suivante :

**R_i : SI x_1 est M_{i1} et x_2 est M_{i2} et ... et x_n est M_{in}
ALORS la donnée appartient à la classe k et non aux autres classes.**

Ce qui se traduit dans le formalisme classique des règles floues par :

**R_i : SI x_1 est M_{i1} et x_2 est M_{i2} et ... et x_n est M_{in}
ALORS $Y_1 = \delta_{u1}$ et $Y_2 = \delta_{u2}$ et ... et $Y_u = \delta_{uu}$ et ... et $Y_c = \delta_{uc}$,
où δ_{ij} désigne le symbole de Kronecker.**

Chaque condition (x_j est M_{ij}) est interprétée comme le degré d'appartenance de l'observation x_j au sous-ensemble flou M_{ij} , c'est-à-dire $\mu_{M_{ij}}(x_j)$. L'inférence floue est alors définie de la manière suivante :

$$\mu_{B'_k}(y) = \perp_{i=1}^{N_r} I(\beta_i, \mu_{B_{ik}}(y)) \text{ avec } \beta_i = \mathbf{T}_{j=1}^n \mu_{M_{ij}}(x_j),$$

où \mathbf{T} est une T -norme symbolisée par le *et* utilisée dans la partie prémisses des règles, β_i représente la valeur d'activation de la règle R_i , I est l'implication floue [BM95], \perp représente l'opérateur d'agrégation et N_r est le nombre total de règles.

Le résultat obtenu en sortie du SIF mesure alors le degré d'appartenance de la forme présentée en entrée, relativement à chacune des classes.

Les systèmes d'inférence floue permettent de concilier une modélisation de nature numérique qui caractérise la plupart des problèmes réels de reconnaissance de forme, avec une modélisation robuste et qualitative des connaissances extraites.

La génération automatique des SIF repose sur une analyse non supervisée des données d'apprentissage. Pour ce faire, nous proposons d'utiliser les concepts récents de la classification possibiliste [2] afin d'extraire et de qualifier un ensemble de propriétés pertinentes. La classification non supervisée de type possibiliste permet de caractériser de manière précise les propriétés intrinsèques de chaque prototype modélisé. Ces propriétés sont alors représentées par des sous-ensembles flous M_{ij} , directement par projection sur les différents axes de l'espace de représentation. Ces différents sous-ensembles flous sont ensuite rassemblés pour former des règles floues dites prototypes qui vont modéliser de manière "lisible" ces propriétés pertinentes.

L'approche méthodologique proposée a été mise en œuvre pour la conception d'un système complet de reconnaissance en-ligne d'écriture manuscrite.

3.4 Modèle et système de perception et d'interprétation (SPI)

Mots clés : perception, interprétation, cycles perceptifs, réseaux de neurones.

Résumé :

Les méthodes classiques de reconnaissance de formes concernent la reconnaissance d'objets isolés. Cette reconnaissance est un processus purement statique. Cependant, le Gestaltisme a mis en évidence le fait que la perception humaine est influencée par les objets qui se trouvent dans le voisinage immédiat de l'objet à reconnaître, c'est-à-dire par son contexte spatial. De plus, les recherches en sciences cognitives ont montré que la perception et l'interprétation visuelle chez l'Homme reposent sur un processus de construction dynamique d'une représentation mentale [Kan79], [EK00]. Ce processus consiste à organiser et à intégrer dynamiquement des données et des connaissances situées à plusieurs niveaux de représentation, en se fondant sur des critères d'organisation perceptive et de cohérence globale afin d'aboutir à différentes interprétations possibles. Ce n'est que depuis quelques années que les problèmes de perception et d'interprétation de signaux et d'images, par ordinateur, sont devenus un enjeu de recherche important [SB94], [BS99]. Il s'agit de doter la machine de nouvelles capacités analogues à celles de l'être humain.

Nous avons étudié une nouvelle méthode de reconnaissance de formes *dynamique et multi-contextuelle* ainsi qu'un système de perception et d'interprétation de formes en contexte (SPI).

Le modèle générique développé repose sur une organisation modulaire et hiérarchique nommée *colonne*, constituée d'un empilement de N niveaux de représentation. Chacun des niveaux comporte des

[Kan79] G. KANIZA, *Computing Perceptual Organization in Computer Vision*, New York: Praeger, 1979.

[EK00] M. W. EYSENCK, M. KEANE, *Cognitive Psychology, A Student's Handbook*, Psychology Press, 2000.

[SB94] S. SARKAR, K. L. BOYER, *Computing Perceptual Organization in Computer Vision*, World Scientific, 1994.

[BS99] K. L. BOYER, S. SARKAR (éditeurs), *Perceptual Organization in Computer Vision*, Special Issue of *Computer Vision and Image Understanding*, vol. 76, n°1, Academic Press Ed., 1999.

interactions intra-niveau et inter-niveaux avec les niveaux voisins inférieurs et supérieurs (cf. figure 5). L'interaction entre plusieurs niveaux de représentation, l'entrée de données contextuelles sur plusieurs niveaux et la fusion de données en contexte se font d'une part, par un mode de représentation unifié sous forme de couches de neurones et, d'autre part, par une succession de *cycles perceptifs* qui modifient de façon itérative les activations des neurones de chacune des couches. Il s'agit donc d'un réseau *dynamique et récurrent*. De plus, compte-tenu de son architecture, ce système est intrinsèquement parallèle et réparti. Un mécanisme de *focalisation d'attention* est également mis en œuvre dans le système. Ce mécanisme consiste en une sélection dynamique des nœuds-hypothèses et des liens correspondants à partir des résultats d'une phase de pré-reconnaissance. Ceci conduit à ne considérer qu'un faisceau restreint dans l'ensemble du treillis d'hypothèses. Il s'agit donc d'un réseau *oncogénique*.

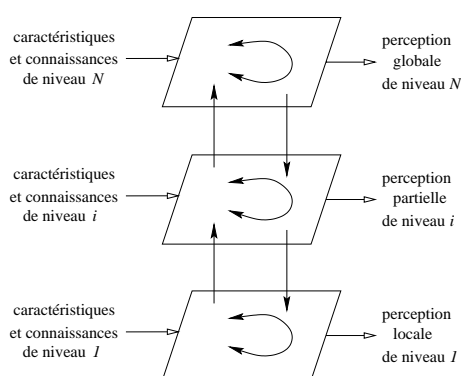


FIG. 5 – Modèle générique d'une colonne de perception et d'interprétation.

Chaque hypothèse d'existence de primitives, de sous-formes ou de formes est modélisée par un nœud-hypothèse dans un graphe valué ou dans un réseau par un neurone avec une valeur d'activation associée.

L'introduction de connaissances contextuelles dans les différents niveaux de représentation se fait en trois temps: création de nœuds représentatifs des connaissances disponibles *a priori*, introduction d'activations initiales pour ces nœuds, création de liens activateurs entre les nœuds modélisant des connaissances compatibles entre elles et de liens inhibiteurs entre les nœuds représentant des informations incompatibles entre elles.

Les données issues de la forme placée à l'entrée du système (sensations visuelles) sont hiérarchisées et entrées dans le système (cf. figure 5) à des niveaux différents selon leurs natures respectives. Elles sont perçues et interprétées en fonction de leurs contextes respectifs. À chacun des niveaux de représentation correspond donc une interprétation possible sur la sortie correspondante.

À chaque niveau (cf. figure 6), le fonctionnement du système comporte trois phases successives : l'extraction et l'entrée des informations issues de la forme à percevoir (*sensations visuelles*) relatives à chaque niveau de représentation, la fusion de ces informations avec les connaissances *a priori* (contextes intra et inter-niveaux), une interprétation en sortie du niveau de représentation considéré.

Les cycles perceptifs consistent en une succession de trois processus qui se répètent :

- un processus de relaxation à l'intérieur de chaque niveau qui augmente les activations latérales des informations compatibles entre elles et qui accroît les inhibitions latérales des informations

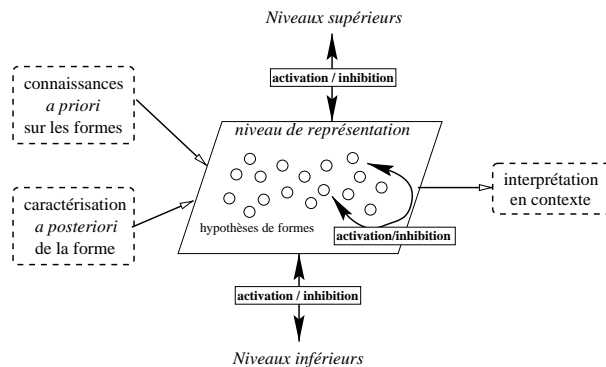


FIG. 6 – Schéma d'un niveau de représentation.

- incompatibles entre elles (mécanisme de perception en contexte),
- un processus ascendant qui consiste en une augmentation des valeurs d'activation ou d'inhibition associées aux nœuds du graphe d'un des niveaux de représentation en fonction de celles de la couche de niveau inférieur (mécanisme d'induction),
 - un processus descendant qui propage les informations des niveaux les plus élevés vers les niveaux inférieurs également selon un processus d'activation et d'inhibition (mécanisme de déduction).

Ces mécanismes d'activation, d'inhibition et de fusion de connaissances servent à gérer la cohérence de l'ensemble des informations disponibles, chacune d'elle étant placée en contexte. Il s'agit dès lors d'une interprétation multi-contextuelle. Ceci se traduit, *in-fine*, par l'émergence d'une interprétation. Cette interprétation apparaît alors comme une *évidence* qui résulte de la *fusion cohérente* de l'ensemble des connaissances disponibles. L'interprétation est partielle si elle correspond à un niveau de représentation intermédiaire et elle est globale si elle intervient au niveau le plus élevé de la colonne.

Le fonctionnement du système possède trois types de critères d'arrêt. S'il apparaît, au niveau de représentation le plus élevé dans la colonne, un neurone dont l'activation est très supérieure à celle de tous les autres, il correspond à l'interprétation finale de la forme observée. Si aucune valeur d'activation ne dépasse les autres, la forme observée est rejetée comme étant inconnue. Si après un nombre de cycles prédéterminé il existe, dans un niveau de représentation donné, un neurone dont l'activation est très supérieure à celle de tous les autres neurones, celui-ci correspond à une interprétation partielle. Ceci permet notamment au système de généraliser et de proposer une interprétation probable d'une forme qui n'a pas été apprise, mais dont tous les éléments constituants sont cohérents entre eux.

4 Domaines d'applications

4.1 Analyse de documents papier

Les domaines d'applications de nos recherches dans le contexte de l'analyse de documents papier sont très nombreux. En effet, l'approche générique que nous avons développée (méthode DMOS) permet une adaptation relativement rapide à de nouveaux types de documents. Les documents sur lesquels nous avons déjà travaillé sont les suivants :

- partitions d'orchestre à portées polyphoniques ;
- formules mathématiques ;
- tableaux/formulaires, avec compréhension de l'organisation hiérarchique ;
- formulaires anciens altérés.

La méthode DMOS s'applique à de nombreux types de documents structurés : formules chimiques, cartouches de plans, pages de journaux, factures, bons de livraison, etc.

En outre, la description grammaticale d'un type de document (grâce au langage EPF) peut être générale afin de reconnaître une classe de documents (les cartouches de plans par exemple) ou bien spécifique pour compenser une absence d'informations (un type de cartouche précis mais très souvent altéré).

Bien entendu, il est possible de fusionner des descriptions grammaticales de documents. Ainsi, nous pouvons très simplement construire un système capable de traiter des tableaux contenant des formules mathématiques.

Les utilisations de ces systèmes de reconnaissances sont également nombreuses (avec l'intégration de la reconnaissance de l'écriture manuscrite) :

- rétro-conversion de documents papier pour éviter une saisie manuelle afin d'obtenir une version électronique exploitable. Cette version électronique peut ensuite servir par exemple à une réédition, à la production de Braille, etc.
- identification de types de documents pour, par exemple, effectuer une gestion automatique de fax ;
- extraction automatique d'index, d'annotations pour permettre ensuite l'accès à des images de documents ;
- localisation de zones particulières dans une image de documents pour en faciliter également l'accès (section 4.2).

4.2 Annotations et méta-données

4.2.1 Annotation de documents numériques

Nos travaux de recherche sur les documents numériques et le web sémantique sont appliqués à l'annotation de documents anciens numérisés contenant du texte manuscrit ou imprimé.

Les expérimentations menées actuellement dans le projet sur la rétroconversion de documents anciens montrent qu'il est possible d'utiliser des techniques de reconnaissance génériques pour localiser des zones caractéristiques dans certaines images de documents. L'utilisation d'une représentation en XML du contenu géométrique de ces images va permettre d'utiliser le langage de transformation XSLT pour produire des documents XML structurés selon leur contenu sémantique.

4.2.2 Bibliothèque virtuelle de typographie

Si de nombreux ouvrages anciens sont souvent cités dans des bibliographies (par exemple en histoire de la typographie), il est en revanche fréquent qu'il soit très difficile de trouver ces ouvrages, ou des rééditions. Il nous a donc semblé intéressant de définir un formalisme permettant la recherche sur le web d'ouvrages (bibliothèques où ils se trouvent, rééditions – parfois électroniques –, accessibilité, etc.) à partir de métadonnées (notices des bibliothèques, auteurs, sites, etc.). Le formalisme prévu est

celui des structures de trait typées : les structures de traits constituent l'entité de base dans de nombreux formalismes linguistiques récents. Dans ces formalismes, une structure de traits est un graphe connexe, orienté, dont les arcs sont étiquetés par des traits et dont les nœuds sont décorés par des types. Les types sont organisés selon une hiérarchie d'héritage.

4.3 Reconnaisseurs d'écriture manuscrite

Les applications en reconnaissance d'écriture manuscrite connaissent un nouvel essor à la fois dans les domaines du traitement automatique de document papier (*reconnaissance hors-ligne*) et dans les nouvelles modalités d'interactions homme machine basées sur l'utilisation d'un stylet couplé à un écran tactile (*reconnaissance en-ligne*).

En *reconnaissance hors-ligne*, les besoins industriels sont aujourd'hui considérables, et par conséquent l'attente de systèmes de reconnaissance robustes et fiables est très grande. L'étendue des applications est vaste ; elle recouvre notamment les problèmes de traitement automatique de l'écriture contenu dans tout type de document papier, notamment les bordereaux de livraison, les feuilles de sécurité sociale, ou encore les télécopies. Dans ce domaine, nos efforts ont porté plus particulièrement sur la garantie d'une grande fiabilité des résultats en cas de traitement automatique effectif de l'écriture pour éviter tout risque d'erreur.

En *reconnaissance en-ligne* de l'écriture manuscrite, un marché très important s'est ouvert suite au développement économique récent des produits suivants :

- ordinateurs de poche (Palm, PocketPC),
- ordinateurs tablettes (ordinateurs sans clavier munis d'un écran LCD tactile),
- téléphones mobiles de nouvelle génération (« smart phone ») qui intègrent les modalités d'interaction des ordinateurs de poche (écran LCD tactile) et leurs principales fonctionnalités (agenda, éditeur de texte, accès à Internet, etc.).

Le logiciel RecifCar est aujourd'hui intégrable à ce type de produit (cf. paragraphes 5.1 et 6.6).

4.4 Interfaces orientées stylo

Les interfaces orientées stylo constituent une nouvelle modalité d'interaction entre l'homme et la machine. Elles sont utilisées sur des produits possédant des écrans tactiles ou des tablettes à digitaliser. Ces interfaces offrent à l'utilisateur la possibilité d'effectuer à main levée à l'aide d'un stylet :

- des saisies naturelles (il peut directement écrire des mots, des caractères ou des symboles sur l'écran tactile) ;
- des dessins de graphe, de croquis, d'organigrammes, etc. ;
- ou encore des commandes gestuelles intuitives (suppression, insertion, déplacement, etc.).

Ces différents tracés sont alors reconnus, interprétés et exécutés par la machine en fonction de leur nature et de leur contexte.

L'étendue des applications basées sur des interfaces orientées stylo est très vaste. On peut tout d'abord mentionner les produits dont la communication, pour des questions d'encombrement et de convivialité, s'appuie uniquement sur un écran tactile : ordinateur de poche (Palm, PocketPC), téléphone mobile de nouvelle génération (« smart phone »), ordinateur tablette. On peut de plus citer les applications qui, de par leur nature, profiteront pleinement des avantages d'une communication basée sur une saisie de symboles ou des graphiques à main levée : CAO, édition électronique, édition

musicale, édition de formule mathématique, etc.

Dans ce domaine, nous avons conçu et réalisé plusieurs prototypes qui ont pour objectif de mettre en évidence les possibilités et la convivialité de ces nouvelles modalités d'interaction homme-machine : éditeur interactif de partitions musicales, éditeur de graphes dessinés à main levée, éditeur de page WEB basé sur la reconnaissance de gestes graphiques.

5 Logiciels

5.1 RESIF, Reconnaissance d'Écriture Manuscrite en-ligne par Systèmes d'Inférence Floue

Participant : Éric Anquetil.

Mots clés : reconnaissance, écriture manuscrite, Palm, PocketPC, smartphone.

RESIF est un ensemble de composants logiciels pour l'analyse, la modélisation et la reconnaissance automatique de l'écriture manuscrite en-ligne. Les quatre composants logiciels principaux sont :

- ResifApp : Système de modélisation automatique par apprentissage des caractères manuscrits ;
- ResifMot : Système d'analyse et de reconnaissance de mots et de lettres manuscrits ;
- ResifCar : Système d'analyse et de reconnaissance de caractères isolés (lettres, chiffres, symboles spéciaux) ;
- ResifSeg : Système de segmentation automatique supervisée de mots en lettres.

Ce logiciel a été déposé dans son intégralité à l'APP avec comme auteurs E. Anquetil et G. Lorette. Le composant logiciel ResifCar est actuellement en phase finale d'optimisation (cf. section 6.6) pour permettre une intégration sur des produits ne comportant qu'un processeur de faible puissance : téléphone mobile de nouvelle génération (*smart phone*), Palm, PocketPC, etc.

ResifCar permet à l'utilisateur d'écrire directement des caractères cursifs (lettres, chiffres, symboles) sur un écran tactile, en conservant son écriture naturelle et sans période d'apprentissage.

5.2 DocRead : générateur automatique de systèmes de reconnaissance de documents structurés

Participant : Bertrand Coüasnon.

Mots clés : reconnaissance, documents structurés, partitions musicales, formules mathématiques, tableaux, formulaires, archives.

DocRead est un générateur automatique de systèmes de reconnaissance de documents structurés, développé grâce à la méthode DMOS. Celui-ci est constitué d'un compilateur du langage EPF (permettant de décrire un document à l'aide d'une grammaire), d'un module d'analyse lié à ce langage, d'un module de vision précoce (binarisation et extraction de segments) et d'un classifieur ayant des capacités de rejet.

Ce générateur nous permet une adaptation rapide à un nouveau type de document. En effet, il faut simplement définir une nouvelle grammaire (à l'aide d'EPF) qui décrit le nouveau type de document et, si nécessaire, il faut effectuer un nouvel apprentissage du classifieur pour lui permettre de reconnaître

de nouveaux symboles. Le système de reconnaissance adapté est alors automatiquement produit par compilation.

Grâce à ce générateur, nous avons défini un certain nombre de systèmes de reconnaissance :

- ScoRead : prototype de reconnaissance de partitions musicales ;
- MathRead : prototype de reconnaissance de formules mathématiques ;
- TabRead : prototype de reconnaissance de structures tabulaires ;
- FormuRead : logiciel qui permet d'extraire automatiquement la structure de formulaires d'incorporation militaire du XIX^e siècle, malgré leur mauvaise qualité. Ce logiciel a été testé sur 5 268 images et a montré ses très bonnes capacités. Nous commençons actuellement le traitement de 100 000 images de documents des Archives de la Mayenne.

5.3 Bibliothèque d'extraction d'indices visuels

Participant : Jean Camillerapp.

Mots clés : Vision précoce, binarisation adaptative, vectorisation, filtrage de Kalman, squelettisation.

Il s'agit de la bibliothèque sur laquelle s'appuie le logiciel DocRead décrit ci-dessus.

Elle comporte différentes méthodes de binarisation adaptative, notamment en incorporant des opérations contours-régions. Certaines de ces méthodes peuvent fournir des résultats avec une résolution sub-pixelle.

Elle réalise l'extraction des segments rectilignes dans des images complexes en s'appuyant sur le filtrage de Kalman décrit dans 3.1.

Elle comporte également une méthode de squelettisation à partir d'images en niveaux de gris.

Bien qu'initialement optimisée pour le traitement des images de documents, cette bibliothèque peut être utilisée dans d'autres contextes de traitement d'images. Elle a ainsi été essayée pour repérer les limites d'un court de tennis dans des séquences d'images vidéo.

6 Résultats nouveaux

6.1 Application de la méthode DMOS à la reconnaissance de structures tabulaires

Participant : Bertrand Couïasnon.

Résumé : *L'aspect générique de la méthode DMOS de génération de systèmes de reconnaissance de document, a pu être validé sur la reconnaissance de structures tabulaires récursives.*

Grâce à la méthode DMOS[16], nous avons pu définir différentes grammaires EPF pour produire par simple compilation, différents systèmes de reconnaissance de partitions musicales, de formules mathématiques [17], de registres militaires du XIX^e siècle [15].

Afin de valider l'aspect générique de la méthode DMOS, nous avons travaillé sur la reconnaissance de structures tabulaires. Nous avons défini une description grammaticale (à l'aide du langage EPF) d'un tableau-formulaire constitué de filets (segments), quels que soient le nombre de colonnes, le nombre de lignes et la dimension des cases.

Cette description commence par la définition de la plus grande structure tabulaire détectable dans un tableau, puis la définition du tableau lui-même. De manière récursive, à l'intérieur de chaque case, un nouveau tableau est recherché. Cette description récursive permet au système produit de reconnaître l'organisation hiérarchique d'un tableau-formulaire quel que soit son placement dans un document.

Certains logiciels du commerce de reconnaissance de documents peuvent traiter des tableaux, cependant, ils ne produisent que la présentation graphique de ces tableaux et ne peuvent détecter leur organisation hiérarchique. Or, cette dernière est primordiale pour pouvoir structurer et accéder aux données contenues dans un tableau.

La figure 7 affiche les cases détectées aux différents niveaux de profondeur de l'organisation hiérarchique du tableau.

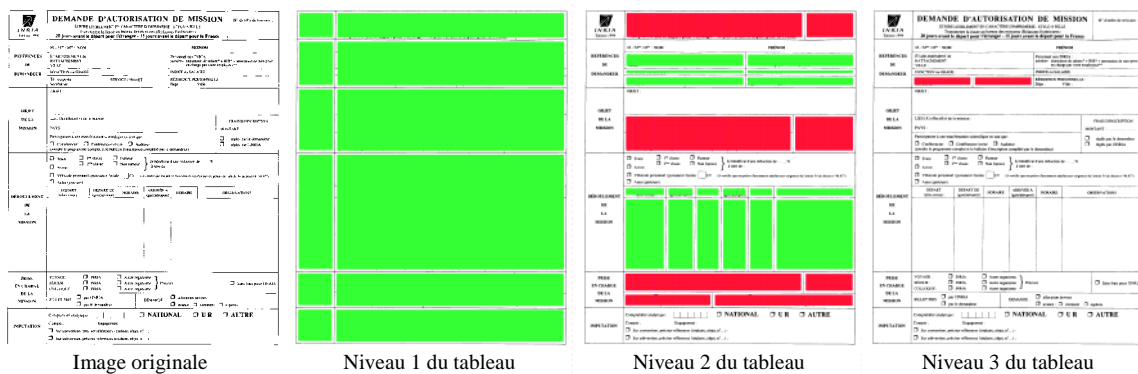


FIG. 7 – Exemple de la reconnaissance de la structure hiérarchique d'un tableau

Il nous reste à compléter ce travail par une phase de validation et par la définition d'une structure pouvant contenir l'organisation hiérarchique reconnue. Nous devons également travailler sur la reconnaissance de tableaux sans files.

6.2 Préparation de l'insertion de la reconnaissance de l'écriture manuscrite dans la méthode DMOS

Participants : Jean Camillerapp, Nicolas Briec.

Dans le traitement des documents anciens, il n'est actuellement pas envisageable d'employer partout des procédés de lecture automatique. Par contre, la méthode DMOS peut permettre de spécifier les parties de document qui doivent faire l'objet d'une reconnaissance des informations manuscrites, afin de préparer une indexation automatique.

Dans ce but, nous avons travaillé sur l'intégration et l'adaptation d'un certain nombre d'acquis antérieurs dans les fonctionnalités de la méthode DMOS. Celle-ci a été conçue, au départ, pour utiliser des images binaires. Il y a deux ans nous avons intégré une méthode de binarisation adaptative, comme première phase du traitement, mais sans toutefois tirer parti des localisations *sub-pixelles*.

Cette année, nous avons complètement repris l'extraction des segments de droite par filtrage de Kalman, en utilisant directement des images en niveaux de gris. La binarisation adaptative sert pour

l'amorçage des filtres et fournit à ceux-ci des mesures de position des traits noirs beaucoup plus précises

Cette approche diminue les effets de quantification spatiale (pas d'échantillonnage) introduits lors de la binarisation. Elle permet également de réinterpréter les niveaux de gris de l'image dans le prolongement de segments ayant un contraste très faible. On peut ainsi extraire des segments que l'œil perçoit mais qui n'avaient été que partiellement détectés lors de la phase de binarisation.

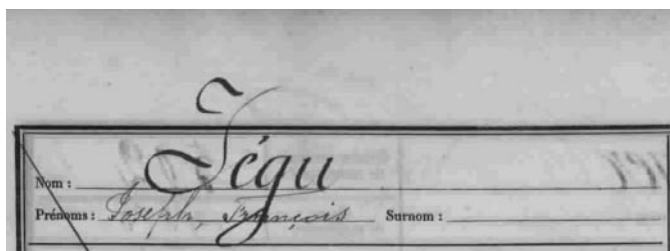


FIG. 8 – Case état civil localisée par DMOS

Ce travail sur le filtrage de Kalman a été l'occasion d'une réflexion sur l'influence des différents paramètres (matrices de covariance des modèles) intervenant dans ces filtres.

Nous avons ensuite abordé la segmentation de l'image dans les régions où le tracé déborde sur les parties préimprimées afin d'isoler le tracé et pouvoir le transmettre à un module de reconnaissance.

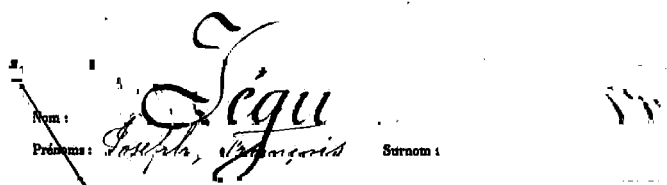


FIG. 9 – Séparation du tracé et du cadre

Nous avons également continué à explorer la voie de la reconstruction de la dynamique du tracé afin de lier reconnaissance hors-ligne et reconnaissance en-ligne. Pour cela, nous avons implanté un algorithme proposé par Kato et Yasuhara [KY99]. A l'usage, il s'avère que les images que nous avons à traiter ne satisfont pas les contraintes assez fortes que cet algorithme impose et nous n'en avons pas trouvé d'extension simple.

6.3 Annotations

Participants : Nicolas Briec, Bertrand Couasnon, Ivan Leplumey, Hélène Richy.

[KY99] Y. KATO, M. YASUHARA, « Recovery of Drawing Orders from Scanned Images of Multi-Strokes Handwriting », in : *International Conference on Document Analysis and Recognition 5*, p. 261–264, Bangalore, 1999.

Résumé : *A partir de la description en format XML de l'image de documents manuscrits, et de la description sémantique de ces images, un système d'annotation est en cours de mise au point et de validation sur des documents fournis par des services d'archives.*

Une première expérimentation porte sur des images de formulaires : des registres de matricules militaires datant du XIX^e siècle. Le découpage automatique concerne plus de 100 000 images de registres matricules militaires numérisés, dans le but de séparer données publiques et données protégées. Les images publiques seront ensuite mises en consultation sur Internet. La méthode de détection automatique de la structure de ces formulaires permet de découper automatiquement les images et de leur associer une typologie prédéfinie (état-civil, signalement, affectation, état des services, ...).

Une seconde expérimentation porte sur l'annotation coopérative de documents anciens. Ces derniers sont de deux types, les registres matricules cités ci-dessus, et les registres paroissiaux et d'état civil de deux cantons (un pour la Mayenne, et un autre pour l'Ille-et-Vilaine). En limitant la reconnaissance de caractère au contenu des zones les plus lisibles des registres matricules, il est possible d'extraire des images les informations nécessaires à la définition de méta-données en XML. Toutefois, pour les registres difficilement déchiffrables, un système d'annotation directe sera expérimenté. Dans le cas des informations d'état-civil la définition préalable d'une ontologie sur les liens de parenté peut permettre de créer automatiquement des liens entre les différentes personnes identifiées dans les actes (mariage, décès, naissance). Cette ontologie conçue comme un graphe RDF peut être décrite en XML.

6.4 Bibliothèque virtuelle de typographie

Participant : Jacques André.

Une version expérimentale est en cours de définition pour une « bibliothèque virtuelle de typographie » pour laquelle la collecte du fond a déjà démarré en relation avec l'Institut Culturel du Livre (Lyon), l'Institut de documentation et communication (UBS-Lorient) et la bibliothèque de l'Irisa. Voir <http://www.irisa.fr/faqtypo/BiViTy.html>.

6.5 Modélisation automatique des connaissances par systèmes d'inférence floue hiérarchisés

Participants : Nicolas Ragot, Éric Anquetil, Guy Lorette.

Résumé : *Dans le cadre de la modélisation automatique des connaissances, nous proposons une nouvelle approche par systèmes d'inférence floue hiérarchisés. Un arbre de décision flou à deux niveaux est utilisé comme structure élémentaire. Le premier niveau de l'arbre hiérarchise les connaissances par des partitionnements flous successifs afin d'établir une première discrimination des classes. Le deuxième niveau complète ce découpage en modélisant de manière intrinsèque chaque sous-classe issue du niveau précédent. Ces différents partitionnements sont formalisés par des systèmes d'inférence floue. Conjointement, le meilleur espace de représentation pour chaque partitionnement est déterminé par un algorithme génétique.*

Dans nos précédentes recherches, un classifieur a été développé pour la reconnaissance de l'écriture manuscrite en ligne [2]. Ce classifieur est basé sur la modélisation de chaque lettre par des sys-

tèmes d'inférence floue (SIF) (cf. section 3.3) organisés hiérarchiquement grâce à des connaissances a priori sur l'écriture. Cette modélisation hiérarchique des connaissances permet de rendre le processus de décision plus stable et plus robuste tout en conservant le caractère interprétable, modifiable et donc optimisable des SIF. Cette méthode donne de bons résultats sur la reconnaissance des lettres manuscrites. Cependant, l'organisation hiérarchique étant fondée sur des connaissances a priori, cette approche n'est pas directement applicable à d'autres problèmes de classification.

Nous proposons donc une nouvelle approche [18, 19] inspirée des travaux précédents avec cette fois un objectif de généralité permettant :

- d'extraire automatiquement des connaissances ;
- de les hiérarchiser automatiquement ;
- de sélectionner dynamiquement les connaissances pertinentes à mettre en jeu lors de la classification de nouveaux individus.

Cette méthode conduit à l'élaboration d'un système dont la structure de base repose sur les arbres de décision flous ^[Mar98]. L'arbre se compose de deux niveaux bien distincts appelés respectivement *niveau de pré-classification* et *niveau de classification*.

Le rôle du premier niveau consiste à trouver des sous-groupes d'individus ayant des caractéristiques communes stables et de les hiérarchiser de manière à effectuer une première discrimination des classes. Pour cela, l'extraction des connaissances est obtenue par des partitionnements flous successifs dans des espaces à n dimensions. Ceux-ci sont déterminés de manière non supervisée par un algorithme de classification non supervisée (CNS) de type *C-moyennes floues* ^[Bez81]. Un algorithme génétique (AG) détermine pour chaque partitionnement le meilleur sous-espace de caractéristiques optimisant la discrimination inter-classes.

Le deuxième niveau travaille de manière supervisée sur chacune des classes afin de modéliser les sous-classes issues des partitionnements précédents. Cette modélisation s'effectue au moyen d'un algorithme de CNS de type possibiliste ^[R.K94]. Pour les deux niveaux, la formalisation (des partitions et des sous-classes) est faite par des SIF formant des unités facilement interprétables et exploitables.

Après l'apprentissage, pour la classification de nouveaux individus, une base de règles floues est extraite dynamiquement de la structure arborescente. Ce SIF hiérarchique permet de n'utiliser que des connaissances pertinentes relativement à un individu donné.

L'originalité de l'approche (qualifiée de *2l-FDT* pour « 2 levels Fuzzy Decision Tree ») repose sur la collaboration des techniques suivantes :

- une hiérarchisation automatique des connaissances au moyen d'un arbre de décision flou basé sur un partitionnement et une modélisation en n dimensions par des algorithmes de CNS floue, une optimisation locale du choix de l'espace de représentation par AG et un double niveau de modélisation ;
- une inférence par un SIF hiérarchique déterminé dynamiquement pour chaque individu à classer.

Pour évaluer le *2l-FDT*, des expérimentations ont été conduites sur des problèmes bien connus de classification (les Iris de Fischer et les Formes d'ondes de Breiman) ainsi que sur un problème plus

[Mar98] C. MARSALA, *Apprentissage inductif en présence de données imprécises : construction et utilisation d'arbres de décision flous*, thèse de doctorat, université de Paris 6, janvier 1998.

[Bez81] J. BEZDEK, *Pattern Recognition with fuzzy objective function algorithms*, Plenum Press, 1981.

[R.K94] R. KRISHNAPURAM, « Generation of membership functions via possibilistic clustering », in: *IEEE World congress on computational intelligence*, p. 902–908, 1994.

complexe de reconnaissance de chiffres manuscrits².

Les tests sur ces bases permettent de situer le *2l-FDT* par rapport à d'autres approches de classification par arbre de décision et par réseau de neurones :

- le *C4.5* [Qui93], un arbre de décision classique très connu ;
- *Quest* [LS97], un arbre de décision plus récent ;
- *Salammbô* [Mar98], une approche récente par arbre de décision flou ;
- un réseau de neurones de type RBFN (Radial Basis Function Network) [ACD00].

Les résultats sont donnés dans les tableaux 1 et 2 où *Taille* est la taille moyenne de l'arbre (nombre de règles déductibles) et *Test* est le taux de classification sur des exemples inconnus (généralisation).

	Iris	
	<i>Taille</i>	<i>Test</i>
<i>C4.5</i>	8.5	95.2%
<i>Quest</i>	3.0	97.3%
<i>Salammbô</i>	4.0	96.0%
<i>2l-FDT</i>	3.0	97.4%

TAB. 1 – Résultats comparatifs sur la base des Iris de Fischer.

	Formes d'ondes	
	<i>Taille</i>	<i>Test</i>
<i>C4.5</i>	399	74.7%
<i>Quest</i>	23	84.7%
<i>2l-FDT</i>	24	84.1%
<i>RBFN</i>		82.3%

TAB. 2 – Résultats comparatifs sur les Formes d'ondes.

Ces résultats montrent que le *2l-FDT* (bien que inférieur à *Quest* sur les Formes d'ondes) est capable de fournir de bons résultats en généralisation, avec une bonne compacité grâce notamment au partitionnement en n dimensions.

²L'objectif ici n'est pas en soi la reconnaissance des chiffres manuscrits, mais uniquement de pouvoir comparer différentes approches de classification.

[Qui93] J. QUINLAN, *C4.5 Programs for machine learning*, Morgan Kaufmann, 1993.

[LS97] W.-Y. LOH, Y.-S. SHIH, « Split selection methods for classification trees », *Statistica Sinica* 7, 1997, p. 815–840.

[ACD00] E. ANQUETIL, B. COÛASNON, F. DAMBREVILLE, « A Symbol Classifier able to Reject Wrong Shapes for Document Recognition Systems », in: *Graphics Recognition, Recent Advances, Lecture Notes in Computer Science, 1941*, Springer, p. 209–218, 2000.

	Chiffres manuscrits	
	Taille	Test
<i>C4.5</i>	545	69.3%
<i>Quest</i>	64	70.1%
<i>2l-FDT</i>	80	84.5%
<i>RBFN</i>		83.7%

TAB. 3 – Résultats comparatifs sur une base de chiffres manuscrits.

6.6 Intégration du système de reconnaissance de caractères manuscrits ResifCar dans des téléphones mobiles

Participants : Éric Anquetil, Laurent Pasquer, Jean Camillerapp.

Le système de reconnaissance de caractères ResifCar³ est un sous-composant logiciel de Resif développé dans l'équipe (cf. section 5.1).

ResifCar est basé sur une modélisation des caractères par SIF hiérarchisés. Cette modélisation repose sur les fondements théoriques de la logique floue décrits brièvement au paragraphe 3.3, et s'appuie sur les connaissances structurelles de l'écriture manuscrite résumées au paragraphe 4.3. Elle est compacte, robuste et explicite (c'est à dire interprétable). Ces propriétés nous ont permis de travailler, en collaboration avec un industriel (ref. contrat : Insa0018), sur un projet de portage du système ResifCar sur des processeurs réduits utilisés dans des téléphones mobiles. Aujourd'hui, le portage et l'optimisation de ResifCar sont achevés et opérationnels. Le système a été embarqué sur des processeurs de type ARM cadencés à 13Mhz.

Le système ResifCar embarqué permet la reconnaissance des 26 lettres minuscules cursives de l'alphabet latin, des chiffres et les symboles spéciaux suivants : @, &, \$, £, ¥ et €.

Le système est capable d'absorber une grande variété de styles d'écriture en offrant la reconnaissance de plusieurs allographes d'un même symbole (cf. tableau 10).

Caractères	e	es	a	ra	es	e	t	a	es
d									
d									
d									
d									

FIG. 10 – Exemples d'allographes reconnus par ResifCar

L'évaluation du système a porté sur plus de 200 scripteurs et les taux de reconnaissance moyens

³logiciel de Reconnaissance d'Écriture manuscrite par Système d'Inférence Floue (SIF)

obtenus sont de l'ordre de 92% en première position et de 97% dans les trois premières positions.

Ce portage a mis en avant de manière concrète les avantages de la modélisation proposée en permettant une optimisation adaptée du système de reconnaissance sur un processeur à capacité réduite tout en conservant de très bonnes performances de reconnaissance. L'intégration de ResifCar sur des téléphones mobiles offre à l'utilisateur la possibilité d'écrire directement ses commandes ou messages sur l'écran tactile du téléphone, en conservant son écriture naturelle et sans période d'apprentissage.

6.7 Expérimentation sur l'analyse en temps-réel de la structure d'une formule mathématique, utilisation de l'approche perceptive SPI

Participants : Laurent Pasquer, Éric Anquetil, Guy Lorette.

Dans l'optique d'une interface de saisie dynamique de formules mathématiques *via* une tablette électronique, nous avons construit un mécanisme d'interaction avec l'utilisateur basé sur une analyse en temps-réel de ses tracés.

Un prototype a été construit et expérimenté. L'interaction avec l'utilisateur se fait selon deux modalités complémentaires :

mode de saisie :

les symboles tracés sont analysés pour être identifiés. L'utilisateur peut continuer à écrire pendant cette identification. Lorsque le système a reconnu un tracé, le symbole correspondant est inséré dans la formule mathématique.

mode d'édition :

l'utilisateur peut à tout moment décider de modifier sa formule, soit en utilisant le classique menu d'édition, soit en utilisant le stylet électronique pour pointer les symboles et décrire les opérations d'édition.

Le mode de saisie s'appuie sur une prédiction dynamique des différents ajouts possibles dans la formule. Cette prédiction est basée sur une description lexicale bidimensionnelle des formules. La *figure 11* présente la description des quatre constructeurs de formules utilisés :

- la *fraction* regroupe deux formules en une seule,
- la *racine* regroupe trois formules, l'expression sous la racine, le radical (racine n-ième) et l'exposant éventuel,
- les constructeurs *triples* (\sum \int \prod) regroupent trois formules,
- enfin chaque symbole peut regrouper jusqu'à quatre formules, exposants et indices gauche et droit.

En sus de ces constructions, une formule est une concaténation de symboles (c'est le rôle des points noirs sur la figure).

En utilisant cette description, le système peut prédire les différents ajouts possibles à une formule existante (concaténation à gauche ou à droite de nouveaux symboles dans une formule ou création de sous-formules *via* les constructeurs, ...). Ces prédictions sont représentées par des zones rectangulaires dans lesquelles les nouveaux symboles pourront être tracés. L'utilisateur peut décider de voir ou non ces zones, leur affichage lui permet de se familiariser progressivement au mécanisme de prédiction. Une fois habitué, il n'aura plus besoin de les afficher pour savoir où tracer ses symboles. Seul ce mode de saisie a été implanté. La modalité d'édition a seulement été étudiée et spécifiée. Pour expérimenter l'interface de saisie, la classification des symboles tracés (leur reconnaissance) a été simulée à l'aide

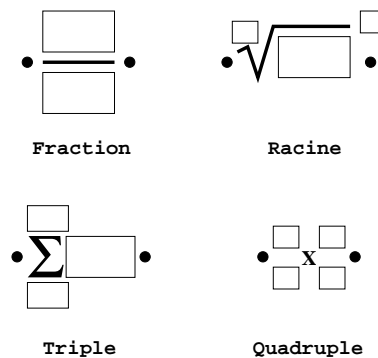


FIG. 11 – Les différents modèles de prédiction syntaxique

d'un menu (l'utilisateur sélectionne le symbole qu'il a tracé). L'expérimentation de cette interface sur quelques utilisateurs a montré l'intérêt de cette approche pour la saisie de formules mathématiques. Le système prédictif permet à l'interface d'afficher à l'utilisateur *en temps réel* les différentes possibilités d'insertion ainsi que le résultat de l'interprétation de sa saisie. Si cette interprétation lui semble incorrecte, l'utilisateur pourra (lorsque le mode d'édition sera implanté) à tout moment l'annuler ou la modifier *via* quelques gestes simples.

Les résultats obtenus nous ayant semblé suffisamment intéressants, nous avons décidé de poursuivre les recherches sur la prédiction dynamique dirigée par une description lexicale bidimensionnelle. Pour rendre *générique* et performante cette prédiction, nous avons décidé d'étendre le modèle SPI (*Système de Perception et d'Interprétation*) précédemment expérimenté sur l'interprétation de mots, vers l'interprétation de *structures graphiques complexes*. Le modèle SPI s'appuie maintenant sur une prédiction dynamique des modifications possibles des structures graphiques, accompagnée d'une modélisation précise de ces structures. Pour ce faire, SPI utilise une description grammaticale de la structure graphique au format XML. Cette description est duale :

- un ensemble de règles précise l'*organisation syntaxique* de la structure à reconnaître ;
- à chaque règle sont associées des contraintes lexicales bidimensionnelles qui servent à la fois à *prédire* les prochaines saisies ainsi qu'à les *valider*.

Deux expérimentations sont actuellement menées sur cette extension du modèle SPI :

- la reconnaissance de formules mathématiques simples,
- la reconnaissance de partitions musicales monodiques.

Ces deux expérimentations sont dans des domaines suffisamment différents pour pouvoir ainsi valider l'aspect *générique* de l'approche utilisée. Une description grammaticale doit être construite pour chacune, la seule difficulté étant de décrire l'enchaînement graphique des structures. La première expérimentation a déjà abouti à un prototype d'interface. Il reste maintenant à terminer l'expérimentation sur la reconnaissance de partitions musicales, ainsi qu'à ajouter la modalité d'édition dans cette interface. Les résultats actuels montrent déjà l'intérêt d'une telle approche : la séparation entre le moteur utilisé par l'interface (le modèle SPI) et la description de la structure graphique nous permettra de réutiliser cette interface *telle quelle* sur d'autres structures graphiques, simplement en changeant la description utilisée.

7 Contrats industriels (nationaux, européens et internationaux)

7.1 Portage du logiciel ResifCar de reconnaissance de caractères manuscrits

Participants : Eric Anquetil, Laurent Pasquer, Jean Camillerapp.

- Référence du contrat : INSA 0018,
- Période : octobre 2000 à juin 2001.

Ce contrat de collaboration avec un industriel a permis de porter le logiciel ResifCar dans des téléphones mobiles basés sur un processeur à capacité limitée. L'objectif est de permettre à l'utilisateur d'écrire ses messages ou ses commandes directement sur l'écran tactile du téléphone en utilisant son écriture naturelle.

7.2 Annotation coopérative des documents d'archives

Participants : Bertrand Coüasnon, Ivan Leplumey, Hélène Richy, Nicolas Briec.

Les relations que nous avons avec les Archives départementales de la Mayenne et celles d'Ille et Vilaine, nous ont permis de présenter à la région un projet fédérateur sur un travail coopératif d'annotation des documents des archives à travers le web.

- Numéro de la convention : INSA 1023,
- Période : octobre 2001 à octobre 2002.

L'équipe IMADOC et les Archives de la Mayenne et de l'Ille-et-Vilaine travaillent au sein d'un projet Mégalis, piloté par les régions Bretagne et Pays de la Loire. Le travail effectué dans le cadre de ce contrat est décrit dans la section 6.3.

8 Actions régionales, nationales et internationales

8.1 Actions régionales

Nous avons établi une convention, sans incidence financière, avec la société VisionObject de Nantes pour l'utilisation d'un logiciel développé conjointement pour l'enseignement et la recherche.

Jacques André fait partie du conseil scientifique du *Musée des transmissions et de l'informatique* de l'Armée de terre (ESAT, Rennes).

Jean Camillerapp fait partie des coordinateurs régionaux d'expertise des conventions CIFRE auprès de l'ANVAR.

8.2 Actions nationales

8.2.1 ICL

Jacques André collabore avec l'ICL (*Institut Culturel du Livre*, Lyon) où il a collaboré à la page « Musées européens de typographie⁴ » et définit une *Bibliothèque virtuelle de typographie*⁵.

8.2.2 GdR

Laurent Pasquer et Guy Lorette participent au groupe de travail « 5.2-Écrit », thème 5 : Communication, du GDR-PRC I³ (Information, Interaction, Intelligence).

Jacques André et Hélène Richy participent au groupe de travail « 7.1-Documents Multi-Médias », thème 7 : Systèmes I³ du GDR I³.

8.2.3 Sociétés savantes

Jacques André est membre du conseil d'administration de l'association GUTenberg.

Jean Camillerapp, Guy Lorette, Eric Anquetil et Bertrand Couïasnon participent aux activités de l'association GRCE : «Groupe de Recherche en Communication Écrite ». Dans ce cadre les activités du projet ont été présentées aux journées de l'ASTI.

Guy Lorette est vice-président GRCE en charge des relations internationales.

Guy Lorette participe aux activités de l'association AFRIF : « Association Française de Reconnaissance et d'Interprétation de Formes ». Guy Lorette est membre du CA et représentant à l'IAPR.

8.2.4 Listes de discussion

Jacques André est responsable de la liste *typographie* hébergée à l'Irisa (sympa@irisa.fr) et archivée sous l'adresse : <https://www.irisa.fr/wws/info/typographie>; et de la liste *GTTS : Groupe de Travail sur la Typographie des Sciences* hébergée à l'Irisa (sympa@irisa.fr) et archivée sous l'adresse : <https://www.irisa.fr/wws/info/gtts>.

8.3 Actions internationales

Jacques André a participé de façon active à la version française⁶ du standard Unicode (norme ISO-10646).

Guy Lorette est responsable des relations internationales de l'IFSIC, membre de la CAI (commission des affaires internationales) de l'Université de Rennes 1 ainsi que le représentant français au *Governing Board* de l'IAPR (*International Association of Pattern Recognition*) et membre du *Constitutions and Bylaws Committee* de cette association.

⁴<http://histoire.enssib.fr/icl-forum/>

⁵<http://www.irisa.fr/faqtypo/BiViTy.html>

⁶<http://hapax.iquebec.com>

9 Diffusion de résultats

9.1 Animation de la Communauté scientifique

9.1.1 Comité éditorial de journaux

Jacques André est :

- co-rédacteur en chef de la revue *Documents numériques* des éditions Hermès, qui aura publié en 2001 :
 - « L'archivage » (vol. 4, n°3-4, 374 pages)
 - « Dossiers numériques » (vol. 5, n°3-4, en préparation) ;
- rédacteur en chef des *Cahiers GUTenberg* qui aura publié en 2001 trois cahiers :
 - « Le document au XXI^e siècle », mai 2001 (224 pages)
 - « Metapost et le tracé de graphiques » (en préparation)
 - « Vignettes, pavages et ornements » (en préparation) ;
- responsable de la rubrique « Documents numériques » de l'encyclopédie *Techniques de l'ingénieur – Informatique* ;
- membre de l'*Editorial board* de la *Typography series* de Addison-Wesley.

Guy Lorette est membre du comité de rédaction des revues :

- *Pattern Recognition*,
- *IJDAR—International Journal of Document Analysis and Recognition*.

9.1.2 Comités de programme de conférences

Jacques André a été ou est membre du comité de programme de GUT2001 (Metz, mai 2001), GUT2002 (Genève), CIDE'01 (Toulouse, 2001), CFD (Tunis, 2002).

Guy Lorette a été ou est membre des comités de programme de : ASTI'2001 (Paris, 2001), 6th IC-DAR (Seattle, 2001), RFIA'2002 (Angers, 2002), IWFHR'8 (Niagara on the Lake, 2002), CFD (Tunis, 2002).

Jean Camillerapp est membre du comité de programme de CFD (Tunis, 2002).

9.2 Enseignement

9.2.1 Enseignement universitaire

L'équipe étant en grande partie constituée par des enseignants chercheurs, ceux-ci sont très impliqués dans les activités d'enseignement, mais comme la plupart ne sont pas rattachées à ce thème de recherche, elles ne sont pas citées ici.

DEA d'informatique, à l'IFSIC, Université de Rennes1, participation de Guy Lorette (cours perception, reconnaissance et interprétation d'images).

Jacques André assure un enseignement de typographie numérique en 5^e année de l'École régionale des Beaux-Arts à Rennes et est membre du « jury » du Diplôme supérieur de création typographique à l'École Supérieure Estienne à Paris.

H. Richy est responsable du DESS Méthodes Informatiques et Technologies de l'Information et de la Communication (Mitic) à l'IFSIC, Université de Rennes 1. Ce DESS accueille une vingtaine d'étudiants depuis septembre 2001.

H. Richy a donné un cours sur les documents numériques en DESS Méthodes Informatiques et Technologies de l'Information et de la Communication, à l'IFSIC, Université de Rennes 1.

Stages effectués au sein du projet :

- Nicolas Briec et Sébastien Decaix (DEA d'informatique, IFSIC).
- Grégory Maitralain et Joseph Thibaut (stage d'été, INSA).

9.2.2 Autres enseignements

Jacques André a organisé une série de cours sur XML pour le service de formation permanente de l'Université de Rennes 1.

Hélène Richy a donné un cours sur « XML and Electronic Documents on the Web » à l'ITI (Le Caire, Égypte), du 9 au 16 mars 2001.

Participation d'Éric Anquetil et de Guy Lorette à la 3^e année de la filière « Systèmes d'information statistique » de l'ENSAI (cours : reconnaissance à l'aide de réseaux de neurones, reconnaissance de l'écrit manuscrit).

9.3 Participation à des colloques, séminaires, invitations

Jacques André a été invité une semaine dans le groupe *Mule (Multilingual Emacs)* à Tsukuba (Japon) et donné une conférence invitée (*Dynamic fonts*) lors de la conférence m17n2001 de *Multilingual Computer Science* à Tokyo.

Hélène Richy a été invitée au colloque « Textualité et nouvelles technologies », organisé par la revue *éc/artS*, en collaboration avec le Bureau de l'Écrit du Ministère français des affaires étrangères et l'AFAA, dans le cadre de « France au Québec, la saison », à Montréal, du 22 au 25 octobre 2001, pour un exposé sur le « Traitement électronique des documents anciens ».

10 Bibliographie

Ouvrages et articles de référence de l'équipe

- [1] J. ANDRÉ, M.-A. CHABIN (éditeurs), *Les documents anciens*, Numéro spécial de *Document numérique*, vol. 3, n°1-2, Hermès, 1999.
- [2] E. ANQUETIL, G. LORETTE, « Automatic generation of hierarchical fuzzy classification systems based on explicit fuzzy rules deduced from possibilistic clustering: application to on-line handwritten character recognition », in : *Proceedings of the International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems, (IPMU'96)*, p. 259–264, Grenade, Espagne, 1996.
- [3] B. COÛASNON, J. CAMILLERAPP, « A way to separate knowledge from program in structured document analysis: application to optical music recognition », in : *Proceedings of the 3rd International Conference on Document Analysis and Recognition, (ICDAR'95)*, p. 1092–1097, Montréal, Canada, 1995.
- [4] V. P. D'ANDECY, J. CAMILLERAPP, I. LEPLUMEY, « Analyse de partitions musicales », *Traitement du Signal* 12(6), 1996, p. 653–661.

- [5] G. LORETTE, E. ANQUETIL, « Théorie des catastrophes, géométrie différentielle et segmentation de l'écriture cursive », in : *Actes du 4ème Colloque National sur l'Écrit et le Document, (CNED'96)*, p. 1–6, Nantes, France, 1996.
- [6] G. LORETTE, J.-P. CRETTEZ, « Reconnaissance de l'écriture manuscrite », in : *Traité informatique, Techniques de l'Ingénieur*, 1998, ch. H 1358, p. 1–15.
- [7] G. LORETTE, Y. LECOURTIER, « Is recognition and interpretation of handwritten texts a scene analysis problem? », in : *Proceedings of the 3rd International Workshop in Frontiers in Handwriting Recognition, (IWFHR3)*, p. 184–196, Buffalo, USA, 1993.
- [8] G. LORETTE, « Handwriting recognition or reading? What is the situation at the dawn of the 3rd millennium? », *International Journal on Document Analysis and Recognition* 2, 1, 1999, p. 2–12.
- [9] J. PETTIER, J. CAMILLERAPP, « Script representation by a generalized skeleton », in : *Proceedings of the 2nd International Conference on Document Analysis and Recognition, (ICDAR'93)*, p. 850–853, Tsukuba, Japon, 1993.

Articles et chapitres de livre

- [10] J. ANDRÉ, « Chronique typographique », *Lettre GUTenberg*, 20, octobre 2001.
- [11] J. ANDRÉ, « Normes et codage des caractères », in : *Traité informatique, Techniques de l'Ingénieur*, décembre 2001, ch. H 7028, p. 1–36, Seconde édition augmentée et corrigée.
- [12] H. RICHY, « Métadonnées et documents numériques », in : *Traité informatique, Techniques de l'Ingénieur*, 2001, à paraître.
- [13] H. RICHY, « Traitement électronique des documents anciens », *Ec/artS*, 3, 2001, à paraître.

Communications à des congrès, colloques, etc.

- [14] J. ANDRÉ, « De Babel Web à Orléans : plurilinguisme, Internet et recherche universitaire », in : *Actes du congrès LexiPraxis 2000*, p. 27–28, Paris, décembre 2000.
- [15] B. COÛASNON, L. PASQUER, « A real-world evaluation of a generic document recognition method applied to a military form of the 19th century », in : *ICDAR, International Conference on Document Analysis and Recognition*, Seattle, USA, septembre 2001.
- [16] B. COÛASNON, « DMOS: A generic document recognition method, application to an automatic generator of musical scores, mathematical formulae and table structures recognition systems », in : *ICDAR, International Conference on Document Analysis and Recognition*, Seattle, USA, septembre 2001.
- [17] P. GARCIA, B. COÛASNON, « Using a generic document recognition method for mathematical formulae recognition », in : *GREC, IAPR International Workshop on Graphics Recognition*, Kingston, Canada, septembre 2001.
- [18] N. RAGOT, E. ANQUETIL, « Modélisation automatique des connaissances par systèmes d'inférence floue hiérarchisés », in : *Actes du congrès LFA'01, Rencontres Francophones sur la Logique Floue et ses Applications*, Mons, novembre 2001.
- [19] N. RAGOT, E. ANQUETIL, « A new hybrid learning method for fuzzy decision trees », in : *Proceedings of the 10th IEEE International Conference on Fuzzy Systems*, Melbourne, décembre 2001.