

## **Rapport d'activité scientifique 1997**

### **Projet Imadoc**

Interprétation et Reconnaissance d'Images et de Documents

Thème Inria 3

Interaction homme-machine, images, données, connaissances

(Version du 26 janvier 1998)

## Table des matières

1	Composition de l'équipe . . . . .	3
2	Présentation générale et objectifs . . . . .	3
	2.1 Contexte . . . . .	3
	2.2 Objectifs . . . . .	4
	2.3 Méthodologies . . . . .	6
	2.3.1 Stratégies d'analyse de documents . . . . .	6
	2.3.2 Modélisation de l'écriture manuscrite . . . . .	6
3	Fondements scientifiques . . . . .	7
	3.1 Segmentation par filtrage de Kalman . . . . .	7
	3.2 Reconnaissance des formes par systèmes d'inférence floue . . . . .	9
	3.3 Introduction des connaissances sous formes grammaticales . . . . .	10
4	Domaines d'applications . . . . .	11
	4.1 Écriture manuscrite . . . . .	11
	4.2 Analyse de partitions musicales . . . . .	12
5	Logiciels . . . . .	13
6	Résultats nouveaux . . . . .	14
	6.1 Analyse et reconnaissance du tracé des mots . . . . .	14
	6.2 Unification de la reconnaissance en-ligne et hors-ligne de l'écriture . . . . .	15
	6.3 Reconnaissance des symboles musicaux . . . . .	17
	6.4 Intégration de connaissance lexicale . . . . .	19
	6.5 Organisation de dictionnaire . . . . .	19
	6.6 Interface homme machine orientée stylo . . . . .	21
7	Actions industrielles . . . . .	21
	7.1 Convention SRTP : étude et traitement d'images de documents en niveaux de gris, CDP 510227-3 . . . . .	21
8	Actions régionales, nationales et internationales . . . . .	22
	8.1 Actions régionales . . . . .	22
9	Diffusion de résultats . . . . .	22
	9.1 Animation de la communauté scientifique . . . . .	22
	9.2 Enseignement universitaire . . . . .	23
	9.3 Participation à des colloques, séminaires, invitations . . . . .	23
10	Bibliographie . . . . .	23

# Projet Imadoc

## Interprétation et Reconnaissance d'Images et de Documents

---

**Localisation :** *Rennes*

**Mots-clés :** analyse de document, apprentissage automatique, interface homme-machine, interprétation d'image, interprétation de signal.

## 1 Composition de l'équipe

### Responsable scientifique

Jean Camillerapp, professeur, Insa

### Assistante de projet

Edith Blin, TR, Inria

### Personnel INSA

Éric Anquetil, bourse MENESR, puis maître de conférences depuis le 1er septembre 1997

Bertrand Coüasnon, maître de conférences

Ivan Leplumey, maître de conférences

### Personnel Université de Rennes 1

Guy Lorette, professeur

Gildas Ménier, Ater, jusqu'au 1er septembre 1997

Charles Quéguiner, maître de conférences

### Chercheur doctorant

Laurent Pasquer, bourse MENRT

## 2 Présentation générale et objectifs

### 2.1 Contexte

En dépit d'annonces insistantes et répétées de l'avènement d'une société *sans papier*, la multitude de documents papiers existants pose encore d'énormes problèmes de transfert sur un support électronique.

Le document papier, une fois mis sous forme électronique, permet une recherche par le contenu, un transport très rapide, un archivage et une gestion beaucoup plus aisée. Par contre, le document sous forme papier permet, *in-fine*, à l'utilisateur de disposer d'une plus grande autonomie, d'une possibilité de lecture détaillée plus confortable et de possibilités d'annotations plus grandes.

Ces deux formes de représentation des documents sont complémentaires, et il est nécessaire de pouvoir passer facilement d'une forme à une autre. S'il s'agissait, soit d'un simple passage du document papier à une image bitmap, soit d'imprimer le contenu de l'image d'un document électronique, le problème serait trivial. En réalité, les problèmes scientifiques sous-jacents sont complexes car l'objectif est de passer, soit d'un document conceptuel à sa réalisation définitive (*synthèse de document*), soit du document papier à son interprétation sémantique (*analyse de document*), [6], [12], [16].

## 2.2 Objectifs

Les recherches menées au sein du projet IMADOC concernent *l'écrit et le document* sous toutes leurs formes (manuscrit, imprimé, image, graphique, multimédia, etc.) ainsi que les activités qui y sont liées, notamment la production de nouveaux documents hypertextes ou multimédia interactifs, la transformation sous forme électronique élaborée de documents papier existants et leur traitement « intelligent » ainsi que l'*Interaction Homme-Document (I.H.D.)*. De manière plus générale, les centres d'intérêt du projet IMADOC touchent à la *communication écrite* sous un triple aspect : *synthèse de documents*, *analyse de documents*, *interaction homme-document* (cf. *figure 1*).

### **Synthèse de documents : « de la sémantique à la génération finale du document électronique »**

Elle concerne la production automatique de nouveaux documents ainsi que la génération dynamique, adaptative, on dit encore *interactive*, de documents électroniques. La synthèse de documents étudie des modèles de documents qui rendent compte à la fois de leur organisation logique ou abstraite (*structure logique*) ainsi que de leur présentation graphique et de leur contenu (*structure physique*). Elle s'étend à l'étude de la structure et de la typographie de documents structurés [AFQ89] hypertextes et multimédia.

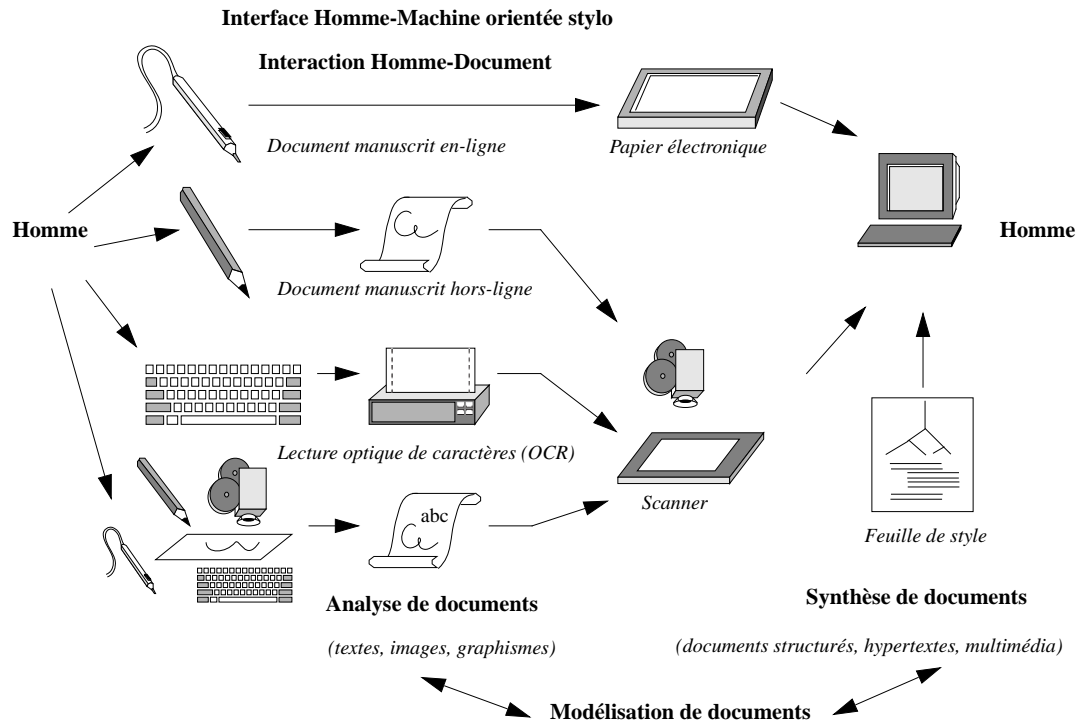
### **Analyse de documents : « du document papier à sa sémantique extraite de sa forme électronique »**

Afin d'aboutir à une interprétation fiable du contenu informationnel d'un document, il s'agit de faire en sorte que l'ordinateur soit capable d'analyser, de manière automatique, les signaux ou les images de documents [OK95], et de reconnaître tous les

---

[AFQ89] J. André, R. Furuta, and V. Quint. *Structured documents*. Cambridge University Press, 1989.

[OK95] L. O'Gorman and R. Kasturi. *Document Image Analysis*. IEEE Computer Society Press, 1995.

FIG. 1 – *Communication écrite*

éléments qui constituent un document et d'analyser leurs relations mutuelles tant physiques que logiques [BW97]. Pour aboutir à une interprétation finale correcte et robuste, il faut utiliser intelligemment plusieurs types de contextes.

De plus, la communication écrite étant une activité humaine, il faut être capable de gérer des connaissances imprécises, incertaines et incomplètes, ce qui crée un degré supplémentaire de complexité. Les recherches sur la reconnaissance et l'analyse du document manuscrit [Pla93] se subdivisent traditionnellement en deux catégories selon l'origine du document : *l'analyse statique* ou *hors-ligne*, dans laquelle le document est à l'origine sur support papier, puis est numérisé au moyen d'un scanner, et *l'analyse dynamique* ou *en-ligne* dans laquelle l'auteur du document est en interaction avec la machine au travers d'une interface orientée stylo.

### Interaction Homme-Document

Cette interaction se fait à l'aide de la conception de nouvelles interfaces homme-machine (I.H.M.) orientées stylo et de nouvelles modalités d'interaction, *papier électronique*, *reconnaissance du geste graphique*, qui sont plus naturelles, plus conviviales et plus ergonomiques que l'utilisation du clavier et de la souris pour l'annotation, la modification et la correction de documents.

[BW97] H. Bunke and P.S.P. Wang. *Handbook of Character Recognition and Document Analysis*. World Scientific, 1997.

[Pla93] R. Plamondon. Special issue on handwriting processing and recognition. *Pattern Recognition*, 26(3), 1993.

## 2.3 Méthodologies

### 2.3.1 Stratégies d'analyse de documents

Si l'action de lire, de reconnaître ou d'interpréter, un document peut sembler simple et naturelle, il s'avère que cette activité est difficile à décrire et à formaliser sous la forme d'une application informatique, car cette action fait intervenir de manière subjective la gestion de plusieurs contextes d'analyse. Il semblerait d'ailleurs que dans ce domaine on soit en présence d'un paradoxe maintenant bien connu : *pour reconnaître il faut segmenter le signal d'entrée, mais pour bien segmenter il faut avoir reconnu.*

L'existence de ce paradoxe implique donc qu'on ne peut pas se restreindre à une analyse ascendante, c'est à dire du signal vers le document interprété, ou au contraire à une analyse complètement descendante. Les mécanismes de reconnaissance que nous allons utiliser vont donc enchaîner des actions de segmentation et des actions de regroupement en mettant en œuvre des stratégies de génération et de validation d'hypothèses. L'approche développée dans le projet consiste également à essayer de mettre en évidence les informations contextuelles utilisées et d'en obtenir une description informatique simple et bien séparée des procédures qui les utilisent.

La première segmentation est une segmentation physique. Elle s'appuie sur des propriétés simples du signal : composantes connexes d'une image binaire, points singuliers dans la trajectoire du stylo, ... Ces propriétés qui n'utilisent qu'une information contextuelle très pauvre et pratiquement indépendante de l'application permettent d'aboutir à une décomposition assez sûre du signal d'entrée, décomposition sur laquelle les étapes suivantes de regroupement et de segmentation logique pourront s'appuyer, sans remise en question majeure.

### 2.3.2 Modélisation de l'écriture manuscrite

Dans le modèle que nous proposons, le tracé d'un mot se décompose en un ensemble d'allographes<sup>1</sup> concaténés. Dans chaque allographe, on distingue les constituants associés au corps de l'allographe et les zones (amorce et terminaison) situées en début et en fin du tracé de l'allographe, qui composent la liaison entre deux allographes juxtaposés. Ces zones de liaison peuvent être invisibles et se réduire, dans certains cas, à un simple lever de stylet.

Dans une phase d'analyse plus fine, on extrait de chaque allographe les structures fondamentales (traits descendants significatifs), leur environnement graphique local dit contexte morphologique (second niveau de modélisation), et les zones de liaison (troisième niveau de modélisation) qui peuvent être, là aussi, invisibles et correspondre à un lever de stylet (cf. *figure 2*).

Cette décomposition reflète la modélisation des lettres manuscrites dont les paramètres sont établis à partir d'un apprentissage automatique. Plus précisément, la détection et la modélisation des différentes zones pertinentes de tracé (appelées primitives) sont établies par classification non supervisée de type *possibiliste* à partir

1. le mot allographe désigne l'un des tracés possibles d'une lettre

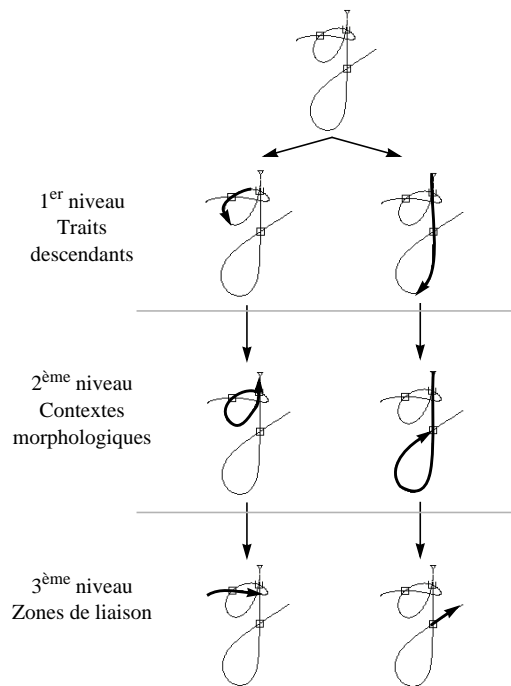


FIG. 2 – Illustration des 3 niveaux de modélisation pour la classe de la lettre *g*.

d'un ensemble de graphèmes extraits du tracé de lettres isolées. Ces graphèmes sont construits à partir de la détection de points d'ancrage visuel correspondant aux singularités du tracé : points de forte variation angulaire, intersections, *etc.* [5]. Les primitives ainsi modélisées sont représentées sous la forme de *prototypes* formalisés de manière qualitative par des *règles floues*. Chaque classe de lettres est alors décrite par un ensemble de *systèmes d'inférence floue* (cf. section 3.2) hiérarchisés d'après la *robustesse* des connaissances modélisées. On obtient ainsi une modélisation compacte et explicite de chaque classe de lettres [11]. Les différents constituants du tracé, mis en évidence dans ces trois niveaux de décomposition, à travers le modèle logique que nous venons de décrire, composent une base d'informations sur laquelle peut s'appuyer l'analyse logique du tracé d'un mot pour développer une méthodologie de reconnaissance.

### 3 Fondements scientifiques

#### 3.1 Segmentation par filtrage de Kalman

*Mots-clés* : filtrage de Kalman.

**Résumé** : *Les traitements d'images liés à la reconnaissance de documents ont souvent besoin d'extraire des structures filiformes, c'est à dire des objets dont la largeur est très faible par rapport à leur longueur. Dans les images de documents, ces structures ne sont généralement pas isolées mais se croisent ou se chevauchent, ce qui en complique la détection. Pour les identifier, nous avons proposé d'utiliser le formalisme*

*du filtrage de Kalman qui fournit un cadre cohérent permettant de bien faire la séparation entre la définition des caractéristiques des objets à reconnaître, et les procédures de segmentation.*

Les structures linéaires tiennent une grande place dans les documents, et elles sont généralement à tendance horizontale ou verticale, aussi peut-on se ramener à une analyse de l'image soit colonne par colonne, soit ligne par ligne. Les structures linéaires présentent une cohérence forte, il est donc possible de prédire leurs caractéristiques dans une colonne (ou dans une ligne) à partir des positions dans les colonnes précédentes. Nous utilisons pour cela le formalisme de Kalman dont on pourra trouver un exposé dans [Sor85] et une description de son application au traitement de documents dans [4].

Dans la colonne  $k$ , chaque structure linéaire est représentable par un vecteur d'état  $S(k)$  qui comporte l'ordonnée dans la colonne, l'épaisseur, la pente et éventuellement une luminosité pour les images en niveaux de gris.

Ce vecteur évolue en fonction de la colonne analysée  $k$  selon l'équation :

$$S(k + 1) = A.S(k) + W$$

dans laquelle  $A$  est la matrice caractérisant l'évolution du vecteur d'état  $S$  d'une colonne à l'autre et  $W$  un bruit de moyenne nulle.

Cette structure linéaire, si elle était isolée, produirait dans l'image une trace, qui se traduirait dans chaque colonne par un empan - ensemble de pixels connexes de la colonne - généralement plus sombre que son voisinage. Les caractéristiques de cet empan fournissent le vecteur de mesure  $X(k)$  qui est relié à l'état  $S(k)$  par l'équation :

$$X(k) = C.S(k) + N$$

dans laquelle  $C$  est la matrice permettant de passer de l'espace des états à celui des mesures et  $N$  le bruit de mesure.

Pour l'analyse de la colonne suivante, le filtrage de Kalman permet de prédire l'état  $\hat{S}(k + 1)$  et la mesure  $\hat{X}(k + 1)$ , ainsi que la matrice de covariance  $H(k + 1)$  de l'erreur de prédiction.

La véritable difficulté de cette approche ne réside pas dans ces équations, mais dans le processus d'affectation des mesures extraites d'une colonne de l'image avec les mesures prédites. En effet, les structures linéaires se coupent et se chevauchent ce qui perturbe les traces dans l'image, il n'y a donc pas une correspondance biunivoque entre les états caractérisant les structures linéaires et les mesures. Nous avons donc rajouté un mécanisme de contrôle qui, en utilisant les matrices de covariance  $H$ , est capable de détecter les empan qui proviennent de la fusion de plusieurs structures et qui associe les empan ou des parties d'empan avec les mesures estimées et donc avec les états. Ce mécanisme de contrôle détecte également l'apparition de nouvelles structures en créant un nouvel état quand il existe des mesures qui ne sont pas interprétables par les prédictions.

Ce formalisme a été appliqué pour :

- la détection des lignes de portées lors de l'analyse des partitions musicales ;



- la détection des traits délimitant des tableaux ;
- la détection et la suppression des lignes de base dans des images de chèques en niveaux de gris.

### 3.2 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 afin de faire face à la modélisation et à l'interprétation des formes, ou encore, à l'intégration et à la fusion de connaissances. 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 et par conséquent très difficilement optimisables.

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* [1]. 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** pas 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**  $Y_1 = b_{i1}$  **et**  $Y_2 = b_{i2}$  **et** ... **et**  $Y_u = b_{iu}$  **et** ... **et**  $Y_c = b_{ic}$ ,

avec :  $\forall k / 1 \leq k \leq C \ k \neq u, \ b_{ik} = 0, \ \text{et} \ b_{iu} = 1.$

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  $T$  est une  $T$ -norme symbolisée par le *et* utilisée dans la partie prémisse des règles,  $\beta_i$  représente la valeur d'activation de la règle  $R_i$ ,  $I$  est l'implication floue,  $\perp$  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 afin d'extraire et de qualifier un ensemble de propriétés pertinentes. Ces propriétés sont alors représentées directement par les sous-ensembles flous  $M_{ij}$  et par conséquent décrites de manière explicite par des règles floues dites prototypes au regard de leur caractère générique [13].

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

### 3.3 Introduction des connaissances sous formes grammaticales

*Mots-clés* : analyse structurelle, grammaire, *Definite Clause Grammar* ou DCG, segmentation, gestion des connaissances *a priori*.

**Résumé** : *Afin de tenter de résoudre les problèmes de segmentation que l'on rencontre en reconnaissance de documents, nous proposons une méthode baptisée DMOS permettant de formaliser grammaticalement la connaissance a priori, pour réaliser une segmentation contextuelle des documents à structure forte. L'introduction du contexte permet d'améliorer la qualité de la segmentation, donc de la reconnaissance.*

Dans le cadre de la reconnaissance optique de documents, la fiabilité est importante afin que l'utilisateur n'ait pas à relire l'ensemble du document pour y détecter et corriger d'éventuelles erreurs résiduelles.

Cette fiabilité peut s'obtenir, d'une part en améliorant la qualité de la reconnaissance, notamment en résolvant les problèmes de segmentation, et d'autre part en faisant détecter par le système lui-même les régions comportant des erreurs de reconnaissance. Pour atteindre ces deux objectifs, il faut utiliser la connaissance *a priori*, qui permet de régler certains problèmes de segmentation et permet également de modéliser la *redondance*, ce qui autorise une détection d'erreurs.

Nous proposons pour des documents à forte syntaxe dans lesquels des règles d'écriture peuvent être connues, une méthode baptisée DMOS (Description et MODification de la Segmentation), constituée d'un formalisme grammatical de position permettant de modéliser la connaissance, et d'un 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. En outre, cette méthode est chargée de faire appel à un classifieur pour reconnaître les symboles pouvant être assimilés à des caractères. Ainsi, des hypothèses de segmentation peuvent être produites grâce au contexte puis validées par le classifieur.

La méthode DMOS offre en plus l'avantage de séparer la connaissance (décrite sous la forme d'une grammaire) des outils de traitement [3], et de produire automatiquement l'analyseur par compilation de la grammaire. Cette décomposition facilite largement la maîtrise de l'introduction de connaissances complexes et l'adaptation à différents types de documents.

## 4 Domaines d'applications

### 4.1 Écriture manuscrite

*Mots-clés* : écriture manuscrite, en-ligne, hors-ligne.

**Résumé** : *Aujourd'hui, les taux de reconnaissance obtenus sur l'écriture manuscrite en-ligne et hors-ligne laissent entrevoir à court terme la possibilité de très nombreuses applications commerciales et industrielles.*

Pour envisager des applications « utilisables » par le plus grand nombre, qui sont donc économiquement intéressantes, il faut parvenir à des taux de reconnaissance pour l'écriture manuscrite très proches de ceux de l'être humain. Face à cet enjeu économique important, les premières études sur la reconnaissance d'écriture ont débuté très rapidement, sans avoir forcément conscience de la complexité des mécanismes intervenant dans la lecture car elle constitue pour l'homme une activité courante et naturelle. Ces premières études se sont très vite heurtées à cette complexité et ont engendré la conception de systèmes de reconnaissance d'écriture manuscrite aux performances médiocres. Ces premiers systèmes, trop rapidement commercialisés, ont alors eu tendance à discréditer les applications envisagées pour intégrer l'écriture comme une nouvelle modalité d'interaction possible entre l'Homme et la machine. Depuis quelques années, les études menées sur la reconnaissance d'écriture

connaissent un nouvel essor avec l'apparition de nouvelles méthodologies pour faire face aux problèmes les plus complexes de la reconnaissance des formes. Dans ce sens, nous avons cherché à développer un système de reconnaissance d'écriture manuscrite omni-scripteurs dont le caractère évolutif et les performances déjà obtenues permettent dès à présent d'envisager la conception de ce type d'application.

En ce qui concerne la *reconnaissance d'écriture en-ligne*, nous nous intéressons au développement d'interfaces orientées stylo permettant d'intégrer de manière homogène des commandes d'édition gestuelles à la saisie directe de commandes ou de textes manuscrits. Ce type d'interface graphique s'inscrit dans le cadre d'applications spécifiques où les fonctionnalités de pointage, de dessin, et de saisie limitée de texte sont fortement couplées. Ces applications recouvrent par exemple, les interfaces pour ordinateur portable, les logiciels de CAO, DAO, PAO ou encore les bloc-notes électroniques, ...

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 du montant des chèques, ou des adresses postales, mais aussi, d'une manière générale, le traitement de tout type de document papier, notamment les bordereaux de livraison, les feuilles de sécurité sociale, ou encore les télécopies.

Il existe déjà de nombreuses applications industrielles qui traitent de manière partielle ces différents problèmes. L'apport d'une reconnaissance fiable, même limitée à 30% des cas (rejet direct d'une grande partie des documents) peut être d'une très grande utilité dans les domaines d'application dans lesquels les quantités de documents à traiter sont considérables. C'est pourquoi, l'effort des systèmes de reconnaissance hors-ligne porte aujourd'hui sur la garantie d'une grande fiabilité des résultats en cas de traitement automatique effectif pour éviter tout risque d'erreur (même si pour cela il faut rejeter a priori un grand nombre de cas et donc les traiter manuellement). Dans cette optique, nos recherches ont pour objectif de concevoir un système de reconnaissance hors-ligne basé sur les concepts de la modélisation en-ligne de l'écriture que nous avons mis au point (cf section 6.2). Ce système devrait permettre, de par son caractère interprétable, de concilier de bons taux de reconnaissance tout en garantissant une grande fiabilité des résultats.

## 4.2 Analyse de partitions musicales

*Mots-clés* : édition de partitions musicales, Braille.

**Résumé** : *Les applications de la reconnaissance de partitions musicales sont importantes pour le développement de l'informatique musicale. Elles interviennent principalement dans le monde de l'édition de partitions et dans la constitution de bases musicales.*

La reconnaissance optique de partitions musicales est un domaine d'étude particulièrement intéressant car très représentatif des difficultés généralement rencontrées en reconnaissance de documents structurés.

En outre, ses applications directes sont relativement nombreuses. Elles s'inscrivent dans le cadre du développement de l'informatique musicale, qui manipule généralement l'information musicale à l'aide de la notation classique : la partition. Cependant, cette notation relativement complexe est difficile à saisir sous une forme numérique. De la même manière que la reconnaissance optique de caractères (*OCR*) pour les traitements de texte, la reconnaissance optique de partitions musicales permet d'éviter cette phase de saisie lente et fastidieuse.

Il est ainsi possible, grâce à un fichier au format MIDI (*Musical Interface for Digital Instruments*) généré par la phase de reconnaissance, de produire par exemple des documents sonores de très faible encombrement pour utilisation sur le Web.

Les musicologues ont également un besoin important en bases de données qui leur sont nécessaires pour analyser le style et la structure musicale (statistiques), indexer des bases de thèmes musicaux, ...

Contrairement aux applications précédentes, celles liées à l'édition musicale nécessitent de reconnaître l'ensemble des symboles rencontrés sur une partition. Lorsque le document analysé est une partition imprimée, les applications ayant trait à l'édition peuvent être par exemple :

- réaliser une nouvelle édition d'une édition ancienne (avec une mise en page différente), qui ne serait peut-être pas éditée pour des raisons de coûts de saisie trop élevés ;
- produire, à partir d'une partition d'orchestre, les partitions pour chaque instrumentiste et inversement ;
- extraire chacune des voix d'une partition à portées polyphoniques, afin de générer une partition avec une seule voix par portée (lorsque par exemple, chaque voix correspond à un chanteur) ;
- adapter une partition pour un autre instrument ;
- traduire une partition en Braille.

Nous travaillons sur la reconnaissance des partitions afin de produire une représentation pivot servant de point de départ aux applications présentées ci-dessus.

## 5 Logiciels

*Participants* : Éric Anquetil, Jean Camillerapp, Bertrand Coüasnon, Ivan Leplumey

**Résumé :** *Le projet a développé et maintient plusieurs chaînes de logiciels :*

- *un système de reconnaissance en ligne de mots tracés sur une tablette graphique ;*
- *des outils de binarisation adaptative et d'extraction du tracé ;*
- *un système de reconnaissance de partitions musicales ;*
- *un outil de détection des segments quasi rectilignes qui se croisent ou qui se chevauchent.*

*Cet ensemble d'outils permet d'évaluer rapidement les difficultés de traitement de nouveaux types de documents.*

## 6 Résultats nouveaux

### 6.1 Analyse et reconnaissance du tracé des mots

*Participants* : Éric Anquetil, Guy Lorette

*Mots-clés* : segmentation, mots manuscrits.

**Résumé** : *En suivant la modélisation générique de l'écriture manuscrite décrite en section 2.3.2, nous proposons une nouvelle approche analytique pour développer un système de reconnaissance omni-scripteur de mots manuscrits. Cette approche utilise directement la modélisation explicite des lettres décrite précédemment en s'appuyant notamment sur une nouvelle approche de segmentation structurelle du tracé des mots.*

Pour aborder la conception d'un système de reconnaissance de mots pouvant intégrer un contexte lexical de plusieurs dizaines de milliers de mots, il est nécessaire de considérer la forme d'un mot non pas comme une forme à part entière (*approche globale* de reconnaissance) mais comme la juxtaposition de plusieurs constituants élémentaires à reconnaître et à interpréter selon différents niveaux de contexte (*approche analytique*). Lorsque l'on envisage ce type d'approche, le problème de la segmentation *a priori* du tracé du mot en lettres apparaît immédiatement.

De nombreuses approches de segmentation s'appuient sur la détection directe de singularités susceptibles de correspondre à des points de segmentation du tracé du mot et nous avons, dans un premier temps, tenté d'absorber la variabilité de l'écriture par utilisation de Modèles de Markov Cachés (M.M.C.) [2].

Dans un second temps, nous avons mis en évidence dans l'écriture, l'existence de zones stables (*zones descendantes*) et de zones discriminantes [7], [8]. De plus, si l'on observe les liaisons inter-lettres, il s'avère qu'elles ne coïncident pas forcément avec une zone singulière. Elles correspondent, le plus souvent, à des zones régulières du tracé qu'il est difficile de délimiter de manière directe. C'est la raison pour laquelle, plutôt que d'opter pour une détection *a priori* de points de segmentation, nous avons proposé une nouvelle approche constructive de segmentation [14]. Elle s'appuie sur la localisation des structures d'ancrage qui correspondent aux zones descendantes stables modélisées pour chacune des classes de lettres. En effet, l'observation du premier niveau de modélisation (cf. *figure 2*) associé aux différentes classes de lettres montre que la *structure fondamentale* d'une lettre ne peut être composée, en principe, que d'une, deux ou trois *zones descendantes* stables. Par conséquent, il est possible de localiser de manière exhaustive l'ensemble des portions de tracé susceptibles de correspondre à une lettre pour en déduire ensuite progressivement les zones de liaison intra et inter-lettres. La *figure 3* illustre le graphe de juxtaposition des hypothèses de lettre organisées de manière générique, graphe qui a été obtenu sur le tracé du mot *try*.

Le processus de reconnaissance consiste alors à parcourir de manière exhaustive les différents chemins de ce graphe, en évaluant pour chacun d'eux une mesure de validité correspondant à la *fusion floue* des mesures d'adéquation de chacune des hypothèses de lettre avec les modèles et des mesures de cohérence du contexte liées

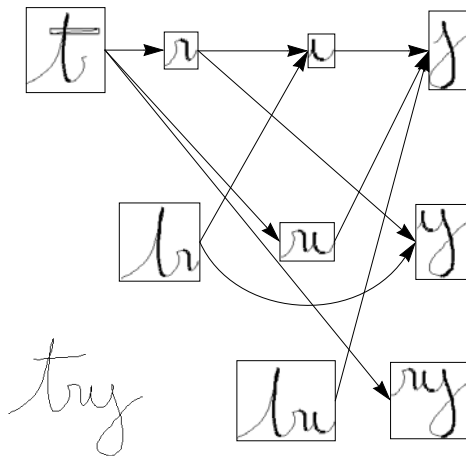


FIG. 3 – Graphe de juxtaposition des hypothèses de lettre.

à l'agencement spatial relatif de ces hypothèses de lettre. Le processus de reconnaissance s'achève par une phase de validation ou de correction lexicale.

L'apprentissage a été effectué sur une base de lettres isolées, tandis que les tests ont été réalisés sur un ensemble de 1128 mots différents écrits par sept scripteurs soit 7896 mots. Les scripteurs n'ont pas participé à la constitution de la base de lettres afin de valider les capacités omniscriteur du système. Dans ce cadre le taux des mots reconnus en première position s'échelonnent entre 87% et 73% en fonction de la taille du lexique des mots possibles, taille qui varie de 1000 à 25000 mots. Il se situe entre 94% et 85% si l'on considère les cinq premières positions [10].

## 6.2 Unification de la reconnaissance en-ligne et hors-ligne de l'écriture

*Participants* : Éric Anquetil, Jean Camillerapp, Charles Quéguiner

*Mots-clés* : écriture manuscrite.

**Résumé :** *Afin de tirer parti des performances intéressantes du reconnaissseur en-ligne développé dans le projet ainsi que pour mettre en évidence une synergie qui pourrait se dégager d'un rapprochement méthodologique entre le en-ligne et le hors-ligne, nous proposons une méthode d'extraction du tracé de l'image sous forme d'une liste ordonnée de points. Bien que prévu initialement pour un autre type de données, nous avons montré que le reconnaissseur était capable de fonctionner convenablement avec des données hors-ligne.*

Les travaux faits antérieurement dans le projet [9] sur l'extraction du tracé dans des images de documents mettent en évidence une ressemblance forte entre les zones régulières que détecte la partie traitement d'images et celles sur lesquelles s'appuie la reconnaissance en-ligne. Comme aujourd'hui la communauté reconnaît que les résultats obtenus en reconnaissance en-ligne dépassent très largement ceux obtenus en reconnaissance hors-ligne nous avons entrepris de développer une approche unifiée de la reconnaissance de l'écriture manuscrite selon le schéma décrit dans la figure 4.

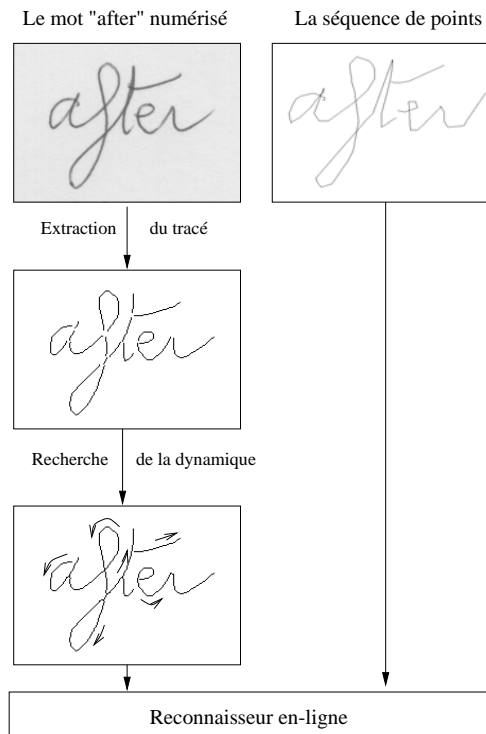


FIG. 4 – *Reconnaissance hors-ligne et en-ligne d'un mot manuscrit.*

Les algorithmes de traitement d'images commencent d'abord par isoler les régions régulières ou filiformes qu'ils représentent par la liste des points constituant leur axe médian. Cet algorithme est capable de définir à chaque extrémité d'une région régulière, une valeur de la luminosité qui constitue un bon seuil local de binarisation. Celui-ci permet d'isoler les régions singulières dont le rôle principal est de connecter les régions régulières. Dans ces régions de petite taille la position du tracé n'a pas besoin d'être déterminée avec beaucoup de précision puisque le reconnaiseur utilise peu les informations provenant de ces parties du tracé.

En utilisant des propriétés de l'écriture manuscrite, les régions régulières sont orientées et reconnectées à travers les régions singulières afin de retrouver l'ordre dans lequel les points ont été tracés. C'est cette liste de points qui est transmise à notre reconnaiseur.

Nous avons fait quelques essais dans lesquels il y avait acquisition du tracé en ligne puis numérisation par un scanner de la trace laissée sur le papier. Malgré un nombre de points presque cinq fois plus nombreux dans le cas du signal extrait de l'image que dans le cas du signal acquis directement en-ligne, nous avons pu montrer que le reconnaiseur pouvait fonctionner sans aucune modification car il continuait à extraire les primitives sur lesquelles il s'appuie (cf *figure 5*). Nous avons également pu faire des tests à partir de documents papier numérisés antérieurement.

Il est clair que ces premiers résultats demandent à être améliorés et validés sur des bases de tests plus importantes, mais la possibilité d'une connexion entre les deux approches nous paraît être acquise.



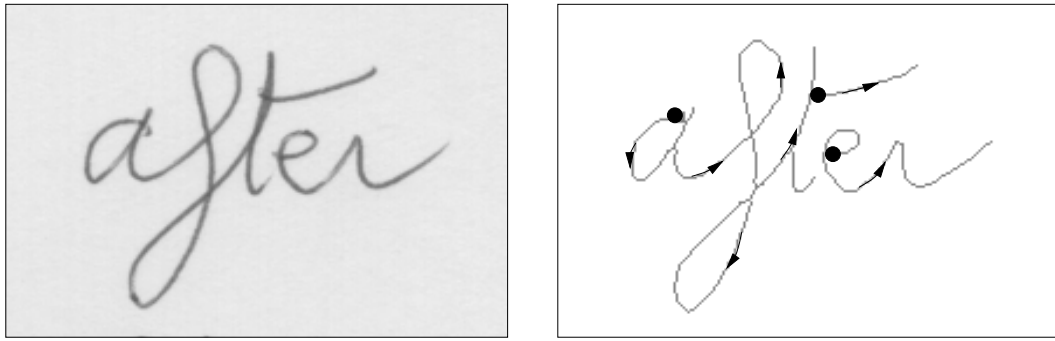


FIG. 5 – *Ordre du tracé des points reconstruit à partir de l'image numérisée. Les points noirs repèrent les débuts de tracé.*

### 6.3 Reconnaissance des symboles musicaux

*Participants* : Éric Anquetil, Bertrand Coïasnon

*Mots-clés* : classification neuronale, *Radial Basis Function* (RBF), rejet, interprétabilité.

**Résumé** : *Dans le cadre du développement d'une chaîne complète de reconnaissance de partitions musicales, un classifieur neuronal de symboles musicaux a été développé. Grâce à une complète interprétabilité, un rejet fiable est possible tout en conservant un bon pouvoir de généralisation. Ceci permet de valider des hypothèses de segmentation contextuelle.*

La reconnaissance de partitions musicales est un domaine d'étude particulièrement intéressant dans lequel la connaissance *a priori* est très structurée et où de nombreux problèmes de segmentation n'ont pas encore été résolus par les méthodes usuelles. En outre, ces problèmes de segmentation sont très représentatifs de ceux que l'on rencontre en reconnaissance de documents.

Pour pouvoir appliquer la méthode DMOS (section 3.3) à ce type de document structuré, nous avons défini une grammaire complète de la notation musicale avec l'aide de Bernard Rétif, professeur, conseiller aux études au Conservatoire de musique de Région, à Caen.

Dans le cadre du développement d'une chaîne complète de reconnaissance de partitions musicales, il est nécessaire de définir un classifieur permettant de reconnaître les symboles musicaux pour valider les essais de segmentation produits par la méthode DMOS grâce au contexte et à l'introduction de la connaissance *a priori*.

Comme les symboles musicaux correspondent à tous les objets pouvant être assimilés à des caractères (clés, altérations...), leur reconnaissance est relativement proche de la reconnaissance optique de caractères. Cependant, la principale différence, et c'est ce qui en fait la principale difficulté, est due aux nombreux problèmes de segmentation rencontrés sur les partitions. Ceux-ci impliquent que le classifieur développé doit absolument être capable de rejeter une forme inconnue correspondant par exemple à un symbole mal segmenté. Le rejet du classifieur permet alors de tenter une nouvelle segmentation contextuelle.

FIG. 6 – *Partition reconstruite (image du bas), à partir d'une image binaire (image du haut).*

Le classifieur développé est constitué d'un réseau de neurones à fonctions à bases radiales (*Radial Basis Functions, RBF*), prenant en entrée des caractéristiques géométriques classiques en reconnaissance de caractères. Son originalité provient de la méthode d'apprentissage, locale à chaque classe et en deux étapes, qui permet de réaliser un rejet bien défini et efficace, tout en conservant de bonnes propriétés en généralisation.

Les premiers tests effectués sur le classifieur ont porté sur une base de 2 100 symboles musicaux, répartis sur 13 classes. Sur cette base, par exemple avec un rejet de 2,7%, on obtient un faible taux d'erreur de 0,3% et donc un taux de reconnaissance de 97%. Ces résultats sont intéressants car les faibles taux d'erreurs sont généralement difficiles à obtenir sans un fort rejet. Ce classifieur est en cours d'intégration dans la chaîne complète de reconnaissance.

Cette chaîne s'appuie sur une détection des segments à l'aide d'un filtre de Kalman (section 3.1), et est déjà capable de reconnaître des partitions d'orchestre à portées polyphoniques, en corrigeant certaines erreurs de segmentation (symboles touchant des notes) et en indiquant les zones comportant des erreurs (la figure 6 donne un exemple de reconnaissance dans lequel la règle pour les demi-soupirs n'est pas implémentée et qui ne comporte pas l'affichage de la détection des erreurs, mais dans laquelle la séparation des voix est réalisée).

La prise en compte de rejets décidés par le classifieur de symboles musicaux devrait permettre de résoudre d'autres problèmes de segmentation, les symboles qui se touchent, par exemple.

## 6.4 Intégration de connaissance lexicale

*Participants* : Guy Lorette, Laurent Pasquer

Les chaînes de reconnaissance en-ligne ou hors ligne ne peuvent arriver à lever certaines ambiguïtés lorsque les lettres sont un peu déformées par rapport à leur modèle théorique. Ainsi on peut passer continuellement d'un *a* à un *u*, selon le caractère plus ou moins fermé de la partie haute de la boucle (cf. *figure 7*).

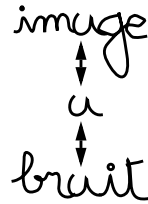


FIG. 7 – *Différence d'interprétation d'un même graphisme en fonction du contexte lexical.*

Pour améliorer les performances de reconnaissance au niveau des mots, il faut s'appuyer sur d'autres connaissances que celles issues du signal d'entrée.

Nous travaillons actuellement sur l'incorporation de connaissances lexicales. Classiquement celle-ci se fait sous la forme d'un post-traitement éliminant de la liste des mots fournie par le reconnaiseur, les mots qui n'appartiennent pas au lexique. Au contraire, nous introduisons directement les connaissances lexicales dans le processus de fabrication de cette liste de mots tel qu'il est décrit dans la section 6.1. Pour cela on fusionne les contraintes lexicales avec les informations graphiques extraites par le reconnaiseur pour arriver à un consensus sur les mots les plus probables en tenant compte, d'un côté, des classes de lettres pouvant être confondues et de l'autre, des lettres dont la reconnaissance est très sûre (lettres clés, points d'ancrage). La fusion intègre aussi l'adéquation entre les différentes successions de lettres proposées par le reconnaiseur.

Cette intégration procède par génération et validation d'hypothèses, ce qui nous permet de ne plus avoir besoin de générer toutes les hypothèses de lettres, puis de rechercher les mots les plus proches.

## 6.5 Organisation de dictionnaire

*Participants* : Guy Lorette, Gildas Ménier

*Mots-clés* : dictionnaire, organisation, accès.

### Résumé :

*Une organisation particulière d'un dictionnaire, qui soit mieux adaptée à la reconnaissance de l'écrit manuscrit, est proposée. Elle permet une recherche beaucoup plus rapide que dans un dictionnaire à structure arborescente.*

Pour obtenir des systèmes de lecture automatique de *l'écrit manuscrit*, dont les performances soient proches de celles de l'homme, il est nécessaire de disposer de

dictionnaires dont l'organisation, le mode d'accès et la rapidité de recherche soient optimisés. L'organisation classique d'un dictionnaire sous forme arborescente ne permet pas d'effectuer une recherche par le contenu lexical à partir d'un sous-ensemble réduit de lettres clés qui constituent autant d'indices visuels. Ces lettres clés sont celles qui, à l'intérieur d'un mot, ont été préalablement reconnues avec l'indice de confiance le plus élevé. Ce type d'organisation arborescente de dictionnaire ne permet pas non plus une recherche du sous-ensemble des mots dont l'orthographe est voisine de celle d'un mot donné.

Nos recherches ont porté sur l'organisation d'un dictionnaire sous forme d'une pyramide multirésolution de cartes de Kohonen [15], en s'inspirant de celle que l'on trouve au niveau du cortex cérébral pour représenter les entrées sensorielles, (cf. *figure 8*). Afin de respecter la notion de voisinage, les cartes 2D de Kohonen sont plaquées sur une structure torique. Cette approche effectue une classification *a priori* de type non supervisé de l'ensemble des mots du dictionnaire. Cette classification est faite à partir des vecteurs qui représentent le contenu lexical de chacun des mots et ce, en utilisant un critère de distance d'édition de Levenshtein [Ste94].

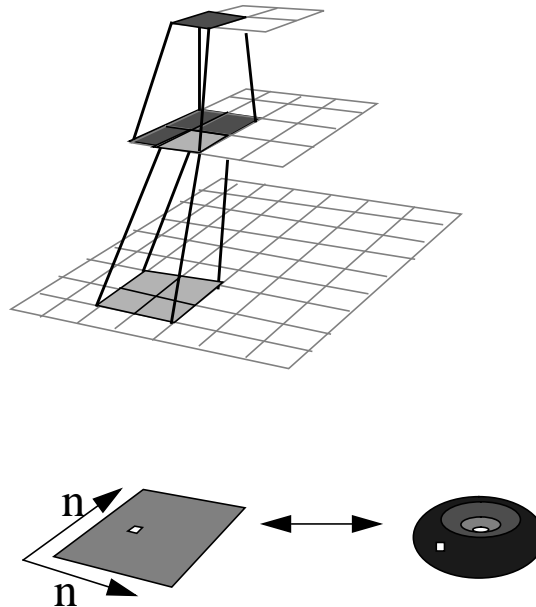


FIG. 8 – *Pyramide multirésolution de cartes lexicotopiques et plaquage d'une carte sur une topologie toroïdale.*

Lors de la phase de post-traitement lexical, l'organisation choisie permet au système de reconnaissance de se focaliser, de manière très rapide, sur un sous-dictionnaire de très petite taille. Ceci se fait par calcul d'un nombre très restreint de distances entre la représentation du mot à reconnaître et celles de mots situés aux différents niveaux de la pyramide. Un premier test a été effectué sur un lexique de 1125 mots. Pour 10 000 accès au lexique, les mots étant pris dans un ordre aléatoire, le nombre moyen de distances calculées passe de 1125 à 56 seulement. Cette organisation permet également de se focaliser, à chaque fois, sur un sous-lexique comportant, *in fine*, moins de 5% de l'ensemble des mots de départ.

---

[Ste94] Graham A. Stephen. *String Searching Algorithms*. World Scientific Publishing Co., 1994.

## 6.6 Interface homme machine orientée stylo

*Participants* : Eric Anquetil, Gildas Ménier

*Mots-clés* : interface orientée stylo, communication homme-machine, éditeur orienté stylo.

**Résumé** : *La conception d'interface Homme-Machine orientée stylo vise à établir de nouvelles modalités d'interaction entre l'homme et la machine qui sont plus conviviales et plus naturelles. Dans ce cadre, notre objectif est de développer des interfaces permettant d'intégrer de manière homogène la reconnaissance aussi bien du geste que de l'écriture manuscrite ou encore d'autres symboles graphiques afin de pouvoir utiliser une modalité de communication directe avec l'ordinateur par l'intermédiaire d'un papier électronique et d'un stylo.*

Nos travaux ont porté plus particulièrement sur deux applications conçues dans un cadre générique de programmation. La première de ces applications concerne le développement d'un éditeur de partitions musicales permettant la saisie directe sur une tablette graphique de notes et de symboles musicaux. Cet éditeur intègre différentes commandes d'édition basées sur la notion de gestes graphiques (suppression, ajout d'une note, etc.). La seconde application concerne le développement d'une interface de saisie d'écriture manuscrite qui a été couplée au système de reconnaissance d'écriture que nous avons développé. La saisie s'effectue à l'aide d'un stylo et d'une tablette électronique en intégrant là aussi des commandes gestuelles simples d'édition (sélection, déplacement, suppression, sauvegarde, lecture, etc.). Ces applications ont été développées en JAVA ce qui a permis une portabilité des interfaces sur différentes plate-formes (station sous Solaris, PC sous Windows95 et Windows NT).

## 7 Actions industrielles

### 7.1 Convention SRTP : étude et traitement d'images de documents en niveaux de gris, CDP 510227-3

*Participants* : Jean Camillerapp, Guy Lorette

Le projet entretient des relations suivies avec le SRTP (Service de recherche technique de la Poste), organisme qui a, en France, un rôle fédérateur important pour les équipes travaillant sur la reconnaissance de l'écriture manuscrite.

Dans ce cadre, une convention a été signée pour étudier l'influence d'une nouvelle méthode de binarisation des images d'adresses postales.

Une très bonne binarisation devient indispensable pour permettre d'accroître encore les performances des systèmes de reconnaissance. En effet, les boucles très serrées ou les traits proches peuvent conduire à des confusions, si la binarisation, telle que peut la réaliser un scanner, est trop brutale. Nous avons fourni un logiciel de binarisation qui utilise des détecteurs de contour classiques pour réaliser une bonne séparation

entre le tracé et le fond dans les portions simples du tracé, et qui les relaye par des méthodes de grossissement de région dans les zones complexes. Celles-ci s'appuient sur une interprétation locale des variations de luminosité au voisinage des extrémités des contours.

Le SRTTP est en train d'intégrer cette méthode de binarisation dans la nouvelle chaîne qu'il développe, et nous attendons ses résultats pour reprendre cette approche.

## 8 Actions régionales, nationales et internationales

### 8.1 Actions régionales

Le projet IMADOC a mis à disposition de l'équipe Traitement du signal et réseaux neuronaux de Supélec, à Rennes, une base de données de caractères manuscrits acquis en-ligne.

Jean Camillerapp encadre Damien Legeard, boursier Cifre au Cemagref (Centre National du Machinisme Agricole, du Génie Rural, des Eaux et des Forêts) sur l'introduction de méthodes de vision numérique dans la mesure automatique de critères de qualité d'un produit agro-alimentaire.

## 9 Diffusion de résultats

### 9.1 Animation de la communauté scientifique

G. Lorette est l'un des deux représentants français au « Governing Board » de l'IAPR (International Association of Pattern Recognition) ainsi que le président du « Technical Committee » TC-11 (Applications in text processing).

G. Lorette est membre du comité de rédaction des revues : « Pattern Recognition » et « International Journal of Document Analysis and Recognition » (IJDAR).

B. Coüasnon, I. Leplumey et G. Lorette participent au groupe de travail GT7, « Écrits et documents » du GDR-PRC Communication Homme-Machine.

J. Camillerapp et G. Lorette participent aux activités de l'association GRCE : « Groupe de Recherche en Communication Écrite ». G. Lorette est vice-président du GRCE et responsable des relations internationales.

G. Lorette est membre du comité de programme du 11ème congrès Reconnaissance des Formes et Intelligence Artificielle (RFIA'98).

J. Camillerapp et G. Lorette sont membres du comité de programme du Colloque International Francophone sur l'Écrit et le Document (CIFED'98).

G. Lorette est membre des comités de programme des : 8th biennial conference of the International Graphonomics Society (IGS'97), 4th International Conference on Document Analysis and Recognition (ICDAR'97), International Conference on Computational Linguistics, Speech and Document Processing (ICCLSDP'98).

## 9.2 Enseignement universitaire

DEA informatique, Ifsic (G. Lorette : analyse d'image et vision).

E. Anquetil, "Introduction to cluster analysis, fuzzy clustering", *Third Erasmus Seminar on Artificial Intelligence in Engineering*, Mai 1997, Rennes.

Stages effectués au sein de l'équipe: E. Pelletier, O. Laurelut (DEA informatique), M. Ravart (DIIC IFSIC), X. Gatellier, K. Montheillet, L. Pucel, D. Chazalviel, C. Dumont-Girard, L. Vincent(INSA)

## 9.3 Participation à des colloques, séminaires, invitations

E. Anquetil, « Génération de règles floues par apprentissage : application à la reconnaissance en-ligne de l'écriture manuscrite », *Séminaire intelligence artificielle et processus de décision*, groupe de travail « Information et Systèmes » de l'AF CET, responsable : Bernadette Bouchon-Meunier, Université Paris VI, juin 97.

E. Anquetil, « Extraction et modélisation automatique de connaissances par génération de règles floues : application à la reconnaissance en-ligne de l'écriture manuscrite », *Les ensembles flous en traitement du signal et des images*, Journée commune GT2-GT5 du GDR-PRC ISIS, responsable : Isabelle Bloch, ENST Paris, avril 97.

# 10 Bibliographie

## Ouvrages et articles de référence de l'équipe

- [1] 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.
- [2] E. ANQUETIL, G. LORETTE, « Reconnaissance en-ligne de lettres manuscrites cursives par chaînes de Markov cachées », *Traitement du Signal 12(6)*, 1996, p. 575–583.
- [3] B. COUASON, 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, Montreal, 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, 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.
- [7] G. MÉNIER, G. LORETTE, P. GENTRIC, « A genetic algorithm for on-line cursive handwriting recognition », *in: Proceedings of 12th International Conference on Pattern Recognition*, p. 522–525, Jérusalem, Israel, 1994.
- [8] G. MÉNIER, G. LORETTE, P. GENTRIC, « A New Modeling Method for Handwriting Recognition », *in: Proceedings of the 3rd International Conference on Document Analysis and Recognition, (ICDAR'95)*, p. 499–503, Montreal, Canada, 1995.
- [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)*, Tsukuba, Japon, 1993.

### Thèses et habilitations à diriger des recherches

- [10] E. ANQUETIL, *Modélisation et reconnaissance par la logique floue : application à la lecture automatique en-ligne de l'écriture manuscrite omni-scripteur*, thèse de doctorat, université de Rennes 1, janvier 1997.

### Chapitres de livre

- [11] E. ANQUETIL, G. LORETTE, « On-line Handwriting Character Recognition System Based on Hierarchical Qualitative Fuzzy Modeling" », *in: Progress in Handwriting Recognition*, World Scientific, 1997, p. 109–116.
- [12] G. LORETTE, J.-P. CRETTEZ, « Reconnaissance de l'écriture manuscrite », *in: Traité informatique*, Techniques de l'Ingénieur, ch. H 1358, à paraître.

### Communications à des manifestations scientifiques

- [13] E. ANQUETIL, G. LORETTE, « Génération de SIF sémantiquement interprétables pour la reconnaissance en ligne de caractères manuscrits », *in: Rencontres Franco-phones sur la Logique Floue et ses Applications (LFA'96)*, p. 203–210, Nancy, France, 1996.
- [14] E. ANQUETIL, G. LORETTE, « Perceptual Model of Handwriting Drawing Application to the Handwriting Segmentation Problem », *in: Proceedings of the Fourth International Conference on Document Analysis and Recognition (ICDAR'97)*, p. 112–117, Ulm, Allemagne, 1997.
- [15] G. MÉNIER, G. LORETTE, « Lexical Analyser based on a Self-Organizing Feature Map », *in: Proceedings of the Fourth International Conference on Document Analysis and Recognition (ICDAR'97)*, p. 1067–1071, Ulm, Allemagne, 1997.

### Divers

- [16] G. LORETTE, « Le traitement automatique de l'écrit et du document, État de la recherche », *Documentaliste, Sciences de l'information*, juillet/octobre 1996, Vol. 33, N°4/5, p. 214-217.