Project-Team IMADOC

Interprétation et Reconnaissance d'Images et de Documents

Rennes

THEME COG

Contents

1 Team

Team Leader

Jean Camillerapp [Professor Insa]

Administrative Assistants

Stéphanie Lemaile [TR Inria to 31 Aug 2005]

Diana Gaudin [TR Inria from 01 Sept 2005]

Staff members (Inria)

Jacques André [Research Director to 31 Jan 2005]

Bertrand Coüasnon [Research scientist (secondment) to 31 Aug 2005]

Staff members (Insa)

Éric Anquetil [Associate Professor]

Bertrand Coüasnon [Associate Professor from 01 Sept 2005]

Ivan Leplumey [Associate Professor]

Yann Ricquebourg [Associate Professor]

Staff members (Université of Rennes 1)

Guy Lorette [Professor]

Charles Quéguiner [Associate Professor]

Hélène Richy [Associate Professor]

Nicolas Ragot [Lecturer to 31 Aug 2005]

Sabine Carbonnel [Lecturer from 01 Oct 2005]

Ph.D. Students

Sabine Carbonnel [Menrt research grant to 30 Sep 2005]

Laetitia Rousseau [Insa research grant]

François Bouteruche [Insa research grant]

Isaac Martinat [Insa research grant]

Harold Mouchère [CNRS/Région research grant]

Solen Quiniou [Menrt research grant]

Aurélie Lemaitre [Menrt research grant from 01 Oct 2005]

Sébastien Macé [Région/Insa research grant from 01 Oct 2005]

Technical staff

Grégory Maitrallain [Contract with the *Conseil Général des Yvelines*]

2 Overall Objectives

The research topics of the project concern *the written communication* and *the engineering of documents* under various aspects: *analyzis, recognition, interpretation of digitized documents, man-document interaction.* This research relates to the writing and the documents under all their forms (manuscript, printed paper form, graphs, images, composite documents, etc.) as well as the linked activities.

Facing the multiplicity and the diversity of the applications, the usual solution consists in bringing answers dedicated to each type of application. The project privileges, on the contrary, the development of generic approaches. With this intention, our research is carried out according to five interdependent topics, two of more methodological nature and three of applicative nature.

- **Image processing, early vision**
  Low level processing, also called early vision processing, has an unquestionable influence on the total performances of the data processing sequences.The digitalization of document images must carry out a compromise between the total size of the images and the apparent size of the characters. In this compromise, the width of the feature can be reduced to two or three pixels. The document image processing thus requires a very good precision: in the assignment of a pixel either to the background or to the layout and in the extraction of the visual indices characteristics.

- **Modeling and management of knowledge**
  In order to develop automatic systems of *interpretation* of signals and of document images, the project studies methods of dynamic and multi-contextual recognition based on visual indices. Three complementary approaches for the modeling of contextual knowledge are considered for this purpose: one based on rules for symbolic knowledge systems, another one based on fuzzy inference systems for vague knowledge, and the third one based on neural networks for numerical data.

  The modeling of knowledge is carried out, using either models of *a priori* knowledge or machine learning from experimental data.

  In order to lead to the final interpretation, contextual knowledge is dynamically managed, according to the respective types of modeling, either by modified syntactic analysis, or by a soft computing method (fusion of fuzzy data or perceptive cycles).

  These two approaches have the advantage of offering several possible levels of reading, to locate where the errors remain and to lead to transparent systems which way of working can be analyzed and which performances can be optimized.

- **Handwriting**
  We study recognition systems for *handwritten* documents, either during their phase of creation (recognition of *on-line* writing), or afterwards (recognition of *off-line* writing).

  The operation of these systems is based on *a priori* general knowledge about handwriting, the styles of writing and the visual indices resulting from the layout.

  This enables us to overcome the complex problem of segmentation of the words in separate letters and facilitates recognition of the words of a large size vocabulary (*analytical approach*). The use of such systems can be of interest to a very large audience (*multilingual and omni-script writers recognition systems*). Nevertheless, it is always possible to specialize a given system to build from it a dedicated application.

- **Numerical document and semantic Web**
  The recent technological developments (development of the Internet, high flow communications, large storage capacities, effective image compression algorithms, policy of digitalization on a large scale of the patrimonial or industrial files) make it possible to place enormous corpora of digitized documents at the public disposal.

  However, an intelligent access to the images (*semantic Web*) supposes the existence of annotations describing with precise enough details the contents of these images.

  The studies which we currently undertake relate to the development of models and techniques aiming at facilitating the automatic annotation of digitized images.

  Our research relates to the transformation of images of existing paper documents (reverse conversion of old documents) into structured numerical documents. The generic approach relates not only to the description of the forms (geometry), but also to the definition of models of annotation and ontologies on the contents. This approach makes it possible to control the various stages of the transformation of the documents:
  - localization: location of zones in images,

- reverse conversion: extraction of information and structuring,
- annotation: analysis and interpretation of the contents,
- indexing: search for information (tools to recompose images, filtering, etc).

- **Pen-based interaction**
  As most electronic devices are now provided with a pen interface, a number of applications are expanded where the pen can be used as a convenient and natural form of input. Moreover, handwriting is a very familiar modality of input. Thus, our work focuses on the adaptation of recognition algorithms for an interactive environment.

3 Scientific Foundations

3.1 Early vision

Keywords: early vision, binarization, skeletonization, Kalman filtering.

Adaptative binarization

Documents are generally digitized at a minimal sampling rate, which implies that elements important for recognition can take only a few pixels. For example, the non-detection of a very small white area can mask the presence, however very informative, of the loop in the letter *e*.

(a) Threshold of 210 : allowing the detection of thin lines

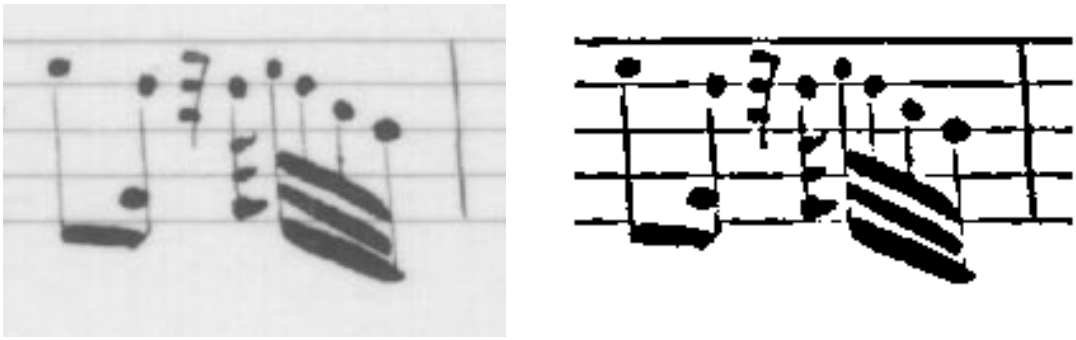
(b) Threshold of 195 : the vertical elements remain

(c) Threshold of 110 : the bars of the notes are separated

Figure 1: Binarization of the gray levels image of the figure 2 with a global threshold.

A digitalization in gray levels allows to implement adaptative binarization algorithms, often based on cooperation between edge detection and region growing. They give a local interpretation of gray values which separate well the various objects in the document (cf figures 1 et 2).

Measurements of objects characteristics are sensitive to space sampling noise. Hence increasing space resolution improves the relevance of parameters related to visual indices. With a local interpretation of gray level, it is possible to locate the edge between the objects and the background with a *subpixel* precision and thus increase the relevance of these parameters.



(a) Initial image in gray level (b) Binarized image

Figure 2: Binarization with an adaptative threshold

Gray level image skeletonization

Skeletonization algorithms have been developed to thin images of lines. But in documents images, it is necessary to distinguish in the layout (figure 3) the area really skeletizable (also called regular area) from the singular areas, principally localised in crossings in which the concept of median axis does not have any practical interest.

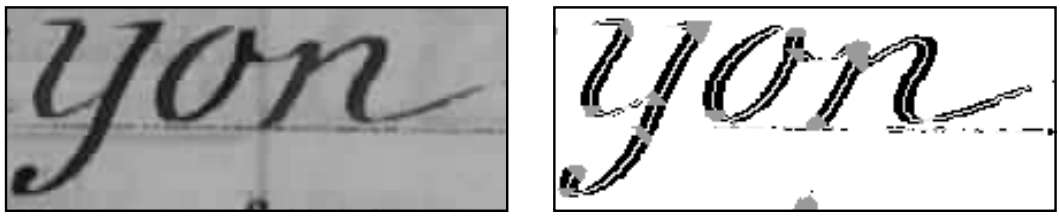


Figure 3: Processing of an image from an old document. The skeleton appears in white in the medium of the line and singular areas are in dark gray.

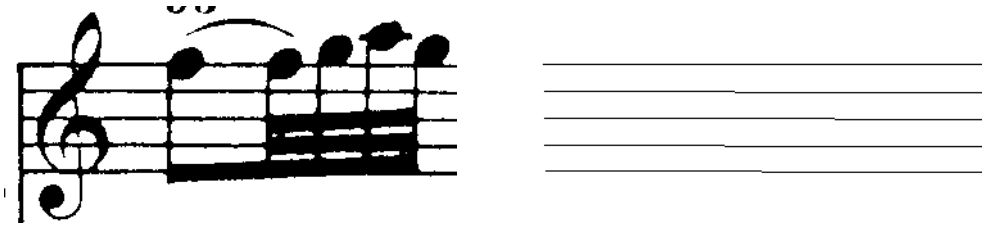
In order to limit the binarization noise, one can obtain a sub-pixel localization of points of the median axis while working on gray level images.

Segmentation with Kalman filters

Linear structures have a great importance in documents, but they are likely to interfere between them or with the drawing. However, the simplicity of their structure makes it possible to plan segmentation and recognition simultaneously.

Kalman filtering is a technique of parameters identification from an ordered series of measures. In the case of lines, the model is reduced to the thickness of the line, its slope and the

equivalent of the ordinate at zero point. Measure results from the position and the thickness of the black run-length[1] orthogonal to the drawing.



(a) Original image (b) Extracted stave lines

Figure 4: Detection of stave lines with Kalman filters

The Kalman filtering also calculates the covariance matrix of its estimate of the parameters of the model. This matrix makes it possible to evaluate on one hand the probability to assign a measure to a segment and on the other hand the probability to continue an assumption disturbed by the presence of another object [5].

The real difficulty of this approach does not lie in the equations of Kalman, but in the association of the extracted run-length from the image with the predicted run-length when linear structures cross or overlap (figure 4).

Thus a run-length in the image can result from overlap of several drawings. At this level the information brought by the covariance matrix is of great importance.

In this approach we can also use a measure of a run-length with a sub-pixel precision.

3.2 Dmos, a Generic Method for Structured Document Recognition

Keywords: syntactical recognition, grammar, definite clause grammar, segmentation, structured document recognition, musical scores, mathematical formulae, table structure, form.

Literature shows that structured document recognition systems are long and difficult to develop. Moreover, we usually find systems specifically developed for one kind of document - for example mathematical formulae - which cannot be easily adapted to another kind of document, like table forms. A new kind of document often requires the complete development of a new recognition system, which is a real loss of energy. Defining a generator of recognition systems for structured documents is a real challenge to avoid this costly rewriting work.

Moreover, improving recognition quality is vital for an industrial processing of documents. This can be done, for example, by solving segmentation problems usually found in documents.

In fact, we can consider that those problems are linked to the management of *a priori* knowledge. Indeed:

[1] Set of consecutive black pixels according to one of the four directions: horizontal, vertical or diagonals

- a generic system must allow to change only what changes from a document to another: the *a priori* knowledge on each document;
- improving segmentation on documents needs to introduce as much as possible *a priori* knowledge in the system.

Therefore, we proposed the Dmos method (Description and MOdification of Segmentation), a generic recognition method for structured documents made of:

- the new grammatical formalism EPF (Enhanced Position Formalism), which can be seen as a description language for structured documents. EPF makes possible at the same time a graphical, a syntactical or even a semantical description of a document;
- the associated parser which is able to change the parsed structure during the parsing. This allows the system to try other segmentations with the help of context to improve recognition.

EPF Language

We can find in the literature various bi-dimensional extensions of mono-dimensional grammars defined for object and document recognition. However, they offer a too weak expressiveness (trees and *web grammars*), or the syntax is too complex for dealing with large and difficult *a priori* knowledge (plex grammars, graph grammars).

We developed the EPF formalism to be able to describe a bi-dimensional structure, in a graphical as well as in a syntactical way. Instead of having characters as terminals in mono-dimensional grammars, EPF uses two kinds of terminals: line-segments and pixel arrays (components, connected or not, which represent a symbol). In EPF several operators are added. Here are some examples:

Position Operator (encapsulated by `AT`):

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

means `A`, and at the position `pos` in relation to `A`, we find `B`.

Where `&&` is the concatenation in the grammar, `A` and `B` represent a terminal or a non-terminal.

Factorization Operator (`##`, in association with the position operators):

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

means `A && AT(pos1) && B` and `A && AT(pos2) && C`

With this syntax it is possible, for example, to describe a simplified beamed note (a beamed eighth note, with only one beam - a segment - that links the notes):

```
beamedNote ::= beam &&
         (AT(leftTip) && noteGr ##
          notesInBetween ##
          AT(rightTip) && noteGr)
```

Where `::=` is the constructor of a grammar rule. The writer of the grammar can define, as much as necessary, position operators as well as he can for non-terminals.

Save Operators (`--->` and `<---`): To be able to represent the same instance of a terminal or a non-terminal `A`, we propose to save an instance of it by using the operator `--->`. This backup of `A` gives then the possibility to refer (using the operator `<---`) to `A` as much as necessary. It is then possible to describe a rectangle:

```
rectangle ::=
   (segV ---> segLeftSide) &&
   AT(touchUp) && segH &&
   AT(touchRight) && segV &&
   AT(touchDown) && segH &&
   AT(touchLeft) &&
   (segV <--- segLeftSide).
```

Associated Parser

The EPF language allows to define a description of a document. From this description, we produce by compilation, a parser with the specific properties needed for parsing bidimensional documents. Compared to classical monodimensional parser, the main properties of the bidimensional parser we develop are:

- changing the parsed structure during parsing for contextual segmentation. The parsed structure can then be changed to become like it should be if there was no segmentation error;
- detecting the next element to parse. Indeed, in classical parsers, the next element is simply the head of the parsed string. In two dimensions, the next element can be everywhere in the image, hence everywhere in the parsed structure. The parser uses the position operators to select the next element to parse;
- dealing with noise. We can consider that dealing with noise corresponds to finding the next element to parse, even when there is a lot of noise in the parsed structure.

In conclusion, with this new EPF language we have developed a generic recognition method for structured documents. We implemented this method to obtain a generator of structured document recognition systems. This generator can automatically produce new recognition systems. It is just necessary to describe the structure of the document with an EPF grammar which is then simply compiled. In this way we have developed various recognition systems: on musical scores, on mathematical formulae, on recursive table structures...

3.3 Pattern recognition by fuzzy inference systems

Keywords: Explicit modeling of the knowledge, fuzzy inference system, learning, classification.

Faced with the variability and the complexity of patterns to be modeled, many recognition systems are often based on "opaque" algorithms which cannot be easily interpreted after the learning process. This "opaque" aspect often hinders an explicit description of the process. Therefore, the modification and the interpretation of error causes of complex system is very difficult. Handwriting recognition is a typical example of this kind of problem. Different methodologies are used to try to solve this difficult problem like approaches based on neural networks or methodologies based on hidden Markov models that lead typically to "opaque" systems. Up to now, it is commonly believed that fuzzy set theory is an interesting approach to design "transparent" modeling (i.e. readily understandable) capable of dealing with the input imprecision. More precisely, we use fuzzy rules to design robust, compact and transparent classifiers. A fuzzy inference system is composed of N_r fuzzy rules to model the C classes. These rules can be expressed explicitly by:

$\boldsymbol{R_i}$: **IF** x_1 is M_{i1} **and** x_2 is M_{i2} **and** ... **and** x_n is M_{in}
THEN the input shape belongs
to the class 1 with the degree $\{b_{i1}\}$ **and** ... **and**
to the class k with the degree $\{b_{ik}\}$ **and** ... **and**
to the class C with the degree $\{b_{iC}\}$.

- the inputs $x = (x_1, ..., x_n)$ are defined in a n dimensional space;
- the premise part of the rule is a conjunction of (x_j is M_{ij}) that formalizes the membership degree $\mu_{M_{ij}}(x_j)$ of the input x_j, to the fuzzy set M_{ij} modeling a property of a class;
- the outputs Y_k are fuzzy sets B_{ik} reduced here to singletons $\{b_{ik}\}$;
- the consequent part of the rule (Y_k *is* b_{ik}) formalizes the membership degree of the input shape, according to the rule R_i and the class k.

Fuzzy inference is then defined as below :

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

where

- $\mathbf{T}$ is a T-norm corresponding to the conjunction in the premise of the rule;
- β_i is the activation degree of the rule R_i;
- I is the fuzzy implication[BM95];

[BM95] B. Bouchon-Meunier, *La logique floue et ses applications*, Addison-Wesley, 1995.

- $\perp$ represents the aggregation operator.

The output of the fuzzy inference system provides the adequacy of the input shape to each class.

The automatic generation of membership functions is still an open problem. We use here the possibilistic clustering developed by Krishnapuram [Kri94] to automatically generate fuzzy rules. Krishnapuram has shown that possibilistic clustering can be successfully used to solve the determination of membership functions which can be interpreted as degrees of "typicality", contrarily to classical clustering where membership functions are relative to each other and interpreted rather as degrees of "sharing". Faced with the problem of complex modeling and to keep an explicit description of each model, we structure hierarchically fuzzy rules relatively to the robustness and to the pertinence of the modeled primitive. The internal structure of the models is automatically generated during the learning phase and is consequently adapted to current class which is going to be modeled. Based on the previous observations, we set out a methodology to generate "transparent" pattern recognition systems based on qualitative modeling by hierarchical fuzzy rules automatically deduced from possibilistic clustering.

This methodology has been applied to the difficult problem of on-line handwritten character recognition. For each character class, robust and stable properties have been automatically extracted and modeled in a qualitative manner. They are organized in a hierarchical structure according to their robustness and guided by *a priori* knowledge on cursive handwriting. This led to the RESIF software specifically elaborated for on-line cursive recognition (cf. section 5.1). Thanks to its compact fuzzy modeling, RESIF is able to absorb the variability of handwriting shapes and can cope with limited computing and memory resources. In 2001, through an industrial collaboration with PurpleLabs company[2], RESIFCar has been successfully integrated into mobile devices (smartphones) nowadays commercialized in Europe. Recently, a new classification approach, named Mélidis, has been elaborated. This new system can be considered as an extension of RESIF. The main goal is to obtain a more generic recognition system able to deal with different kinds of classification problems without needed *a priori* knowledge and without loss of transparency so that optimization can be done afterwards by experts (cf. section 6.7).

4 Application Domains

4.1 Paper document analysis

Applications of our research in the context of paper document analysis are very wide. Indeed, the generic approach we have developed (the Dmos method) allows a quite fast adaptation to new kinds of documents. Documents we have already worked on are:

- Orchestra scores with polyphonic staves;

[2]PurpleLabs : cf. http://www.purplelabs.com

[Kri94] R. Krishnapuram, "Generation of membership functions via possibilistic clustering", *in : IEEE World congress on computational intelligence*, p. 902–908, 1994.

- Mathematical formulae;
- Table structures, forms with recognition of the hierarchical organization;
- Archives documents: old forms more or less structured.

The Dmos method could also be applied to various other kinds of documents: chemical formulae, title blocks, newspaper pages, invoices, delivery notes...

Moreover, a grammatical description of a kind of document (made with EPF) can be either general, in order to recognize a class of document (title blocks for example), or specific, to compensate a lack of information (damaged title blocks).

Of course, it is possible to merge grammatical descriptions of documents. For example, we can simply build a system able to recognize table structure containing mathematical formulae.

With the help of handwriting recognition those recognition systems can be used in many ways:

- Retroconversion of paper documents to avoid a manual input to get a usable electronic version. It can for example be used for a new edition, or to produce a Braille document;
- Kind of document identification, to make for example an automatic management of faxes;
- Automatic production of indexes and annotations for an automatic access by content to documents (section 6.1);
- Detection of specific areas in a document to improve its access.

4.2 Handwriting recognition systems

Recently, there was a new increasing of the applications of handwriting recognition in both domains of automatic processing of paper documents (*off-line recognition*) and in the new modality of man-machine interaction based on the use of a pen and a touch-screen (*on-line recognition*).

To date, in *off-line recognition*, industrial needs are huge and for this reason the design of robust and performant recognition systems is highly needed. The application domain range is very large. It concerns the problem of the automatic processing of every kind of paper documents, e.g. order lists, social security forms or fax. In this domain, we mainly brought our efforts to guarantee a high degree of robustness and confidence in the results in the case of automatic processing of handwriting. This was done to avoid any risk of error.

In *on-line recognition*, a huge market has arised due to the recent economic development of the following devices:

- pocket computers (Palm, PocketPC),
- tabletPCs (computers without keyboard and with a sensitive LCD),
- mobile phones of the new generation (¡¡ smartphones ¿¿) which integrate the interaction modalities of the pocket computers (LCD coupled with a sensitive screen) and their principal functions (agenda, text editor, Internet access, etc.).

To date, the ResifCar software was embedded in such devices (cf. subsections 5.1, 5.2 and 7.2).

4.3 Pen-based interaction

More electronic devices with pen interface are now available for entering and manipulating information. A number of applications are expanded where the pen can be used as a very convenient and natural form of input. Handwriting is an input modality that is very familiar for most users since every one learns to write at school. Pen-based interfaces capture information as the user composed it, including text, graphs or commands.

The Imadoc project focuses on the interpretation of pen input (electronic ink) as captured by the pen based interface system. The pen input may be available on any kind of device where pen interface is available: mobile-device's screens, tactile screens, electronic paper, or tablet.

Topics of interest of these experimentations are the adaptation of the recognition algorithms for an interactive environment, the validation of the generic approach, and the analysis of the benefits of multimodal interface. So, several applications are developed and experimented using gestual commands, or intuitive editing commands:

- writing a musical score,
- drawing a graph,
- annotating documents (photo, archive) for indexing (digital library),
- editing web pages.

Most applications will provide the users with a multimodal interface that enables the choice among various modes of commands.

5 Software

5.1 RESIF : Handwriting Recognition by Hierarchical Fuzzy Inference Systems

Participant: Éric Anquetil.

Keywords: Handwriting Recognition, smartphone, fuzzy logic.

RESIF technology is today composed of three main software to analyze, model and recognize handwritten characters and words:

- RESIFCar is specialized to recognize handwritten characters: latin alphabet, digits and special symbols.
- RESIFMot is the software for unconstrained cursive handwritten words recognition.

- RESIFApp is the automatic learning process that generates from a handwritten character database the hierarchical fuzzy models used by the recognition systems: RESIFCar and RESIFMot.

RESIFCar and RESIFApp are already in their fourth version. In 2001, through an industrial collaboration with PurpleLabs company[3], RESIFCar has been successfully integrated into mobile devices (smartphones) which are characterized by their limited computing and memory resources. Theses mobile phones are nowadays commercialized in Europe.

Work is in progress to extend the capability of the recognition system RESIFMot which is today in a beta version. The aim is to optimize lexical post-processing to deal with large vocabulary and to reduce time computing and memory resources.

5.2 DIGIME : a DIGital Ink Micro Editor

Participants: François Bouteruche, Éric Anquetil.

Keywords: Human-computer interface, pen-based interaction, handwriting input interface.

DIGIME is a handwriting input interface for small-size devices. Use jointly with RESIFCar (see section 5.1), DIGIME allows to enter text on this kind of devices. Its first goal was to demonstrate the RESIFCar capabilities. However, with our researches on the design of pen-based interfaces for small-size mobile devices (see section 6.8), DIGIME has become a full application and is in its second version.

The originality of DIGIME is to offer an handwritten character input in entire word context (see figure 5). It means that the previously entered characters are displayed in the input area until the user requests to clear them. Moreover, to enhance the user experience, a set of handwritten gestures allow to perform accentuation, punctuation and edition tasks without using buttons and menus. An optional feature of DIGIME is the use of a lexical post-processing which is a lightened version of the one proposed in section 6.11.

It is currently running on Windows Mobile Pocket PC 2003 and Pocket PC 2003 2nd Edition operating system. Thanks to its architecture, it can be easily adapted to other mobile operating system such as Windows Mobile Smartphone 2003 or Palm OS.

Figure 5: Screenshot of DIGIME.

[3]PurpleLabs : cf. http://www.purplelabs.com

5.3 DocRead : an Automatic Generator of Recognition Systems on Structured Documents

Participant: Bertrand Coüasnon.

Keywords: Recognition, structured document, musical scores, mathematical formulae, table structures, forms, archives.

DocRead is an automatic generator of recognition systems on structured documents. It has been developed thanks to the Dmos method. It is made of a compiler of the EPF language (with which it is possible to describe a document), a parser associated to this language, an early vision module (binarization and line segments detection) and a classifier able to reject.

This generator allows us a fast adaptation to a new type of document. Indeed, it is only necessary to define a new grammar in EPF, which describes the new type of document. Then, if necessary, a new learning of the classifier is done to make it able to recognize new symbols. The new recognition system adapted to a new structured document is produced by compilation.

With this generator, we already have been able to produce recognition systems of structured documents:

- ScoRead: a prototype for musical scores recognition;
- MathRead: a prototype for mathematical formulae recognition;
- TennisRead: a prototype for tennis court detection in videos;
- TabRead: a prototype for table structures recognition;
- DecRead: a prototype for naturalization decrees of the 19th century recognition. These are fully handwritten forms;
- FormuRead: a software for military forms of the 19th century recognition despite their deterioration. This software has been successfully tested on 88,954 pages of the *Archives de la Mayenne* and *Archives des Yvelines* (section 6.3);

5.4 Library to extract visual indices

Participant: Jean Camillerapp.

Keywords: Early vision, adaptative binarization, vectorization, Kalman filtering, squeletonization, multi-resolution.

It is the library on which the DocRead software described in the section 5.3 is based.

It is composed of various methods of adaptive binarization. Those use either local determinations of threshold or edge detection and region growing cooperation. Some of these methods can provide results with a sub-pixel resolution.

Using Kalman filters, as described in section 3.1, this library carries out detection of rectilinear segments in complex images.

It also contains a skeletonization method from gray level images.

The selected resolution for the digitalization of the documents is not always adapted at best to the recognition of all the structures. This is why the library also provides possibilities of multi-resolution processing.

5.5 FDT-L : Fuzzy Decision Trees Library

Participants: Nicolas Ragot, Éric Anquetil.

Keywords: Fuzzy decision tree, fuzzy logic, fuzzy clustering.

The FDT library is a classification software for extracting and exploiting fuzzy decision trees. Fuzzy sets are automatically extracted by fuzzy clustering algorithms in one dimension or more. Depending on the clustering algorithm used, they can either describe the classification problem based on a discrimination point of view or on an intrinsic point of view. The resulting classifiers can be more accurate than other traditional classifiers such as radial basis function network or multi-layer perceptron with a higher legibility. This first version is also able to construct a forest of decision trees to be more accurate on multi-classes problems.

FDT-L is currently used by the Mélidis system (cf. section 6.7). Work is also in progress to extend the capability of the library by taking in account more clustering algorithms and to provide a visual interface.

5.6 A platform for image document annotation

Participants: Bertrand Coüasnon, Grégory Maitrallain, Ivan Leplumey.

Keywords: Annotation, access by content, document retrieval, handwritten documents, archives.

We propose a platform to improve the access by content on archives documents with handwritten text. To make this access, it is necessary to associate annotations to the images of documents. Annotations for archives documents can be geometric or textual. With those annotations, it is then possible to make an automatic selection of images. The platform we propose to manage annotations has the interest of producing annotations in two complementary ways: automatically with document recognition and collectively with the help of the readers during their readings.

We propose to build a platform for archives document retrieval which could deal with textual and geometric annotations at the same level. Moreover this platform is able to manage relations between textual and geometric annotations to specify that a textual information is at this specific location in the image of document. In a different approach, a textual annotation can be linked to different locations in different pages of document *e.g.* in different images.

With the platform, on a web browser, a user can leaf through images of archives documents. When a page is displayed, all the associated annotations are presented on the interface: geometric annotations are drawn on the image, the textual annotations are presented in tabs for the forms (marriage certificate ...) and in field boxes for the fields in the form (name, date ...). The reader can consult annotations, add or modify annotation (if he has the right

to), but is limited by the allowed annotation structure given by the DTD configuration file, according to the kind of document. The system can also store various interpretations if readers do not agree.

Structured search or full text search is possible on all the annotations whatever the way they have been produced: automatically or manually.

Since January 2004, the platform is running in the reading room of the Archives départementales des Yvelines, and also on Internet. 430,000 military forms and 1,450,000 double pages of register of births, marriages and deaths are available for the public, through the graphical user interface of this platform for annotation.

The platform is applied presently on archives documents, but of course it could be used on every kind of images.

A pen-based interface has been defined on this platform. Using specific gestures (section 4.3) and on-line handwriting recognition (RESIFCar, section 5.1), it has been possible to design a new way to interact with digital documents and paper documents.

5.7 Handwriting Ink Grabbing Application

Participants: Grégory Maitrallain, Éric Anquetil, Sabine Carbonnel.

Keywords: Data Base Aquisition, Handwriting, InkML, Pen-based Interface.

To improve its different character and word recognition techniques, our project needs much handwriting data, from numerous writers. To gather these data, an ink managing library and handwriting grabbing application have been developped in Java.

The ink managing library permits to handle and store extended ink data in two file formats, a raw data file format (Imadoc), and a XML based file format (ImadocML). This ImadocML format includes such data as timestamped strokes, data about the writer (sex, age ...), and can deal with references to other traces, in order to handle segmentation data.

The handwriting ink grabbing application uses the ImadocML file format to design pen-input forms for differents types of inputs (letters, words, sentences, dates ...). The writer fills in the forms, and the data are then saved in ImadocML format.

A transcoding module has also been developed in order to handle raw Imadoc file format and external formats such as Unipen or InkML.

5.8 Pen-based Musical Score Editor

Participant: Sébastien Macé, Éric Anquetil.

Keywords: Handwriting recognition, musical scores, pen-based interface.

We designed a system for on-line musical score editing. Thanks to the use of a pen, the user can write musical symbols on the touch screen, the same way he does on paper. His strokes are interpreted progressively by the system and changed by their corresponding recognized symbols.

The system allows the user to edit on the staffs bass or treble clefs, whole, half and quarter notes. Thanks to additional lines, these notes can be realized above or below the staff. Notes

can have a durational dot. All the accidentals are available for the notes and for the clefs. Nuances are also present in this version of the application. Line bars can be realized on the staff. Half, eighth, sixteenth, thirty second and sixty fourth rests can already be edited. These symbols can be selected and moved from one part of a staff to another, or deleted. The system is able to deal with documents with as many pages as needed. The user can add a page anywhere in his document and move from one page to another easily and just using the pen. The system proposes some of the usual editing possibilities, like undoing the last stroke, zooming in or out. It is also possible to save the document in a file, and to load a saved document. In the current version, the saved file can only be used in the application, and it is yet impossible to export a file so as it could be used in another, more traditional, musical score editor.

Our aim is to design a system as close as possible to the user's needs. To do so, we collaborate with the MIAC (Music and Image: Analysis and Creation) laboratory of the University of Rennes 2 (see section 7.5): our functionalities and ergonomic choices are validated by professional musicians.

Figure 6 presents a screenshot of the system.

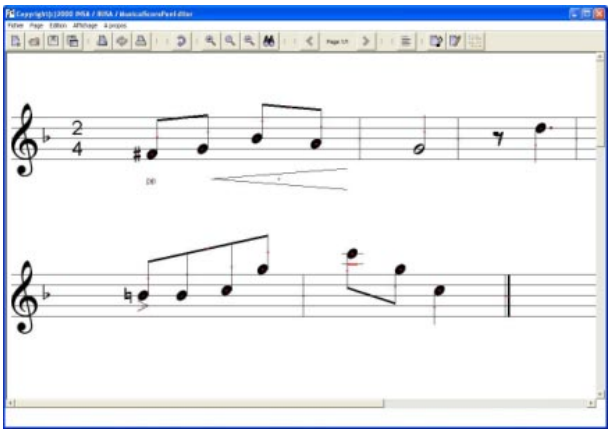


Figure 6: Screenshot of the pen-based musical score editor.

6 New Results

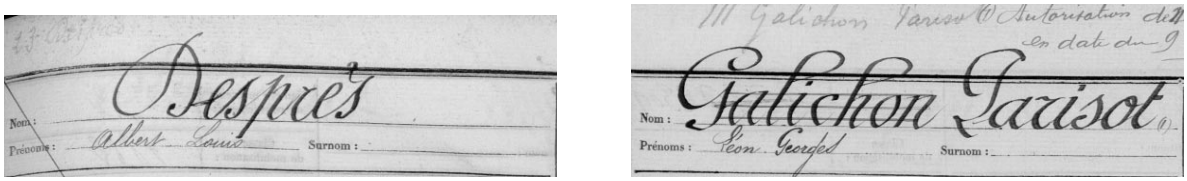
6.1 Improvements in indexing old printed forms with handwritten last name

Participants: Jean Camillerapp, Nicolas Ragot.

Digitizing old documents is important to protect originals and to display them on the Internet. However, it does not offer any improvement on document retrieval by their content. To do so, it is necessary to apply document image analysis and handwriting recognition. Handwriting recognition in old documents is, most of the time, extremely difficult or impossible

to do. Nevertheless, on some kind of old documents it is possible to think about automatic indexing of handwriting.

Military forms of incorporation of the 19th century, of which there are millions of pages, are an example of documents indexable by their handwritten contents. Indeed, these forms have a *name* cell (see figure 7) in which the last name is written by using an alphabet quite standardized: the slanting round-hand. In this handwriting style, the letters are systematically decomposed using a very limited vocabulary of shapes: the *graphemes*.



(a) Indexable image (b) Damaged image

Figure 7: Examples of name cell

To retrieve images which correspond to a given patronymic, a user makes his request with the string of characters which spells the name. Then each character of the string is systematically translated into *grapheme*. The request is then compared to indexes of all images using an edit distance. It is then possible to order the list of images according to this distance to select the images closest to the request.

We work on this subject since 2003. This year we have introduced a distinction between the a priori decomposition of letters in *grapheme* and the *forms* which the classifier can recognize. For example letter *n* breaks up into *n1, n2* and letter *m* into *m1, m2, m3*. But the visual form of *n1* and *m1* are very close. We have also modified characteristics used by the classifier.

So, a new training set of characteristics by *grapheme* and not by *form* was automatically built using a base of 60000 documents [4] whose patronym is known and the 2004 version of the application which gives some good results and thus makes it possible to align the sequence of *graphemes* of a patronym with the sequence of *forms* recognized in the image.

The weights used in the edit distance are then deduced from the confusion matrix between *graphemes* and *forms* calculated on the training set. We then could study the impact of the choice of characteristics and of the grouping of some *graphemes* in a same *form*.

Table 1 shows the improvement of the performances. One can see that those do not depend

[4] This base was manually indexed by the *Archives départementales de la Mayenne*

Version	Base	question	images	in 1^{st} position		≤ 5		≤ 10	
2003	Mayenne	363	363	120	33%	190	52%	220	61%
2004	Mayenne	363	363	201	55%	260	74%	280	77%
2005	Mayenne	363	363	226	62%	292	80%	309	85%
2004	Yvelines	496	496	271	55%	399	80%	433	87%
2005	Yvelines	496	496	296	60%	402	81%	439	89%

Table 1: Variation of the number of answers in different configurations.

on the origin of the images (Mayenne or Yvelines). The work done on the classifiers improved the number of answers in 1st position. On the other hand, the number of answers in the first 5 positions is more stable. A detailed study of badly retrieved images highlights small errors in image processing. It is on this point that we work.

This research was carried out with a grant of the *Conseil Général des Yvelines* (section 7.6). A version of this application is available for public at the *Archives départementales des Yvelines*.

6.2 Recovery of the drawing order

Participants: Laëtitia Rousseau, Jean Camillerapp, Éric Anquetil.

A way to do off-line handwriting recognition is to recover the drawing order in a letter or a word image, to generate an equivalent on-line signal and to use an on-line handwriting recognition system. Our approach [24] [23] to recover the drawing order consists in using handwriting knowledge.

The image extractor (cf. section 3.1) applies a thinning algorithm to extract a skeleton in the ribbon portions of the drawing. The handwriting drawing is then formalized by a graph where nodes correspond to intersections, end points or high curvature, and edges correspond to ribbon drawings.

Description of the method

Search for the starting and the ending points
Several candidates are selected as starting and ending points. The search for these candidates consists in two steps. First, their localization is determined: the edges extremities corresponding to them are chosen. The second step establishes the drawing direction: which point is the starting and which one is the ending.

Reconstruction algorithm

Our reconstruction algorithm is inspired from Kato and Yasuhara [KY99] [RAC04].

Best path choice
Once the reconstruction algorithm has been applied, we can still have several possible paths. Five criteria from *a priori* handwriting knowledge are used.
1 - A very curved edge cannot be traversed more than once.
2 - The loops have to be drawn in the good direction.
3 - The highest downstroke has to be maximized.
4 - The easiest directions have to be favored.
5 - The curvature of the letter has to be minimized.
Several paths are firstly eliminated with criteria 1 and 2. Then, a RBFN classifier is used to combine criteria 3, 4, and 5. The path with the best score is proposed to the on-line recognition system.

Evaluation

This approach has been validated on a database of handwritten letters including both on-line and off-line signals [VGLKB99]. An on-line recognition system has been used to recognize these letters (RESIF, cf. section 5.1).

We have removed multi-strokes letters from the database (23% of the database), stained images (4%), unexpected models (6%) like upper case letters, and letters where the starting and the ending points are hidden (9%). Our method has been tested on 5 800 letters images. Table 2 contains the on-line recognition system rate for the real on-line signal (On-line) and for the on-line signal deducted from the image (Off-line).

On-line	Off-line
88%	82%

Table 2: On-line recognition system rate

Future works will focus on handwriting made in several strokes, letters and words. This research is carried out with a grant of the *Conseil Général des Yvelines* (section 7.6).

6.3 Knowledge for Archival Table Structure Recognition

Participants: Isaac Martinat, Bertrand Coüasnon.

[KY99] Y. Kato, M. Yasuhara, "Recovery of Drawing Order from Scanned Images of Multi-Stroke Handwriting", *in : International Conference on Document Analysis and Recognition 5,*, p. 261–264, Bangalore, 1999.

[RAC04] L. Rousseau, E. Anquetil, J. Camillerapp, "Reconstitution du parcours du tracé manuscrit hors-ligne de caractères isolés", *in : 8ème Colloque Internationnal Francophone sur l'écrit et le Document, (CIFED'04)*, p. 123–127, La Rochelle, France, juin 2004.

[VGLKB99] C. Viard-Gaudin, P. M. Lallican, S. Knerr, P. Binter, "The IRESTE On/Off (IRONOFF) Dual Handwriting Database", *in : Proceedings of the Fifth International Conference on Document Analysis and Recognition (ICDAR'99)*, p. 455–458, 1999.

We designed a system that recognizes tables in archival documents. Many works were carried out on table recognition but very few on tables of historical documents. These documents are difficult to analyze because they are often damaged due to their age and conservation. On archival documents rulings can be broken, curved, and ink bleeds through paper, so flip side rulings are visible.

A general system was already built in EPF formalism (section 3.2) to analyze all kinds of table-forms. We showed [12] that this general system was unable to recognize these damaged archival tables, there is information missing in these document images, so more precise descriptions are necessary to recognize them. Specific systems for archival documents[CCL04] have already been designed in EPF (section 3.2), but they can be applied to a document structure type. To recognize other structure types, other descriptions must be given in EPF. EPF allows to create a system much faster than to develop completely a new recognition system, but descriptions in EPF are still quite long to define and accessible only to document analysis specialists.

To recognize an important number of different archival table types, the knowledge introduced by the user has to be fast and simple to write, so an archivist can give this specification. We want to adapt quickly the recognition system from one table kind to another. Therefore, we proposed a minimal and sufficient way of introducing external knowledge for archival table recognition [19]. We defined a system in EPF that takes in argument user knowledge. The specification given by the user is according to the quality of the document. If the document is of good quality, user needs to give few information, but if the document is of bad quality, more information has to be given to the system.

This new specification is easy and fast to write like `[ rowNumber 25 , colNumber 3 ]` The specification to give sizes for all rows and all columns is `[ rowMin 20, rowMax 150, colMin (cm 1.0), colMax (cm 8.0), rowNumber 25, colNumber 3 ]`. In this example column and row sizes are more constrained, but they can have some variations between documents. Sizes for a specific column can also be specified by the user.

To build a recognition system, we need to choose constraints like gap size between two line segments to form the same ruling. This choice is difficult because if we choose a gap size too small, documents can be undersegmented, otherwise if we choose gap size too big, documents can be oversegmented. For the detection of rulings separating columns and rows, at the first step, gap size is very small. If the system does not correctly recognize the document, the results do not match with the specification given by the user. Therefore the system reattempts the processing increasing gap size until to recognize with the user specification.

We show on a synthetic example (figure 8) with missing information, where row and column number specification is sufficient to recognize the document. A system not adapted to archival documents will recognize only two rows for this example but our system with user specification can detect the line segment for a row separator ruling.

In [19] we show how simple user specifications allow system to recognize very damaged documents.

[CCL04] B. Coüasnon, J. Camillerapp, I. Leplumey, "Making Handwritten Archives Documents accessible to Public with a Generic System of Document Image Analysis", *in: International Workshop on Document Image Analysis for Libraries (DIAL'04)*, p. 270–277, Palo Alto, USA, jan 2004.

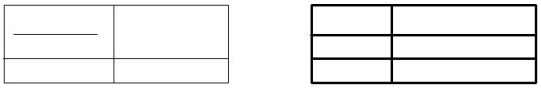


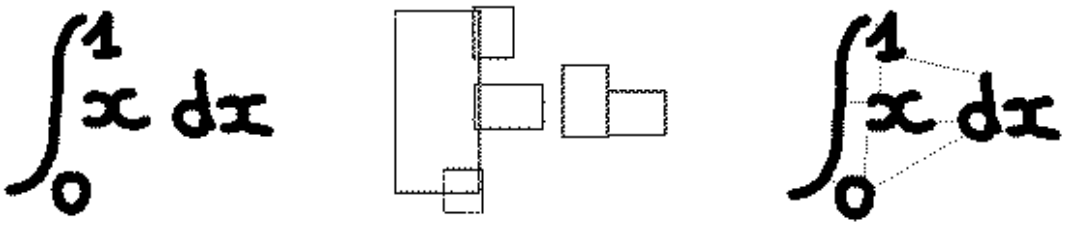
Figure 8: left: synthetic image illustrating missing information, right: structure with 3 rows and 2 columns to recognize.

This work has been done in cooperation with the *Archives départementales des Yvelines* in France, with the support of the *Conseil Général des Yvelines* (section 7.6).

6.4 Using neighbourhood graph in DMOS method

Participants: Aurélie Lemaitre, Bertrand Coüasnon, Ivan Leplumey.

DMOS method is based on a bi-dimensional description of a kind of documents. In the standard version of this method the relative position of elements is given with an approximation of components to their bounding box, which is sometimes not precise enough. The aim of our work has been to introduce new position operators, more precise than bounding boxes. These new operators are based on Voronoï tessellation, which offers a precise notion of relative position.



(a) Studied document

(b) Former bounding box based method

(c) New neighbourhood graph based method

Figure 9: Example of control of components in a document

Using a neighbourhood graph based on discrete distance

This implementation has been detailed in [LQ02]

[LQ02] I. Leplumey, C. Quéguiner, "Un graphe de voisinage basé sur l'utilisation des distances discrètes", *in : Actes du Colloque International Francophone sur l'Écrit et le Document (CIFED'02)*, Hammamet, October 2002.

The principle is to apply chamfer distance by propagation on the initial image. We obtain, as a result, a neighbourhood graph, labelled with distances, more complete than a mere Voronoï tessellation. This graph can be exploited for document analysis.

Integration of neighbourhood graph in DMOS method

The work has consisted in inserting, in the existing formalism, the data that could be extracted from the graph and that could be interesting, generally speaking, for structured document recognition [16].

The classical component detection mechanism, based on the common use of the operators `AT` and `TERM_CMP`, was sometimes mistaken. That is why our first idea was to propose another component detection mechanism based on neighbourhood graph. In order to complete the expression of component detections, we have introduced new basic conditions based on the graph.

Then, we wanted to use the fact that method DMOS makes it possible to analyse documents locally, to propose new information, based on the graph, depending on the context. Consequently, we have set up a bi-directional communication between the analyser and the neighbourhood graph. Indeed, the graph transmits numerical data to DMOS that translates it into symbolic data. Besides, DMOS sends information to the graph by specifying the context of analysis. This is one of the main interests of the method.

Validation of neighbourhood graph integration

The different tools based on neighbourhood graph, integrated with method DMOS, have been applied for the description of various kinds of documents. The aim was to prove their genericity and to determine the cases in which the neighbourhood graph could be useful.

We have first shown the interest of the basic operators. Then, we have experimented the statistic tools on the detection of words in papers. We have finally studied a definition of a grammar using both bounding boxes and neighbourhood graph on handwritten register of the 19th century.

Conclusion

We have set up a cooperation between neighbourhood graph that contains numerical information, and DMOS method, that can convert this data into symbolic information. It would be interesting to improve this use of neighbourhood graph by applying it on non-terminals of the grammar. Thus, we could deal with group of components' relative position.

6.5 An assessment of the structure recognition of 165,000 documents

Participant: Bertrand Coüasnon.

The DMOS method has been applied at now on the structure recognition of 165,000 military forms from the Archives départementales de la Mayenne and from the Archives départementales des Yvelines. With such an amount of documents, we have been confronted

with the reality of difficulties found in ancient documents. Indeed, it is quite impossible to get a representative sample of the difficulties on structure recognition. Therefore, even with a large sample built with archivists, documents we had to deal with were much more damaged than anticipated (see figure 10).

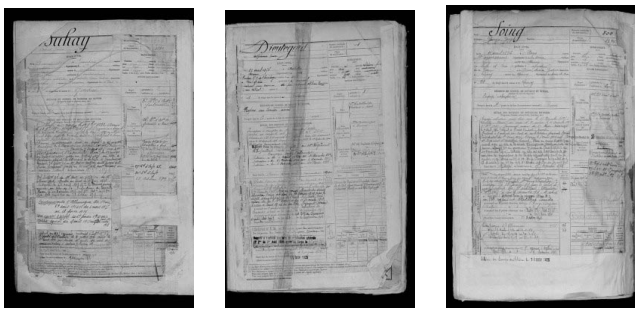


Figure 10: Examples of really damaged military forms. All the structure of these examples are nevertheless correctly recognized

Moreover, the way documents were digitized changed. Initially a flat-bed scanner was used, producing correct images with density corresponding to the one announced. Then a camera based scanner has been used, producing non flat images, introducing curves into the document geometry, with sometimes projective distortions (see figure 11). Moreover, images produced with these camera based scanners had a strong variation of the density compares to the one announced. Even worse, the variation of density was not the same on the x-axis and the y-axis.

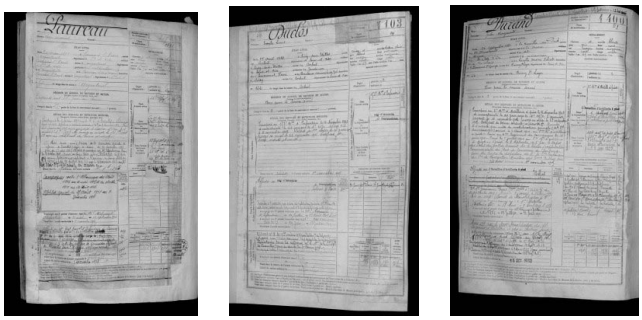


Figure 11: Examples of distortions produced by camera-based scanning. All the structure of these examples are nevertheless correctly recognized

All these strong and unanticipated difficulties could be overcome easily thanks to the genericity and adaptability of the DMOS method. Results on the cells position detection in the military forms were quite stable with previous tests: on 164,479 pages, 98.7% were correctly detected to separate public cells from protected cells (containing medical information). These results were obtained with no error. This experiment at a very large scale showed, once again,

the importance of generic methods in document structure recognition.

6.6 ILib 2.0: a feature extraction library

Participants: Yann Ricquebourg, Sébastien Rius[5], Éric Anquetil.

Concerning the studies of efficient classification and recognition methods, the Imadoc team lacks its own library of low-level extraction of information that should feed the developed classifiers. Therefore, we decided to federate our works and join our algorithms extracting and computing features from images in a common framework. This library has to be complete enough, with a stable API and easily reusable.

This platform represents a missing link in the components of the team to build entire processing chains.

To achieve this goal, we have been working the previous years on the definition and the progressive C++ implementation of such a library. After these early works, we decided this year, while still integrating new features, to deeply "clean" the first version and build a so-called ILib 2.0 version, focusing on:

- a more user-friendly restructuration of the library packages
- a portability to our target platforms (Solaris, Linux, Windows)
- a now stable and well-defined API
- a deep debugging and a large validation test stage

Those main purposes have been achieved and the library now proposes in a generic and stable way "classical" features often cited as required by common recognition systems (as surface, gravity center, curvature...), as well as more original or promising characteristic functions as Zernike moments, (concerning plain pixel images), or elliptic Fourier descriptors (concerning contour curves). Both these last functions have interesting hierarchical information properties, thus leading to a progressive separation of major factors and details, associated with a high power of information concentration since only few descriptors lead to a quite total representation of the characterized prototypes (see Figure 12). We recently added extraction of features concerning the holes and hollows determination, based on chanfrein distance and convex envelope.

Moreover, through a generic approach of our implementation, any extractor functions can be involved in a useful classical "zoning" computation, without requiring complementary programming of the extractor itself.

Additionally, the library also includes auxiliary data structures that are required by some feature extractors (Freeman chains, run-length representation...) as well as utility algorithms (adaptative binarisation, connected region labelling...).

Besides, we are developing a graphical user interface for the test and validation of those feature extraction functions. The GUI intends to be as generic as possible since it has been based on the common API we defined in the framework of the library.

[5] 4 months internship

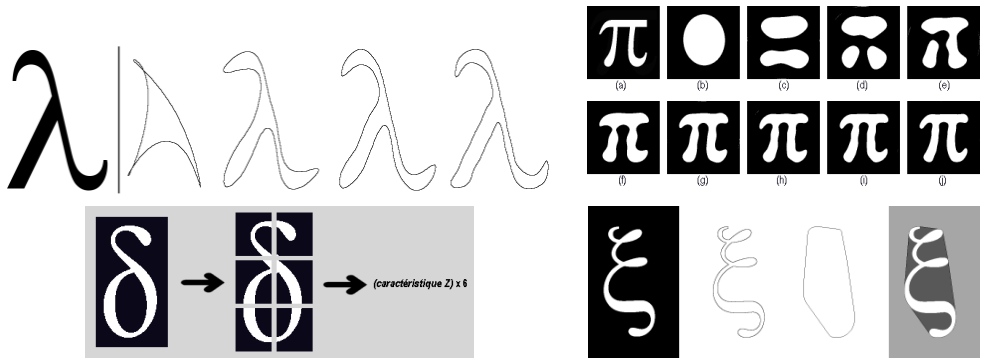


Figure 12: Some of the ILib functionalities: progressive decomposition using elliptic Fourrier descriptors, using Zernike moments, zoning, holes and hollows determination

6.7 Feedbacks and new perspectives for the Mélidis recognition system

Participants: Nicolas Ragot, Éric Anquetil, Guy Lorette.

The main drawback of generic classification approaches such as neural networks or support vector machines is that they act as monolithic black-boxes. Consequently, the results could be suboptimal for some problems because all the data is treated in the same way and because it is generally difficult to find the best parameters. The Mélidis [Rag03] recognition system was elaborated to offer another alternative for complex classification tasks. The first originality of the system comes from its specific hierarchical architecture that provides a better legibility of the system thanks to a high modularity and a formalization by a set of rules. Thus, optimization can be done by experts after the learning stage to adapt the system more specifically to the given problem. Moreover, the architecture and the formalism used allow the system to deal with a wide variety of problems since it can efficiently handle noise and variability, it can integrate a rejection process for unknown or ambiguous data and it was designed to be as compact as possible for integration on devices with limited resources such as PDAs or Smart Phones. All this was possible thanks to the previous researches on RESIFcar [2].

The Mélidis classifier, is based on an original hybrid architecture composed of both complementary modeling and a focusing mechanism that drives the learning and the recognition stages. The first level models each class to be recognized according to its intrinsic characteristics. Using this kind of modeling, the system is able to find automatically stable subclasses that compose the initial classes. The intrinsic models (description of the subclasses by fuzzy prototypes) are generated by an adapted fuzzy clustering algorithm like the Possibilistic C-Mean (PCM) on each class separately. Based on this modeling, the focusing mechanism determines which examples of a given class are a source of conflict with other classes. Indeed, not necessarily all the examples of a conflicting class are a source of confusion, but, more generally,

[Rag03] N. Ragot, *MÉLIDIS : Reconnaissance de formes par modélisation mixte intrinsèque/discriminante à base de systèmes d'inférence floue hiérarchisés*, PdD Thesis, Université de Rennes 1, October 2003.

the subclasses are. Finally, the aim of the second level of the hybrid system is to operate a progressive discrimination of the classes based on the result of the focusing mechanism. This is done by specific Fuzzy Decision Trees (FDT) dedicated to the discrimination task. For the recognition of unknown shapes, both levels are formalized homogeneously by fuzzy inference systems whose decisions are merged for the final decision.

Previous experiments have shown the ability of the Mélidis system to adapt itself to different classification problems and to provide high recognition rates with few parameters. These experiments have also highlighted the interest of each part of the system and especially of the focusing mechanism. It can be seen that the second level is always more or equally accurate than the first level and in all cases, the final decision provides always higher recognition rates than both levels. Moreover, as expected, the high performances of the second level are not only a consequence of the classification algorithm used by itself (FDT). They are also the result of the focusing mechanism used to drive the recognition process. Indeed, this mechanism provides at least 5% and up to 40% of relative error reduction, which represents an average of 24% on four different classification problems.

Recent experiments have also shown that the effect of the focusing mechanism is more limited when there are numerous classes that overlap strongly in the feature space. A solution to this problem that we are investigating, is to use, at the first level, a clustering algorithm that works in local feature subspaces. This is yet what is done at the discriminant level, where the FDT are implicitly working in there own feature subspace. The main objective is to find prototypes that are more specific and thus to give more robust information to the focusing mechanism so that it can simplify the discrimination at the second level.

Another subject we are still working on is the integration of the adaptation mechanisms described in section 6.10. The works made on this subject gave recently very interesting results using algorithms similar to the ones used at the first level. This adaptation mechanism can be used to handle the shapes that are varying with the time, but it can also be used to improve the performance of the first level, by using this new kind of learning after the original learning stage. Then, the next step will be to adapt this mechanism to the FDT of the discriminant level.

6.8 Design and evaluation of pen-based interfaces for small-size mobile devices

Participants: François Bouteruche, Éric Anquetil, Guy Lorette.

A part of our research focuses on the design of handwriting input methods for small-size devices such as smartphones or PDA. These methods are used to enter text on devices which are equipped with a sensitive screen and a pen. Two main parts compose these methods : a handwriting recognizer and a handwriting input interface.

The handwritten cursive character recognizer RESIFCar (see section 5.1) is the result of our researches on handwriting recognition. It has been embedded in smartphones currently sold in Europe. The hardware constraints (low CPU and memory) have implied the choice of an handwritten cursive character recognizer. Indeed a handwritten cursive word recognizer would currently be too expensive in term of resources to obtain acceptable response time from

a user point of view.

Despite the good performances of this recognizer on classical handwritten character benchmark, our industrial feedback shows a problem of acceptability due to the associated interface. Consequently, we focus our current works on the design of user-friendly handwriting input interfaces for small-size mobile devices. Our guideline is to include the user inside the design process. To do so, we collaborate with the Research Center in Psychology, Cognition and Communication of the University of Rennes 2 (CRPCC, see section 7.4). The design process follows an iterative implementation-evaluation cycle. At each step of the cycle, our implementation choices and the influence of the interface features are evaluated thanks to user experiments conducted by the CRPCC.

The first interface have been designed following a set of design principles extracted from our industrial feedback, the analysis of existing handwritten character input interfaces and works in cognitive psychology [14, 15]. It is called DIGIME 1 (DIGital Ink Micro Editor, see section 5.2). Its main features are an input in entire word context, a visual feeback area closed to the input area and a set of accentuation, punctuation and editing gestures.

previously written and recognized character until the user requests to clear them. The expected benefit is to prevent the share of attention between the input area and the application area when the user wants to check what he has previously written. Indeed, the share of attention between several sources of information has a negative effect on the user performance. closed to the input area as possible. As large distances between sources of information have negative effects on the user, the goal is to minimize the distance between the input area and the area allowing the user to check the result of the recognition (usually the application area).

The evaluation of DIGIME 1 has shown some problems in our implementation choices. First, the expected benefits of the input in entire word context, which was to prevent the share of attention between the application area and the input area, have been cleared because of the ink persistence. Indeed, keeping displaying the digital ink encourages the user to write cursive words, although RESIFCar is a cursive character recognizer. Consequently, a lot of recognition errors occur and the user's performances are low. In the same way, the presence of a visual feedback area has no effect on the user's performances. The negative effect of an additional source of information (the visual feedback area) is not counterbalanced by the benefits of the minimization of the distance between the sources of information. Finally, the functionalities of DIGIME 1 are not correctly exploited because of a lack of explanation. As a consequence the user can't reach maximal performances.

To overcome these problems and enhance the user's experience, we have proposed DIGIME 2 [15](see section 5.2). This interface merges the ink persistence and the visual feedback to provide a more user-friendly solution to the input in entire word context and the visual feedback. Moreover, on-line helps have been added to allow an easy start with the input method. These evolutions must now be evaluated.

6.9 Using the relative spatial context of stroke to drive the recognition process

Participants: François Bouteruche, Éric Anquetil, Guy Lorette.

Currently, on-line handwriting recognition systems are only able to recognize isolated entities (words, characters ...). As a consequence they don't completely fit the needs of a real-world pen-based application. For instance, in a handwriting note-taking application offering sketch recognition capabilities, it is impossible to discriminate strokes representing circles, letters 'o' and numbers '0' without any information on their context.

A new direction of our works focus on the explicit exploitation of the relative spatial context of strokes to drive the recognition of its shape and its interpretation. We propose a new and automatic approach based on the idea that only a sub-set of all possible symbols can be written in each specific spatial context. For example, if a stroke is written above a previously recognized Latin character, it is more likely an accent than another character. So, the recognition system can focus on accent recognition. Thus, we propose a hierarchical recognition method that we call "context driven recognition method" (CDR method) [13]. It is composed of two levels. The first one (the context level) models the spatial context of strokes and the second one (the shape level) models their shape. For a given stroke, the context level select a sub-set of all possible symbols so that the shape level focus only on the discrimination of these selected symbols. The aim is to simplify the recognition problem at the shape level by reducing the number of symbols to discriminate. Indeed, the more there are symbols to discriminate, the more the recognition problem is complex.

The main challenge of this work is to define an automatic modeling scheme to extract and formalize the spatial context. The aim of the proposed approach is to learn and model the relative spatial context of a stroke without any prior information on the application. To validate the method, we apply it on the problem of handwritten gesture recognition in an handwriting input method (see section 6.8). These gestures allow the user to modify or correct previously inputted characters. A former designed system used empiric rules to handle the spatial context. Each evolution of the input method implied tedious modifications of these rules. We replaced them by our CDR method and report in this paper the experimental results.

Since a classical Radial Basis Function Network achieve only a recognition rate of 59.12% using only shape features whereas it achieves a recognition rate of 96.24% using both context and shape features, there is no doubt on the usefulness of the spatial relative context in our recognition problem of 16 handwritten gestures. The proposed CDR method achieves a recognition rate of 96.64% that is equivalent to the RBFN performances but with a successful decomposition and simplification of the recognition process.

Our future works will focus on testing this method on more complex recognition problems and on the improvement of the context detection and modeling which is the core of the problem. A greater collaboration between the two level during the recognition process will keep our attention too.

This work is a part of a PhD thesis granted by France Telecom R&D (see section 7.3).

6.10 On-line writer adaptation for handwriting recognition using fuzzy inference systems

Participants: Harold Mouchère, Éric Anquetil, Nicolas Ragot, Guy Lorette.

The Imadoc team has already designed powerful fuzzy systems for on-line handwriting recognition like RESIF (see section 5.1) and MELIDIS (see section 6.7). We propose [21, 20] to improve the recognition rate by designing a technique for on-line adaptation to the writer style. Due to the on-line context, the adaptation has to be transparent, fast and embeddable on mobile device.

In these studies we use a simplified version of Fuzzy Inference System (FIS)(cf. section 3.3) which is composed of N rules, each rule defined by:

$$\mathbf{R}_i : \textbf{IF} \;\; X \;\; \text{is } P_i \; \textbf{THEN } s_1^i = a_1^i \; \textbf{and} \; \ldots \;\; \textbf{and } s_C^i = a_C^i$$

The premise is defined by the fuzzy prototype P_i. It represents the intrinsic properties of a class and is defined by a center (a vector) and a shape (a matrix) in the feature space. Its activation β_i represents how much the form to recognize X belongs to the prototype P_i. The conclusion describes each class by associating them a score S_C^i. The final decision of the FIS is taken by a combination of conclusion scores and premise activations to compute a score s_c for each class.

Once the FIS is learned, the adaptation affects both the conclusions of the rules with the gradient descent algorithm and the premises with the ADAPT method. ADAPT is based on on-line learning methods (LVQ [LHOK98] for the re-centering and covariance matrix estimation [Sch96] for re-shaping) but using the ADAPT learning rate δ_i (where b_c is the target score for the class c):

$$\delta_i = \beta_i * \left(\sum_{c=1}^{C} (b_c - s_c) * s_c^i \right).$$

The system was tested on 8 different writers. We compare [21] the ADAPT method with the existing methods LVQ and FLVQ[CL94]. The classical re-centering methods reduce the error rate by 56% but the re-centering ADAPT method reduces it by 70% and the complete ADAPT method (re-centering and re-shaping) achieves 85% of reduction (i.e. 4 on 5 errors are avoided) [20]. In average ADAPT can inprove the recognition rate from 88% to 98.2%. Furthermore the results have shown that our on-line adaptation strategy based on a data buffer is as efficient as an off-line strategy.

This work is granted by the CNRS and the Brittany Region (grant BDI CNRS/Region).

6.11 Lexical Post-processing for Handwritten Word Recognition

Participants: Sabine Carbonnel, Éric Anquetil, Guy Lorette.

We study a lexical processing for an analytical handwriting word recognition system based on three steps in order to improve the recognition result. The first step consists to introduce a

[LHOK98] J. Laaksonen, J. Hurri, E. Oja, J. Kangas, "Comparison of Adaptive Strategies for On-Line Character Recognition", *in: Proceedings of ICANN'98*, p. 245–250, sep 1998.

[Sch96] J. Schürmann, *Pattern Classification: a unified view of statistical and neural approaches*, Wiley-Interscience, 1996, ch. Recursive parameter estimation.

[CL94] F. Chung, T. Lee, "Fuzzy Competitive Learning", *IEEE Transaction on Neural Network 7*, 3, 1994, p. 539–551.

language model (characters n-grams) in the segmentation graph of the analytical recognition system. The induced error reduction rate is 8% for a 3-gram model.

The second and third steps are a lexical post-processing of the recognition proposals. The aim is to correct a proposal list according a lexicon and an edit distance. The second step is a lexicon reduction in order to reduce the search space during the post-processing. We propose two lexicon modellings based on word global features : the first modelling group together words according their shape and the second group together words according numerical characteristics with the HCA clustering algorithm. Accuracy and reduction rate obtained with this two lexicon modelling are high and the recognition rate is improved in comparison with a lexicon without specific modelling. The error reduction rate is 6% for the first modelling and 8% for the second, but this two modellings induce an increase of the memory requirement ($\times 3$ for the first modelling and $\times 2.4$ for the second).

The third step corrects the recognition proposal using an edit distance specific for handwriting errors correction. A classic edit distance does not allow to solve this specific problem. Some adaptations have been carried out to take this problem into account [CA03]. Even if good results are obtained, an empiric and manual estimation of the different edit operations and their respective costs is needed. We propose an automatic method to define edit operations and costs according to the recognition system properties: so it is possible to follow the evolution and improvement of the recognition module. This learning is based on the Boosting principle. The automatically learned distance permits to obtain slightly better results (error reduction rate: from 1% to 10% according the test set) than with the distance empirically computed and better results than a classic edit distance (error reduction rate: from 9% to 10%).

6.12 Statistical Language Models for On-line Handwritten Sentence Recognition

Participants: Solen Quiniou, Éric Anquetil, Sabine Carbonnel, Guy Lorette.

We investigate the integration of a statistical language model into the on-line recognition system RESIFMot (see section 5.1) to improve word recognition in the context of handwritten sentences. This technique is the most frequently used to incorporate linguistic knowledge and comes from speech recognition.

Sentence recognition aims at finding the most likely sentence $\hat{W}$ between candidate sequences W given the signal S of the handwritten sentence:

$$\hat{W} = \arg\max_{W} \log\left[p(S|W)\right] + \gamma \log\left[p(W)\right], \quad (1)$$

where $p(S|W)$ is the a posteriori probability of the signal S for the given sentence W and is estimated by the recognition system (here, RESIFMot) and $p(W)$ is the a priori probability of

[CA03] S. Carbonnel, E. Anquetil, "Lexical Post-Processing Optimization for Handwritten Word Recognition", *in : International Conference on Document Analysis and Recognition (ICDAR'03)*, p. 477–481, Edinburgh, August 2003.

the sequence W, given by the statistical *language model* (since these probabilities are small, their decimal logarithms are used instead). Furthermore, a *language weight* γ is introduced to balance the influence of the language model against the recognition system.

In [22], we show the impact of the language weight γ on the word recognition rate. We also compare n-gram (bigram and trigram) and n-class models (biclass and triclass whith statistical classes of words automatically built using Brown algorithm [BPdSL92]) in terms of word recognition rate as well as of number of parameters.

The Susanne Corpus [Sam00] was used for both construction of language models (about 5 500 sentences) and tests (80 sentences). For the tests, 20 sentences were written by 4 writers, using the handwriting ink grabbing application (see section 5.7). These sentences were then manually segmented into words in order to introduce no bias due to incorrect segmentation.

The language weight γ was seen to be important since the word recognition rate with the bigram model is 86.3% when $\gamma = 1$ (recognition system and language model are equally important) whereas it reaches 90.3% with its optimal value 0.3 (in the experiments, γ was set empirically).

Finally, we show the influence of a language model since with a bigram one the word recognition rate was improved from 82.5% (without language model) to 90.3% and with the trigram one it achieves 90.4%. The same recognition rates were arised with biclass model and triclass one respectively (the 2 models being made of 400 statistical classes) but the n-class models have twice as less parameters as the n-gram models.

6.13 Using generic method to design pen-based document composition systems

Participants: Sébastien Macé, Éric Anquetil, Bertrand Coüasnon.

The IMADOC team has designed powerful systems for on-line handwritten drawing recognition, making it possible to interpret characters, figures, geometrical shapes, etc. in isolated contexts. The next step is then to exploit these systems in order to interpret complex documents constituted of elements of different natures, which can not be interpreted by a single recognizer. Therefore, we started working on a new generic method taking the structure of documents into account in order to interpret it. Moreover, the analysis of the document must be realised directly during its composition, and display progressively its results to the user.

This new method is based on a formalism making it possible to describe the composition of a nature of structured documents, and a parser exploiting this knowledge and which interprets the user's drawings directly while the user composes his document.

[BPdSL92] P. F. Brown, V. J. D. Pietra, P. V. de Souza, J. C. Lai, "Class-Based N-Gram Models of natural Language", *Computational Linguistics 18*, 4, 1992, p. 467–479.

[Sam00] G. Sampson, "The Susanne Corpus", 2000, available at "http://www.grsampson.net".

A formalism to describe the composition of documents

We proposed a first version of formalism with a syntax based on context free grammars. As this formalism models the composition of on-line documents, we added new operators allowing:

- the management of chronological information;
- the representation of the document spatial structure;
- the recognition process driven by the structure analysis;
- the pen-based human-computer interaction.

We presented these new operators more in details in [17]. This formalism is externalised from the system, which means that it can be easily modified, without having to look at the rest of the code. Once compiled, it can be directly used in the system we developed.

A parser of handwritten drawings

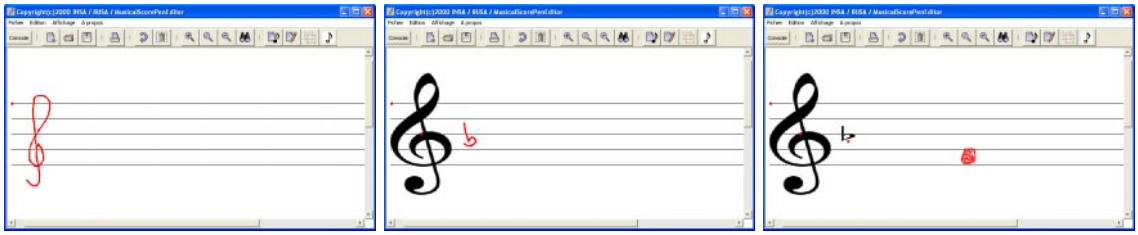
We developed a handwritten drawing parser which analyses the drawings of the user directly during the composition of documents thanks to the use of the knowledge modelled in the formalism. As it is able to deal with incomplete documents, it does not work the same way as a classical parser and must exploit chronological information. Given the structural context of a drawing, it is able to choose, which recognizers it is pertinent to call. The answer of the analysis can then be displayed to the user.

Using this method

So far, we only used this new method for one particular applicative context, which is the composition of musical scores. We developed a system making it possible to write musical scores in quite a same way as if it was on paper. Although it is still just a prototype, the editor already incorporates a lot of musical symbols (see section 5.8). In order to develop a system as user-friendly and as adapted to the musician needs as possible, the system is validated in collaboration with professional musicians from the MIAC (Music and Image: Analysis and Creation) of the Rennes 2 University. We presented a first version of this prototype in [18]. Figure 13 presents five screenshots of the system, corresponding to the composition of a document: the strokes are replaced with their corresponding symbols progressively. The fifth screenshot corresponds to a "novice mode" of the fourth one: it shows the user the areas where specific symbols have to be drawn.

Conclusion

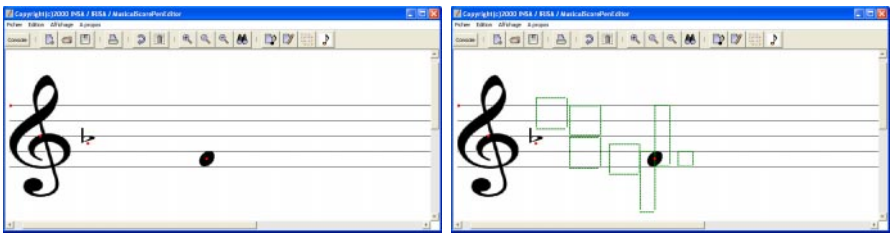
The future work will focus on improving the modelling of the document structure. To validate the genericity of the presented approach, we are going to work on the development of other pen-based application systems, such as mathematical formulas or electronic figures.



(a) First drawing

(b) Recognition of a treble key and second drawing

(c) Recognition of a flat and third drawing



(d) Recognition of a quarter note

(e) Novice mode corresponding to figure (d): boxes indicate where to draw specific symbols

Figure 13: Screenshots of the pen-based musical score editor prototype.

7 Contracts and Grants with Industry

7.1 EVODIA

Participants: Bertrand Coüasnon, Éric Anquetil, Jean Camillerapp, Guy Lorette, Grégory Maitrallain.

The Evodia company stems from the IMADOC team that developed solutions that are now mature:

- process complex documents (ancient degraded media, handwritten contents, mathematical formulas, and so on)
- to integrate pen interfaces in the management of digital documents.

Evodia has been created in October 2005, after being prizewinner - in the "emerging" category - of the 7th national competition to aid in the creation of innovative technology companies organized by the Ministry for Research, with the participation of Oséo-Anvar and the European Social Fund. Evodia his hosted by the Emergys incubator.

The target markets are first of all, the public, national and international, archive market,

then the market of industrial and mobile documents. The software solutions to address the archive market are now finalized and brought up to industrial standards.

Contact: yvan.ride@free.fr

7.2 Evaluation of the RESIF technology : Integration into Smartphone Device and Extension of the On-line Handwritten Word Recognition System to Large Word Vocabulary.

Participants: Éric Anquetil, Sabine Carbonel.

- Company: France Telecom R&D
- Contract: INSA 3012, INSA 3015

A collaboration with France Telecom R&D company started in 2003. The aim of this research collaboration is to evaluate RESIF technology according to two main axes:

- The first axis consists in the study of the integration possibilities of the on-line handwriting recognition systems RESIFCar and RESIFMot (cf. section 5.1) into smartphone devices;
- The second axis focuses on the task of lexicon post-processing (cf. section 6.11) to deal with large word vocabulary in the context of limited resources. The aim of this work is to explore the combination of different approaches to reduce time computing and memory resources involved in lexicon post-processing.

7.3 Automatic Context Modelling for On-line Pen-Based Interface Design.

Participants: Éric Anquetil, François Bouteruche.

- Company: France Telecom R&D
- Contract: INSA 4016

This new collaboration with France Telecom R&D company started in October 2004. We work on the interpretation of pen input (electronic ink) as captured by pen based interfaces (cf. section 4.3). The interpretation of handwritten shapes requires the knowledge of their specific context : the recognition process is driven by the context. The aim of this research is to study the possibilities of an automatic context modelling and its impact on the interpretation process. The first considered approaches are based on fuzzy logic modelling and especially fuzzy decision trees.

7.4 Evaluation of Pen-based Interface Ergonomic Quality.

Participants: Éric Anquetil, François Bouteruche.

- Company: Research Center in Psychology, Cognition and Communication of the University of Rennes 2
- Contract: INSA research collaboration

We have worked in collaboration with the Research Center Psychology, Cognition and Communication of the University of Rennes 2 since January 2004. The aim of this research collaboration is to validate experimentally the quality of the pen-based interfaces for handwriting input on mobile devices. This year, the study focuses on the impact of two components on user's performances and satisfaction(cf. section 6.8): the spatial contiguity respect between the visual feedback area and the input area and the conservation of the previously written character in the input area.

7.5 Pen-based Musical Score Editor.

Participants: Éric Anquetil, Sébastien Macé.

- Company: Research department in Arts, Humanities and Communication of the University of Rennes 2 (MIAC: Music and Image: Analysis and Creation)
- Contract: INSA research collaboration

The collaboration with the research department in Arts, Humanities and Communication of the University of Rennes 2 started in January 2004. The aim of this research collaboration is to validate by musical experts the functionalities and the ergonomic choices we made, in order to design a system as close as possible to the musician needs. This collaboration with experts on music has conducted this year to optimize the prototype of pen-based musical score editor we have developped (see section 5.8).

7.6 Access to Handwritten Archives Documents

Participants: Bertrand Coüasnon, Éric Anquetil, Jean Camillerapp, Ivan Leplumey, Grégory Maitrallain, Isaac Martinat, Laëtitia Rousseau.

- Partner : *Conseil Général des Yvelines, Archives des Yvelines*
- Contract : INSA 3009, INRIA 1 02 C 0602 00 31406 01 1

This research project started 2002 and is planned until October 2007. It is about making handwritten archives documents accessible to public. In collaboration with the *Archives des Yvelines*, Imadoc is working on three axes:

- document structure recognition and handwriting recognition to automatically produce annotations. These annotations offer the possibility to access by handwritten content to archives documents;
- a web platform to manage images, automatic annotations and collective annotations made by readers;
- a pen-based interface build on this platform to offer a new way of interacting with digital documents and paper documents.

The platform for image document annotations (section 5.6) with an automatic access by handwritten last names on military forms (section 6.1) is available in the reading room of the *Archives départementales des Yvelines* and on Internet (`http://www.archives.yvelines.fr`). Readers, after retrieving the right document by a request on handwritten names are able to add collective annotation on it. For the moment 105,000 pages have an automatic access by handwritten last names among 430,000 pages of military forms which are available to collective annotation. 1,450,000 double pages of register of births, marriages and deaths are also available on this platform for annotation.

The pen-based interface of the platform is also running on two reading desks of the future. These prototypes are made of a graphical tablet built in a wooden desktop, and a digital pen to interact with the interface.

New results on this project are presented in sections 6.1, 6.2, 6.3 and 6.5.

7.7 Integration of large textual annotations in the platform for annotation

Participants: Bertrand Coüasnon, Ivan Leplumey.

- Partner : *Ville de Lannion, Archives*
- Contract : INSA 5009

The collaboration with the Archives of the city of Lannion will test a way to link 52,000 textual births, marriages and deaths certificates with images of pages from original registers. Linked annotations will then be introduced in the platform for collective annotation.

7.8 Access to naturalization decrees

Participant: Bertrand Coüasnon.

- Partner : *Archives de France, Centre Historique des Archives Nationales*
- Contract : INSA 5010

This research project is about making a fast leaf-through capability using automatic annotations on naturalization decrees from the end of the nineteenth and the beginning of the twentieth centuries. The structure of these documents is very weak - merely paragraphs of handwritten text. Automatic annotations on the position of handwritten last names will be validated on 150,000 documents.

8 Other Grants and Activities

8.1 National initiatives

8.1.1 ACI Madonne (Ministry Grant)

Imadoc is involved in the ACI (*Action Concertée Incitative*) Madonne (*MAsse de DOnnées issues de la Numérisation du patrimoiNE*). This project is about large data set produced by digitizing cultural heritage. Partners of this ACI are: L3I (La Rochelle), PSI (Rouen), LI (Tours), Loria (Nancy), LIRIS (Lyon), Irisa (Rennes).

8.1.2 RTP

Participation of Imadoc to the *Réseaux Thématiques Pluridisciplinaires* :

- RTP-DOC (STIC) :
  - Atelier "numérisation" : Bertrand Coüasnon, Isaac Martinat and Laetitia Rousseau

8.2 International initiatives

Guy Lorette is in charge of international relationships at IFSIC, and member of the international relations group of the Rennes 1 University.

9 Dissemination

9.1 Leadership within scientific community

9.1.1 Editorial board

Jacques André is in the editorial board of:

- *Document numérique*,
- *Les Cahiers GUTenberg*,
- Tools for Computer Typesetting, *Typography series* from Addison-Wesley.

9.1.2 Programme comittee

Members of the team are in Programme comittee of:

- EuroTEX'2005 (30 th TEX European Conference), March 2005,
- GREC 2005 (International Workshop on Graphics Recognition), Aug. 2005
- IGS 2005 (International Graphonomics Society), June 2005.

- ICDAR 2005 (International Conference on Document Analysis and Recognition), Aug. 2005.
- AXMEDIS 2005 (International Conference on Automated Production of Cross Media Content for Multi-channel Distribution), Nov. 2005.
- CIFED 2006 (Colloque International Francophone sur l'Écrit et le Document). Sept.2006.
- ICDAR 2007 (International Conference on Document Analysis and Recognition), Aug. 2007.

Members of the team are reviewers of: GRETSI 2005, Sept. 2005 and Advanced Concepts for Intelligent Vision Systems ACIVS 2005, Sept. 2005.

9.1.3 Organizing comittee

- Guy Lorette and Éric Anquetil are in the organizing comittee of: IWFHR'10 (International Workshop on Frontiers in Handwriting Recognition), Oct. 2006, `http://www.irisa.fr/iwfhr10`.
- Solen Quiniou and Harold Mouchère are in the organizing comittee of: MajecSTIC'05 (manifestation des Jeunes Chercheurs francophones dans les domaines des STIC), Nov. 2005.

9.1.4 Member of scientific society

Jacques André is a member of the board of directors of the society GUTenberg.

Guy Lorette is :

- member of the scientific council of PSI (Université-INSA de Rouen) and of E3I computer laboratory (E3I-Université de Tours);
- member of the ASTI society.

Éric Anquetil, Bertrand Coüasnon, and Guy Lorette take part in the working group " 5.2-Écrit ", topics 5: Communication, of GDR-PRC I^3 (Information, Interaction, Intelligence).

Jean Camillerapp, Guy Lorette, Éric Anquetil, and Bertrand Coüasnon take part in the activities of the society GRCE : " Groupe de Recherche en Communication Écrite ".

9.1.5 Forum list

9.2 University education

The team is mainly made up of teachers and those are very implied in activities of teaching. But a majority of lectures are not rattached to this research topic, so they are not quoted here.

- Jacques André gives a lecture on Unicode at MASTER, university of Paris Val de Marne (Paris XII)
- Hélène Richy is in charge of the MASTER-PROFESSIONAL MITIC (*Méthodes Informatiques et Technologies de l'Information et de la Communication*) at IFSIC, University of Rennes 1 and gives lectures in this cursus.
- Guy Lorette gives lectures at MASTER-RESEARCH *d'informatique* and MASTER-PROFESSIONAL MITIC at l'IFSIC, University of Rennes 1.

9.3 Participation in conferences, seminars, invitations

- Éric Anquetil was invited for a presentation to the Innovation Days of France Telecom R&D, June 2005;
- Éric Anquetil was invited for a presentation to the IRCAM's forum. Oct. 2005;
- Bertrand Coüasnon, Isaac Martinat and Laetitia Rousseau participated to the joint workshop between the Atelier Numérisation du Programme Société de l'Information "Document Numérique" and the ACI Masses de Données MADONNE. May 2005;
- Bertrand Coüasnon was invited for a presentation to the Web Conference of the RTP Doc - Atelier "Numérisation". Jan. 2005;
- Bertrand Coüasnon was invited for a presentation to the national conference "Numériser et valoriser en région le patrimoine écrit et graphique". Sep. 2005;
- Bertrand Coüasnon organized a presentation day for archivists on "Nouveaux systèmes d'accès aux documents manuscrits numérisés" at Inria-Rocquecourt and at the *Archives des Yvelines*. Jun. 2005.

10 Bibliography

Major publications by the team in recent years

[1] J. ANDRÉ, M.-A. CHABIN (editors), *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, Spain, July 1996.

[3] B. COÜASNON, J.-P. DALBÉRA, H. EMPTOZ (editors), *Numérisation et patrimoine, Numéro spécial de Document Numérique, 7*, 3-4, Hermes-Lavoisier, 2003.

[4] B. COÜASNON, *Dealing with noise in DMOS, a Generic Method for Structured Document Recognition: an Example on a Complete Grammar, LNCS, 3088*, Springer-Verlag, 2004, p. 38–49.

[5] V. P. d'Andecy, J. Camillerapp, I. Leplumey, "Analyse de partitions musicales", *Traitement du Signal 12(6)*, 1996, p. 653–661.

[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, "Handwriting recognition or reading? What is the situation at the dawn of the 3^{rd} millenium?", *International Journal on Document Analysis and Recognition 2*, 1, 1999, p. 2–12.

[8] L. Pasquer, G. Lorette, "Système de perception et d'interprétation de formes structurées (spi)", *RTSI-TSI (Technique et science informatiques)*, Vol.22/7-8, 2003, p. 879–902.

[9] N. Ragot, E. Anquetil, "Système de classification hybride interprétable par construction automatique de systèmes d'inférence floue", *Technique et science informatiques 22*, 7-8, 2003, p. 853–878.

Books and Monographs

[10] F. Mittelbach, M. Goossens, *LaTeX Companion, deuxième édition*, Pearson, Paris, 2005, translator Jacques André and Jean-Côme Charpentier.

Doctoral dissertations and "Habilitation" theses

[11] S. Carbonnel, *Intégration et modélisation de connaissances linguistiques pour la reconnaissance d'écriture manuscrite en-ligne*, PhD Thesis, INSA de Rennes, dec 2005.

Articles in referred journals and book chapters

[12] B. Coüasnon, "DMOS, a Generic Document Recognition Method: Application to Table Structure Analysis in a General and in a Specific Way", *International Journal on Document Analysis and Recognition, IJDAR*, 2005, to appear.

Publications in Conferences and Workshops

[13] F. Bouteruche, E. Anquetil, N. Ragot, "Handwritten gesture recognition driven by spatial context of strokes", *in : Proceedings of the 8th International Conference on Document Analysis and Recognition*, B. Werner (editor), *2*, IEEE Computer Society, p. 1221–1225, Seoul, Korea, August 2005.

[14] F. Bouteruche, G. Deconde, E. Anquetil, E. Jamet, "Conception et évaluation expérimentale d'interfaces de saisie stylo pour systèmes mobiles de petites tailles", *in : Actes de la 17e Conférence Francophone sur l'Interaction Homme-Machine*, p. 235–238, Toulouse, France, September 2005.

[15] F. Bouteruche, G. Deconde, E. Anquetil, E. Jamet, "Design and evaluation of handwriting input interfaces for small-size mobile devices", *in : Proceedings of the 1st Workshop on Improving and Assessing Pen-Based Input Techniques*, p. 49–56, Edinburgh, Scotland, September 2005.

[16] A. Lemaitre, B. Coüasnon, I. Leplumey, "Using a neighbourhood graph based on Voronoï tessellation with DMOS, a generic method for structured document recognition", *in : proceedings of GREC, Sixth IAPR International Workshop on Graphics Recognition*, p. 260–271, Hong Kong, China, August 2005.

[17] S. Macé, E. Anquetil, B. Coüasnon, "A generic method to design pen-based systems for structured document composition: Development of a musical score editor", *in : Proceedings of the 1st Workshop on Improving and Assessing Pen-Based Input Techniques*, p. 15–22, Edinburgh, September 2005.

[18] S. Macé, E. Anquetil, E. Garrivier, B. Bossiss, "A Pen-based Musical Score Editor", *in : Proceedings of ICMC*, p. 415–418, Barcelona, Spain, September 2005.

[19] I. Martinat, B. Coüasnon, "A Minimal and Sufficient Way of Introducing External Knowledge for Table Recognition in Archival Documents", *in : Proceedings of GREC, Sixth IAPR International Workshop on Graphics Recognition*, p. 194–205, Hong Kong, China, August 2005.

[20] H. Mouchère, E. Anquetil, N. Ragot, "On-line Writer Adaptation for Handwriting Recognition using Fuzzy Inference Systems", *in : Proceedings of the 8th International Conference on Document Analysis and Recognition (ICDAR)*, B. Werner (editor), *2*, 1075-1079, IEEE Computer Society, Seoul, Korea, August 2005.

[21] H. Mouchère, E. Anquetil, N. Ragot, "Writer Style Adaptation of On-line Handwriting Recognizers: A Fuzzy Mechanism Approach", *in : Proceedings of the 12th Conference of the International Graphonomics Society (IGS)*, A. Marcelli, C. De Stefano (editors), p. 193–197, Salerno, Italy, June 2005.

[22] S. Quiniou, E. Anquetil, S. Carbonnel, "Statistical Language Models for On-line Handwritten Sentence Recognition", *in : Proceedings of the 8th International Conference on Document Analysis and Recognition (ICDAR'05)*, *II*, p. 516–520, Seoul, Korea, August 2005.

[23] L. Rousseau, E. Anquetil, J. Camillerapp, "Reconnaissance de l'écriture manuscrite hors-ligne par reconstruction de l'ordre du tracé", *in : Proceedings of MAJECSTIC*, Rennes, France, Novembre 2005.

[24] L. Rousseau, E. Anquetil, J. Camillerapp, "Recovery of a Drawing Order from Off-Line Isolated Letters Dedicated to On-Line Recognition", *in : Proceedings of ICDAR*, *2*, p. 1121–1125, Seoul, Korea, August 2005.