Project-Team: IMADOC

Interprétation et Reconnaissance d'Images et de Documents
Theme Cog
Rennes

February 12, 2007

Contents

1 Team

Team Leader

Jean Camillerapp [Professor Insa]

Administrative Assistants

Stéphanie Lemaile [TR Inria]

Diana Gaudin [TR Inria to 01 Jul 2006]

Staff members (Insa)

Éric Anquetil [Associate Professor]

Bertrand Coüasnon [Associate Professor]

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]

Sabine Carbonnel [Lecturer to 31 Aug 2006]

Laetitia Rousseau [Lecture from 01 Oct 2006]

Ph.D. Students

Laetitia Rousseau [Insa research grant to 30 Sep 2006]

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]

Sébastien Macé [Région/Insa research grant]

Christophe Renaudin [Menrt research grant from 01 Oct 2006]

2 Overall Objectives

The research topics of the project concern *the written communication* and *the engineering of documents* under various aspects: *analysis, recognition, interpretation of digitized documents, man-document interaction, etc.* . 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 related activities.

Facing the multiplicity and the diversity of applications, the usual solution consists of 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, of which two are more methodological in nature and three are more application oriented.

- **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 should make 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 task of 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 for *interpretation* for signals and of document images, the project studies methods of dynamic and multi-contextual recognition based on visual indices. For the modeling of contextual knowledge three complementary approaches are considered: one based on rules for symbolic knowledge systems, another one based on fuzzy inference systems for vague knowledge, while 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 obtain 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 advantages of offering several possible levels of reading, to locate where the errors stay remain and to lead to transparent systems of which the working modes can be analyzed and the performances can be optimized.

- **Handwriting**
  We study recognition systems for *handwritten* documents, either during their generation (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 areas of interest (*multilingual and omni-script writers recognition systems*). Nevertheless, it is always possible to specialize a given system and to build from it a dedicated software.

- **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 put enormous corpora of digitized documents at the public disposal.

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

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

  Our research is concerned with 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 as follows:
  - 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 focusses 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 digitalized 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 digitalization in gray levels allows to implement adaptative binarization algorithms, often based on cooperation between edge detection and region growing. They provide a local

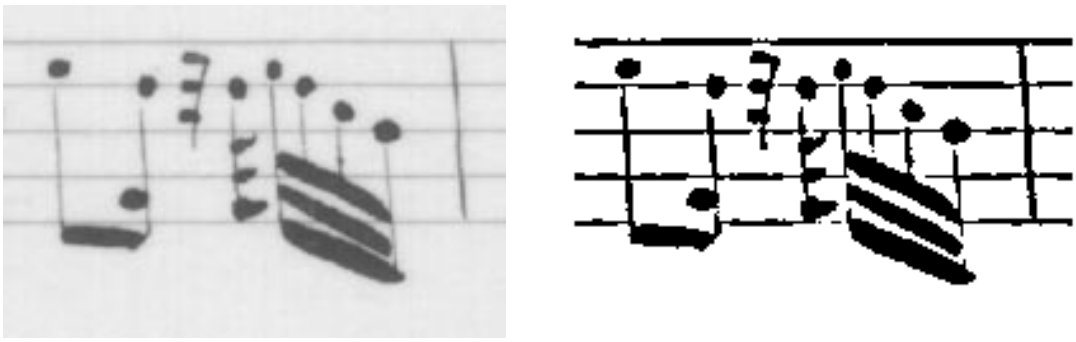
(a) Threshold of 210 : allowing the detection of thin lines

(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.

interpretation of gray values which separate well the various objects in the document (cf figures 1 et 2).



(a) Initial image in gray level

(b) Binarized image

Figure 2: Binarization with an adaptative threshold

Measurements of object 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.

Gray level image skeletonization

Skeletonization algorithms have been developed to thin images of lines. But in document images, it is necessary to distinguish in the layout (figure 3) the area really skeletonizable (also called regular area) from the singular areas, mainly localised in crossings in which the

concept of median axis does not have any practical relevance.

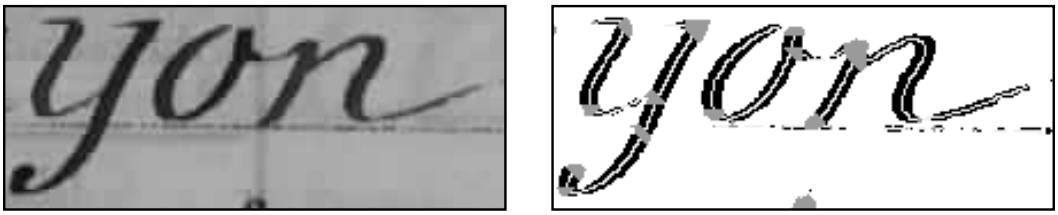


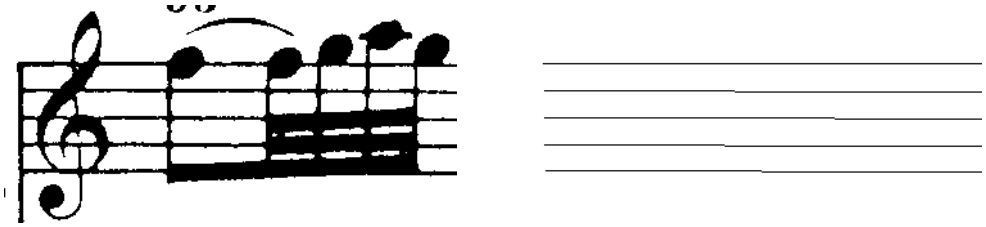
Figure 3: Processing of an image from an old document. The skeleton appears in white in the medium of the line and singular areas are in dark gray.

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 great importance in documents, but they are likely to interfere among 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 parameter identification from an ordered series of measures. In case of lines, the model is reduced to the thickness of the line, its slope and the equivalent of the ordinate at zero point. The measure is computed 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].

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

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

Thus, a run-length in the image can result from overlap of several drawings. Because of such situation, the information obtained from the covariance matrix is of great importance.

In this approach we can also use a measure of 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 e.g. 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 time-consuming task. Defining a generator of recognition systems for structured documents to avoid this costly rewriting work is a real challenge for the researchers.

Moreover, improving recognition quality is vital for an industrial processing of documents. This can be done, for example, by solving document segmentation problems more efficiently.

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

- a generic system should be able to change according to changes from one document to another: the *a priori* knowledge on each document;
- improving segmentation on documents needs to employ *a priori* knowledge in the system as much as possible.

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

- the new grammatical formalism EPF (Enhanced Position Formalism), which can be seen as a description language for structured documents. EPF makes it possible at the same time, to generate a graphical, a syntactic or even a semantic 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 forms of segmentation with the help of context and improve the recognition accuracy.

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 the 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` relative to `A`, we find `B`.

Here `&&` 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 that we developed are:

- changing the parse structure during parsing for contextual segmentation. The parse 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 parse 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, etc.

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 cause of errors causes is very difficult for complex system. Handwriting recognition is a typical example of this kind of problem. Different methodologies are used to try to solve this difficult problem. The approaches based on neural networks or methodologies based on hidden Markov models lead typically to "opaque" systems. At present, it is 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 follows:

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

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];
- $\perp$ 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 values which can be interpreted as degrees of "typicality", contrary 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 maintain an explicit description of each model, we hierarchically structure fuzzy rules with a view 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 that is going to be modeled. Based on the previous observations, we set out

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

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

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 improved. This new system can be considered as an extension of RESIF. The main goal is to obtain a more generic recognition system that is able to deal with different kinds of classification problems without needing *a priori* knowledge and without loss of transparency, so that optimization can be done afterwards by experts [RA04].

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 types of documents. Documents we have already worked on are:

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

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

Moreover, a grammatical description of a kind of document (made with EPF) can be either general, in order to recognize a class of documents (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 systems can be used in many ways:

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

[RA04] N. Ragot, E. Anquetil, "MELIDIS: Pattern recognition by intrinsic/discriminant dual modeling based on a hierarchical organization of fuzzy inference systems", *in : 10th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems, (IPMU'04)*, *3*, p. 2069–2076, Perugia, Italy, juillet 2004.

- 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 indices and annotations for an automatic access by content to documents (section 5.6);
- Detection of specific areas in a document to improve its access.

4.2 Handwriting recognition systems

Recently, there has been a new increase in the applications of handwriting recognition in the domains of automatic processing of paper documents (*off-line recognition*) as well as 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 accurate 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 faxes. In this domain, we have mainly concentrated our efforts to guarantee a high degree of robustness and confidence in the results to be obtained on automatic processing of handwriting. This was done to avoid any risk of error.

In *on-line recognition*, a huge market has arisen due to recent cheap availability 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 systems. The pen input may be available on any kind of device

where pen interface is available: mobile-device screens, tactile screens, electronic paper, or tablet.

Topics of interest for 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 gesture 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 isolated handwritten characters: Latin alphabet, digits and special symbols.
- RESIFMot is the software for unconstrained cursive handwritten word 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.

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

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 [BDAJ05]. Used 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 6.5), DIGIME has become a full application and is in its second version.

The originality of DIGIME is to offer 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.

It is currently running on Windows Mobile Pocket PC 2003, Pocket PC 2003 2nd Edition and Pocket PC 2005 operating systems. Thanks to its architecture, it can be easily adapted to other mobile operating systems, such as Windows Mobile Smartphone 2005 or Palm OS.

Figure 5: Screenshot of DIGIME.

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 having also a reject option.

[BDAJ05] 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.

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 reading 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 6.2);

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[4], Éric Anquetil.

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

[4]Today associate professor at University of Tours

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 networks or multi-layer perceptrons and moreover with a higher legibility. This first version is also able to construct a forest of decision trees to be more accurate on multi-class problems.

FDT-L is currently used by the Mélidis system [RA04]. Work is also in progress to extend the capability of the library by taking into account more clustering algorithms and to provide a visual interface.

5.6 A platform for image document annotation

Participants: Bertrand Coüasnon, Grégory Maitrallain[5], 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 maintain relations between textual and geometric annotations in order 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 move 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 annotations (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.

[5] Today ingenior at Evodia Company (cf. section 7.1)

[RA04] N. Ragot, E. Anquetil, "MELIDIS: Pattern recognition by intrinsic/discriminant dual modeling based on a hierarchical organization of fuzzy inference systems", *in: 10th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems, (IPMU'04)*, *3*, p. 2069–2076, Perugia, Italy, juillet 2004.

Structured search or full text search is possible on all annotations whatever 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 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 [MAC05]. 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.

[MAC05] 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.

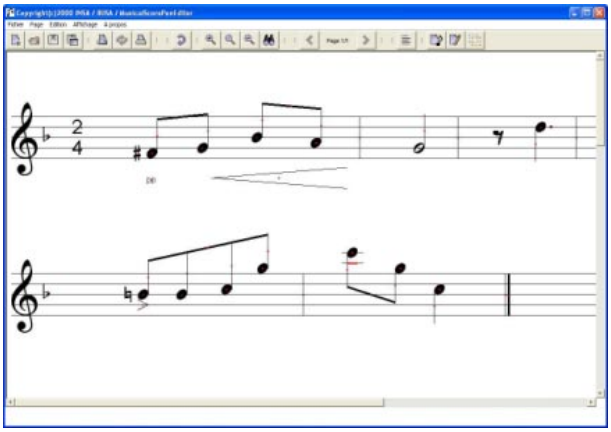


Figure 6: Screenshot of the pen-based musical score editor.

6 New Results

6.1 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 final purpose is the indexation of archive documents by their handwriting content.

Our approach [31] to recover the drawing order consists in using handwriting knowledge. For the moment, it concerns only isolated lowercase letters.
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 or end points, and edges correspond to ribbon drawings.

Description of the method

Letters can be mono-stroke or multi-strokes.

Search for the starting and the ending points
Several candidates are selected as starting and ending points. First, the graph is decomposed. Secondly, several localizations are proposed. Then, the drawing direction is established.

Reconstruction algorithm
The reconstruction algorithm generates one or several paths for each couple of starting and ending points. It is inspired from Kato and Yasuhara [RAC04].

[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*

Best path choice
Once the reconstruction algorithm has been applied, we can have several possible paths. Criteria from *a priori* handwriting knowledge are used to eliminate impossible paths.
1 - A very curved edge cannot be traversed more than once (paths eliminated : 8.3%).
2 - The loops have to be drawn in the good direction (paths eliminated : 13%).

Paths are sorted combining several criteria.
1 - The highest downstroke has to be maximized.
2 - The easiest directions have to be favored.
3 - The curvature of the letter has to be minimized.

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]. The on-line recognition system used to recognize these letters is RESIFCar (cf. section 5.1).

Models of letters have been learned on 6 410 letters. Our method has been tested on 2 139 letters images. Table 1 contains the on-line recognition system rates for the real on-line signal (On-line) and for the on-line signal constructed from the image (Off-line).

On-line	Off-line
88%	81.3%

Table 1: On-line recognition system rates

Future works will focus on words and indexation of archive documents. This research is carried out with a grant of the *Conseil Général des Yvelines* (section 7.6).

6.2 Language for Archival Table Structure Recognition

Participants: Isaac Martinat, Bertrand Coüasnon, Jean Camillerapp.

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 system needs to have an user specification to succeed in recognizing these difficults documents.

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.

A first system [25] has been developped to recognize tables by using external knowledge, but this system was only for tables with ruling separators between columns and rows. The user could not specify the hierarchy of columns or rows and the number of columns and rows had to be fixed.

Thus we propose a new language more general to describe tables. Many descriptions in different languages exist for tables like in XML but they are for edition. For edition, a description is fixed and every element of a table (cell,row,column) is described and every cell is associated with data. Like Wang's model [Wan96] we propose a description composed of two parts, a logical one and a physical one. We want to describe tables with the row and column organization without data of a flexible way. The description given by the user can be very precise (the number of columns and rows is fixed, size of columns and rows are given) or general (table containing n columns and n rows). The language we propose allows us to define a logical part with the hierarchy for rows and columns, logical relationships between columns and rows. It also defines an optional physical part with sizes and separator types. The specification of the physical part is visual to make easier to write it and to read it. We give an example to describe census tables (Fig.7) , words in capital letters are specific to the language :

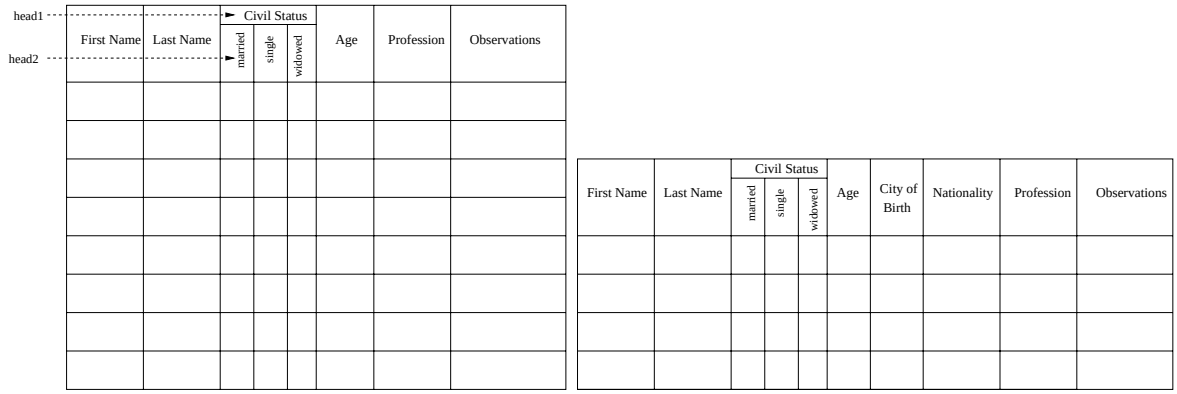
```
//Hierarchy of columns and rows
TABLE census = cols * rows.
COL cols = firstName, lastName, civilStatus, info.
COL civilStatus = married, single, widowed.
ROW rows = boxHead, person.
ROW boxHead = head1 , head2.

//Logical relationships between cols and rows and
//specification for each row the columns that are contained in.
COLS_IN_ROW head1 = civilStatus.
COLS_IN_ROW head2 = married, single, widowed.
COLS_IN_ROW boxHead = firstName, lastName, civilStatus, info.

//Physical specification
//columns separated by rulings represented by |
//rows separated by rulings represented by -(---------)
SEPCOL cols = firstName | lastName | civilStatus | REPEAT+(1, info |) .
SEPCOL civilStatus = married | single | widowed.
SEPROW rows =
         boxHead
    -(----------------)
         REPEAT+(1, {
           person
    -(----------------) }).
```

[Wan96] X. WANG, *Tabular abstraction, editing, and formatting*, PdD Thesis, University of Waterloo, 1996.

```
SEPROW boxHead =
         head1
    -(------------)
         head2.
```



(a) (b)

Figure 7: Examples of Census tables which can be recognized with the same description in our proposed language

The system uses the user specification to recognize document images. From the images, the system gets a set of detected rulings and from this set, it has to recognize a table. This set is not totally representative of the specification. Some false rulings are detected, flip side rulings, rulings from the handwritting and some rulings can be missing. In the case of broken rulings, several rulings are present instead one. We used final intersections, intersections with one or two extremities to select rulings from this set in using user specification. Thus from the user specification, a set of intersections is generated. This set is composed of the intersections that would have to be found on the image. We used these final intersections because they are strong dependencies and thus, less sensitive to noise.

The system tries to match for each row, specification intersections and image intersections to detect column separator positions. The system can use the previous column positions detected from the previous rows to search in precise image zones the intersections of the analyzed row. When the matching fails, it is delayed and when it is released, it can succeed in using the column positions that have been detected during the time between the delay and the release. This system is being tested on sample examples and we have to validate it on a more important number of documents and in using different descriptions.

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).

6.3 Introduction of multiresolution in document recognition

Participants: Aurélie Lemaitre, Jean Camillerapp, Bertrand Coüasnon.

In the field of document structure recognition, we propose to introduce a notion of perceptive vision. Indeed, when you watch a document, you will not see the same thing if you watch it very close or very far. However, both vision can be interesting for structure recognition. That is why we suggest to combine both informations extracted at low and high resolution.

More precisely we first worked on the extraction of text lines at low resolution. Then, we combined this result with high-resolution-elements to describe and recognize archive documents.

Text-line recognition as segment extraction on low resolution images

This work has been presented in detail in [21].

We proposed a method to extract text lines in handwritten documents. Indeed, line extraction is a first interesting step in document structure recognition. Our method is based on a notion of perceptive vision: at a certain distance, text lines of documents can be seen as high-density zones of black pixels, that is to say line segments.

Therefore, we propose to detect text line using a line segment extractor on low resolution images. The chosen extractor is presented in part 3.1.

The main interest of our method is that is makes it possible to deal with difficulties met in ancient damaged documents: skew, curved lines, overlapping text lines ... . We obtained results of text line extraction on archive documents from the 18th and 19th century.

Multiresolution description for archive document recognition

The second step of our work has been to set up a cooperation between the low image analysis and the usual high resolution image description. We worked in method DMOS (see 3.2).

Our aim was to make it possible a multiresolution description of any kind of document. As an example, we worked on naturalisation decree register from the 19th century, but assuming the generality of our method.

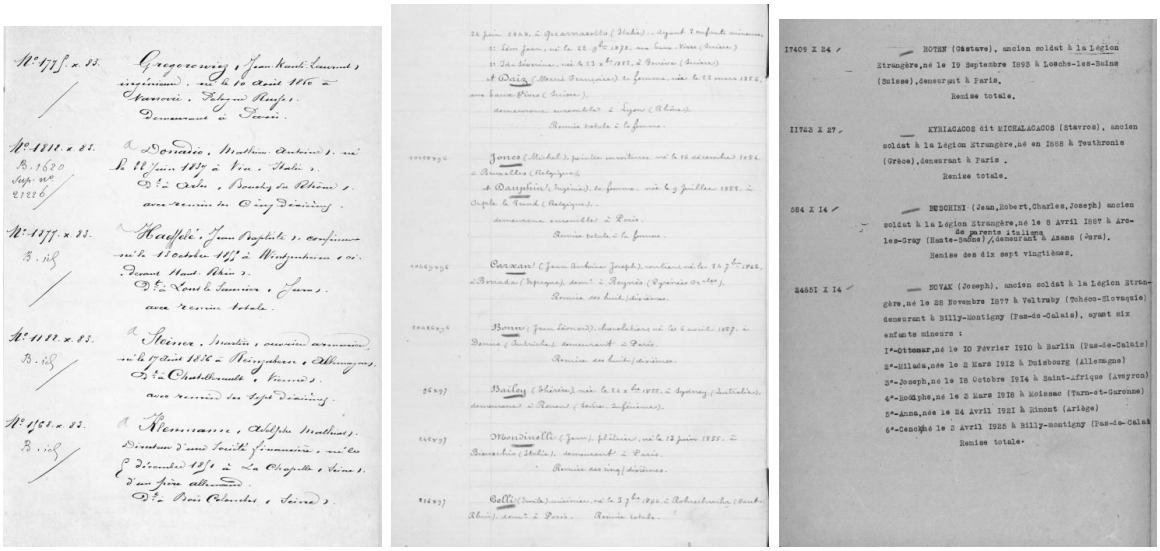
In classical EPF description, we manipulate a single image associated to its extracted connected components and segments. We set up a new operator for EPF grammar that makes it possible to deal now with a set of images and associated elements from different resolution. Consequently, the grammatical description can be based on several resolutions.

We present examples of treated documents on figure 8.

The aim is to detect register numbers in the margin and surnames in the fronting acts. We propose a cooperative description mechanism as presented below :

Step	Low resolution	High resolution
1	Text line extraction	
2	Margin computation	
3		In the margin, find a number
4	In front of a number, find a text line	
5		Detail the text line as connected components
6		Go to step 3 in order to study each number

The key point of our method is that the different resolution levels are always available during the analysis and are used not merely successively but simultaneously. Indeed, the cooperative analysis is guided by the knowledge description.



(a) (b) (c)

Figure 8: Example of treated documents

Thanks to these description, we obtain a better recognition of these kind of document in a general case, mainly because of the intuitive description of margin position.

We applied this system on 89000 images of naturalisation decree registers. We study the results on 2666 images and we manage to find the couples number-name in 98.62% cases. This results are detailed in part 6.4.

Conclusion

We set up a new mechanism in order to combine information extracted from different resolution levels. The next step will be to introduce a cooperation between the symbolic extracted information and the numerical data.

6.4 Large scale recognition of naturalization decree registers

Participants: Aurélie Lemaitre, Bertrand Coüasnon.

This work is a part of a partnership with Archives Nationales. Our goal was to extract structure on both handwritten and printed documents: naturalization decree registers.

Document base

The documents are presented figure 8. The base in composed of 85088 printed or handwritten pages dated between 1883 and 1930. The goal is to extract in such documents each act composed of a registration number in the margin and the associated paragraph in the text body.

A previous work had been realized on 1242 images from 1883 and 1884. However this first base was exclusively composed of handwritten documents and was not representative of the difficulties met with the global base.

That is why we continued this work including a notion of perceptive vision thanks to a multiresolution description of the document. More details are presented in part 6.3.

Results

Here are a few results about our validation:

- 15699 registers
- 85088 treated pages
- 0 untreated page
- 433 230 detected acts {number, name, paragraph}, which represents about 5 per page
- 106 empty pages that should not, that is to say 0.1% omission.

We set up three validation bases in order to estimate our recognition rate.

The first base, called *various* is composed of 320 pages taken in 47 years, with approximatively the same number of images for each year. It is composed of both handwritten and printed document.

The second base, called *representative*, is build with taking in the chronological order one image out of 250. Consequently, it is representative of the ratio handwritten/printed and the different problems of the whole base.

The last *handwritten* base is composed with handwritten images only, dated between 1883 and 1884.

We obtained results presented in table below.

Base	Page number	Acts number	Recognition	Miss	False recognition
Various	320	2706	98.93%	1.07%	4.99 %
Representative	347	3186	98.31%	1.69 %	4.33%
Handwritten	1999	13896	98.63%	1.37%	11.77 %

The most precise results are given by the *representative* base. In our application, the false recognition rate is not the priority as the produced results aim at making a quickly reading.

Conclusion

This work is a large validation of our multiresolution description of document structure.

An access to the result will be set up thanks to consultation platform (see part 5.6 for more information) at the CHAN, *Centre Historique des Archives Nationales*. Thanks to our extraction, the registers will be easily accessible by name or by number. The whole data will be completed by the user when necessary thanks to the annotating part of the platform.

6.5 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 input methods are used to enter text on devices which are equipped with a sensitive screen and a pen. Two main parts compose these input 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.

Our previous works have come up with the design of DIGIME 1 (DIGital Ink Micro Editor, see section 5.2) and its experimental evaluation. Its main features are an input in entire word context, a visual feedback area closed to the input area and a set of accentuation, punctuation and editing gestures. Its evaluation has lighted up 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 [BDAJ05] (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.

Some experiments are under way to test the proposed evolution of DIGIME. Moreover, another set of experiments are conducted to evaluate the user's perception of the usefulness and the usability during its first use of DIGIME. Theses experiments involve two different groups of user. The first one is composed of non-technophile whereas the second one is composed of technophile. The aim is to investigate a possible difference of immediate perception between these two user profiles.

6.6 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.

Our works focus on the explicit exploitation of the spatial context in the recognition process. We have already proposed a first method that determines the different spatial contexts and learns the associated shape recognizers from training data. During the recognition, the spatial context of the stroke is compared to the different spatial context models and the shape recognizer associated to the most relevant one is called. This method is similar to a system using expert's rules to define the contexts and to select the appropriate shape recognizer, but the expert's rules are replaced by an automatic learning and modeling process.

This hierarchical approach is based on the idea that only a subset 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 a diacritic than another character. So, the recognition system can focus on diacritic recognition. Despite the intuitive aspect of this approach, one of its major problems is the *a priori* choice of the knowledge hierarchical order.

Considering this, we propose a new approach [18] using three points of view based on three different compositions of two sources of knowledge: the spatial context and the shape. The first one uses the spatial context to focus the system on groups of symbols sharing similar contexts and based its final decision on the shape. On the contrary, the second point of view uses the shape to focus the system on groups of symbols sharing similar shapes and based its final decision on the spatial context. The third point of view uses jointly both the sources of

[BDAJ05] 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.

knowledge to take its decision. The global system relies on these three points of view that we combine to optimize the global decision. The general framework of the approach relies on the fuzzy sets theory.

The main challenge of this work is to automatically generate the points of view and combine them in order to optimize the global performance of the system. The aim is to propose a method allowing extracting and combining knowledge on the spatial context and the shape to perform contextual shape recognition without any prior information on the targeted domain. To validate the method, we apply it to on-line graphic gesture recognition in an input method for PDA. These gestures allow editing the previously inputted characters.

Since a classical Radial Basis Function Network achieve only a recognition rate of 52.90% using only shape features whereas it achieves a recognition rate of 92.73% using both context and shape features, there is no doubt on the usefulness of the spatial relative context in our recognition problem of 18 handwritten gestures. The proposed method achieves a recognition rate of 93.56% and so reduces the error rate by 11.42% compared to the RBFN using all the features.

Our future works will focus on testing this method on more complex recognition problems. The improvement of the combination of the source of knowledge will keep our attention too.

These works are parts of a PhD thesis granted by France Telecom R&D (see section 7.3).

6.7 Fuzzy relative positioning for on-line handwritten stroke analysis

Participants: François Bouteruche, Sébastien Macé, Éric Anquetil, Guy Lorette.

One of the major issues of on-line handwritten stroke interpretation is the context evaluation, which enables to associate the semantics to a stroke. As explained in section 6.6, it allows discriminating symbols having the same shape but different meaning depending on their context. Moreover, the context identification allows driving the shape recognition process [18, 23].

We have started to work on this problem of context evaluation and more particularly on the spatial context modeling. The challenge is to describe qualitatively and robustly the relative position of objects, *i.e.* how they are located one relatively to another.

So far, this problem is not very present in the literature. Existing methods are mostly empirical and *ad hoc*, which is not satisfactory, since they can not be adapted to other domains dealing with handwritten stroke interpretation. Moreover, they often give an "all-or-nothing" answer, which we believe is not suitable. To deal with the imprecise aspect of this problem, the fuzzy set theory [Zad65] is an interesting modeling framework.

Our first contribution to this problem is the adaptation of a robust and well formalized approach used in the domain of image analysis [Blo99] to the on-line handwritten stroke analysis case. This approach exploits fuzzy mathematical morphology to model relative directional relationships. Due to the non-adjacent property of the points of the on-line signal, the direct

[Zad65] L. A. Zadeh, "Fuzzy sets", *Information and Control 8*, 1965, p. 338–353.

[Blo99] I. Bloch, "Fuzzy Relative Position between Objects in Image Processing: a Morphological Approach", *IEEE Trans.on PAMI 21*, 7, 1999, p. 657–664.

application of this method produces a comb effect (see figure 9, on the left). We propose an approximation that overcomes this problem (see figure 9, on the right).

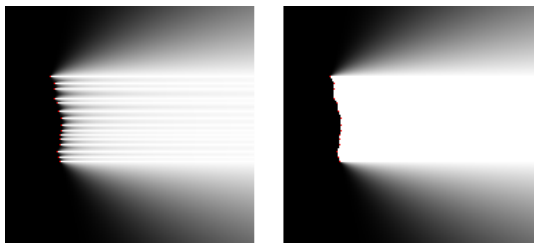


Figure 9: Fuzzy landscape computation of the "on the right" relative directional relationship: effects of the direct application of the method (on the left) and of the proposed approximation (on the right).

Our second and main contribution is an original method to automatically learn fuzzy relative position relationships, which is as far as we know completely new, even in image analysis. We learn the structuring element function that is the core of the definition of a relative positioning relationship. It allows modeling a relative position relationship for each class of symbol that have to be positioned relatively to a reference. These relationships evaluate how typical of the position of the correponding class is the position of a new sample relatively to its reference. Figure 10 presents some of the learnt fuzzy relative position relationships.

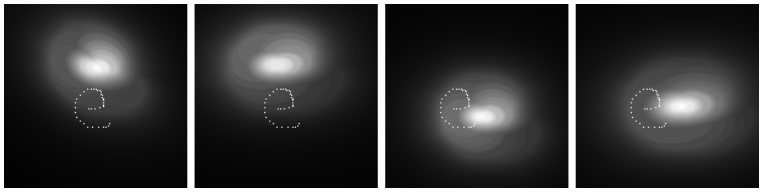


Figure 10: Fuzzy relative position relationships for the grave, dieresis, dot and carret return classes relatively to a stroke.

6.8 Handwriting character synthesis for on-line writer adaptation

Participants: Harold Mouchère, Éric Anquetil, Guy Lorette.

The Imadoc team has already designed powerful fuzzy systems for on-line handwriting recognition like RESIF (see section 5.1). We have proposed in previous works [MAR05b,MAR05a]

[MAR05b] 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.

[MAR05a] H. Mouchère, E. Anquetil, N. Ragot, "On-line Writer Adaptation for Handwriting Recog-

and a recent one [16] to improve the recognition rate by designing a technique for on-line adaptation to the writer style. In a more recent work [28] we propose to use the synthesis of on-line characters to improve the adaptation speed and quality. In this first study we compare different synthesis strategies.

To synthetise handwriting recognition we use different reshaping procedures. We have compare classical reshaping coming from the off-line character recognition field [CPCAL02] (x-, y-scale, slant) with new reshaping solutions adapted for the on-line characters from our contribution. We propose to change the speed and the curvature of the strokes as shown respectively by the figures 11 and 12.

(a) Original, (b) Increased speed, (c) Decreased speed.

Figure 11: Examples of reshaped character by changing the speed.

(a) Original, (b) Decreased curvature, (c) Increased curvature.

Figure 12: Examples of reshaped character by changing the curvature.

In our experiment we have compare this different reshaping strategies in the context of learning a mono-writer recognizer. In this way we can evaluate if the reshaping keep the writer style but generate enough variation to avoid over fitting. We learn a RBFN with just 3 characters of each classes (the minus latin letters) for 12 writers. The reshaping process

nition using Fuzzy Inference Systems", *in : Proceedings of the 8th International Conference on Document Analysis and Recognition (ICDAR)*, B. Werner (editor), *2*, IEEE Computer Society, p. 1075–1079, Seoul, Korea, August 2005.

[CPCAL02] J. Cano, J.-C. Pérez-Cortes, J. Arlandis, R. Llobet, "Training Set Expansion in Handwritten Character Recognition.", *in : Proceedings of the 9th International Workshop on Structural and Syntactic Pattern Recognition (SSPR) and 4th Statistical Pattern Recognition (SPR)*, p. 548–556, 2002.

Strategy	Mini. (%)	Average (%)	Maxi. (%)
Reference	72.3	79.4	89.2
Off-line	67.7	75.5	83.9
Speed	82.2	86.3	92.3
Curvature	72.8	80.9	86.6
All	80.0	86.9	93.7

Table 2: Recognition rates for the different reshaping strategies with only 3 original characters used per class.

is used to synthetise a learning database of 360 characters per class from these 3 characters. This recognizer is tested on 10 original characters per class from the same writer. A reference classifier is compute using 30 original characters per class. The table 2 resumes the results.

We can see that the on-line strategies bring new useful reshaping as the recognition increases from 75.5% for the off-line strategy to 86.9%. Further more with only 3 characters per class we can generate a better classifier as using simply 30 original characters per class which permit a recognition rate of 79.4%.

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

6.9 A unified strategy to deal with different natures of reject

Participants: Harold Mouchère, Éric Anquetil, Guy Lorette.

The interest of reject for classifier optimization has been shown many times. The diversity of the applications requiring this concept makes us to distinguish two main natures of reject with distinct goals: the confusion reject and the distance reject. We present in [26, 29] a unified strategy to deal with these two different reject options.

The aim of the confusion reject is to improve the accuracy of the recognizer by rejecting pattern on which the classifier can strongly make a misclassification. These errors are near the decision boundaries because class scores of two (or more) are nearly equal.

The distance reject allows to delimit the knowledge of the used classifier. In this way, it can reject patterns which do not belong to learned classes. Hence, if a shape is too far from the knowledge it must be rejected. Thus it can be used for outlier detection and rejection.

The proposed reject option uses knowledge of the target classifier through the reliability functions. So this approach do not modify the target classifier on which the reject option is applied. A reliability function ψ in $\Re$ depends on the used classifier and of the nature of the wanted reject option. It allows to determine the reliability you must have in the result of the classifier. The more a pattern must be rejected, the less is the reliability function. Contrary to [DSV00] we permit to define a set of N reliability functions $\{\psi_i\}$ which allow more precision in the reject as shown in [FRG00] and [29].

[DSV00] C. De Stefano, C. Sansone, M. Vento, "To Reject or Not to Reject: That is the Question - An Answer in Case of Neural Classifiers", *IEEE Transaction on Systems, Man and Cybernetics 30*, 1, feb 2000, p. 84–94.

[FRG00] G. Fumera, F. Roli, G. Giacinto, "Reject option with multiple thresholds", *Pattern Recogni-*

Thus the reject is defined with a set of N thresholds $\{\sigma_i\}$ each one associated to a reliability function ψ_i. Then to have a reject, all functions must be lower than their respective threshold:

$$\forall i = 1..N, \psi_i \leq \sigma_i. \quad (1)$$

The learning phase of the reject option consist on finding the suitable set of threshold $\{\sigma_i\}$. We propose an iterative algorithm called AMTL (Automatic Multi-Threshold Learning). This algorithm starts from a set of thresholds which reject all the learning data and next decreases one threshold at each step until the wanted reject rate is achieved. Different variants of AMTL are described precisely in [26, 29, 27].

In [27] we focus on the problem of outliers rejection. The aim of this paper is to study the capacity of different rejection strategies to deal with the generalization of a learned reject option. Three different cases can be distinguished depending on outliers to reject during the use (generalization phase) compared to those available during the learning phase:

- the reject option is learned with a set of classes A and then the classifier will have to reject these same classes A, it is called the *A→A problem*;
- the reject option is learned with a set of classes A and the classifier will have to reject another set of classes B, it is called the *A→B problem*, in a limit case A can be empty;
- the reject option is learned with a set of classes A and the classifier will have to reject both classes from A and B, it is called the *A→AB problem*, this is an intermediate problem between the *A→A problem* and the *A→B problem*.

Two different reject options are compared in [27] in these 3 different context problems. The first one uses a rejection class (RC) directly added in the main classifier and the second one uses our proposed solution with reliability functions (RF) and AMTL algorithm. The results show that RC solution is better in a *A→A problem* but has less generalization capacity in the *A→B problem* than our RF solution. In the *A→AB problem* the both solutions have quite similar results. But our RF solution has three main advantages:

- the original classifier is not modify and so it keeps its recognition performance whereas the RC solution increases the complexity of the classification problem,
- any operating point can be achieve easily with RF solution whereas the RC solution only permit some operating points,
- our RF solution can be learned with a empty outliers database A for the *A→B problem* which is not possible with the RC solution.

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

6.10 Statistical Language Models for On-line Handwritten Sentence Recognition

Participants: Solen Quiniou, Éric Anquetil, Guy Lorette.

---

tion 33, 12, dec 2000, p. 2099–2101.

We investigate the integration of statistical language models 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:

$$\hat{W} = \arg\max_{W} \log\left[p(S|W)\right] + \gamma \log\left[p(W)\right], \tag{2}$$

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 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 [30], we study the use of language models at different steps of the recognition process (n-gram and n-class models with classes either built automatically or predefined) as well as the combination of different language models. In fact, we consider 3 cases :

- all language models are directly used during the recognition step (performed by the Viterbi search), which becomes expensive with long range history models;
- the language models are used, in post-processing step, to rescore the list of N-best sentences produced by the recognition system. The main drawback of this approach is that the improvement brought by the use of the language models strongly depends on the sentences of the N-best list produced by the recognition system;
- the latter approach is somewhere between the two previous ones. Indeed, some language models are used during the recognition step, to produce the N-best sentenc list, while the remaining models take part in the N-best list rescoring. Thus, this approach tries to take into account the advantages of both previous approaches. The idea is to use models with small size histories to generate the N-best list of sentences and then to reorder them, using models with longer size histories.

The Brown Corpus [FK79] was used for both the construction of language models (about 47 000 sentences ie 900,000 words) and the tests. The test set includes 260 sentences (4,137 words) written by 7 writers. These sentences were then manually segmented into words in order to introduce no bias due to incorrect segmentation.

In the first approach, the language model which achieves the best tradeoff between word error rate and number of parameters is the biclass model with 1,500 statistical classes (ie built automatically using the Brown algorithm [BPdSL92]). The recognition rate reached with

[FK79] W. Francis, H. Kucera, *Brown Corpus Manual*, Brown University, 1979.

[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.

this model is 91.8% which corresponds to a 50.2% decrease of the word error rate. This first approach was shown to be better than the second one since the corresponding recognition rates were 3% above the ones obtained with the second approach. Furthermore, the language models with statistical classes perform better than the ones with predefined classes (corresponding to *Part of Speech* categories) and the combination of both types of class models doesn't lead to a significant improvement.

Finally, in the third approach the biclass model with 1,500 statistical classes is used in the recognition process to produce the list of 100 best sentences which is then reordered with different language models. The considered language models used to reorder the list are longer range language model ie their history is made of 3 or 4 words (corresponding to 4-gram, 5-gram, 4-class or 5-class language models). This only leads to slight improvements since only a small percentage of such n-grams or n-classes are estimated in the corresponding language models and are present in the test set.

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

Participants: Sébastien Macé, Éric Anquetil, Guy Lorette.

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 non-isolated contexts. The work presented in this section aims at interpreting strokes of hand-drawn structured document.

We work on a new generic method for pen-based system development. It is based on a framework constituted of three main components:

- a set of graphical functions (display, zoom, *etc.*) and pen-based editing functions (selection, deletion, *etc.*). These functions are domain-independent: they can be exploited by any pen-based system;
- a formalism to model on-line document interpretation of documents from a given domain. This component is domain-dependant because it can be adapted to various domains. Its main originality is to model the global structure of the analyzed document, which allows to filter the possible interpretations of a hand-drawn stroke depending on its positioning relatively to the other elements of the document. This global vision is coupled with more classical recognition systems, based for instance on structural interpretation or statistical and morphological interpretation. The formalism is based on the modelling of the relative positioning of the document elements. In order to take into account the imprecision of hand-drawn strokes, we use the fuzzy logic framework to evaluate the degree to which a given relative positioning is satisfied. This method is presented more in detail in [24].
- an analyzer that interprets the hand-drawn strokes and drives the recognizers by using the knowledge modelled by the formalism. This component is domain-independent.

The components of the system are well separated and can be modified and adapted independently of the others. Thanks to this method, the design of pen-based systems only requires

writing interpretation rules, designing the necessary hand-drawn shape recognizers and specifying the graphical information (for instance visual aspect of the document elements). This generic approach has been presented more in detail in [23] and in [15]. For this last paper, published in the Sixth International Conference on Computer-Aided Design of User Interfaces (CADUI'2006), we have obtained the *best paper prize*.

The presented methodology has been used to develop several pen-based editor prototypes; figure 13. The first one makes it possible to write musical scores in quite a same way as on paper (see section 5.7). 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 (see section 7.5).

We have also designed a pen-based diagram editor prototype. This system has been declined in particular for UML class diagram editing. We have also started working, in collaboration with the EVODIA (see 7.1), on the design of a pen-based system for electric diagram editing.

6.12 A Generic Method of Segmentation / Recognition Collaboration

Participants: Christophe Renaudin, Yann Ricquebourg.

The "basic" offline handwriting recognition scheme can be described in 4 phases:

1. preprocessing: binarization, squeletization, correction, etc. on the image;
2. segmentation: the symbols are isolated in order to recognize them separately;
3. features extraction: different kinds of describing values are calculated from the image of each character;
4. recognition: a classifier recognizes each character from the extracted features.

This scheme is a sequential scheme. Its main drawback is the fact that if one of the phases does not success, in particular the segmentation phase, the result of the whole process won't be correct. This difficulty is summed up in the Sayre's paradox (1973): a letter cannot be segmented before having been recognized and cannot be recognized before having been segmented.

A solution for this paradox is to replace the sequential scheme (first segmentation, then recognition) by a collaboration between the segmentation and the recognition. So, we develop a method based on the evaluation of several segmentation candidates produced by grouping elements of an over-segmentation (figure 14). This method was used to segment and recognize pairs of touching and overlapping digits and touching musical symbols.

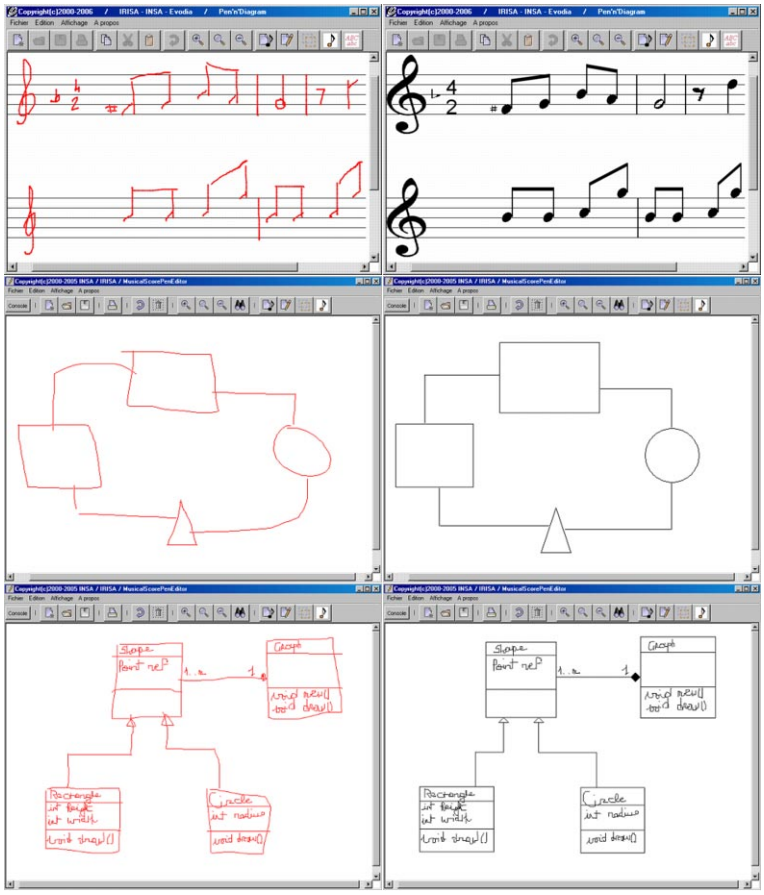


Figure 13: Screenshots of the developed pen-based prototypes; from top to bottom, musical score editor, diagram editor, UML class diagram editor; on the left the hand-drawn strokes, on the right the corresponding interpreted documents.

Over-segmentation

In order to over-segment pair of symbols, we use an algorithm of stroke decomposition in regular and singular areas [RAC04]. The regular areas are connected each other through the singular areas. So we can represent the decomposition by a undirected graph where vertices represent regular areas, and edges represent singular areas.

Candidates generation

Using this decomposition in regular and singular areas in a graph, we are able to create candidates as appropriate subgraph of the structure: they consist in different jointures of regular

[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.

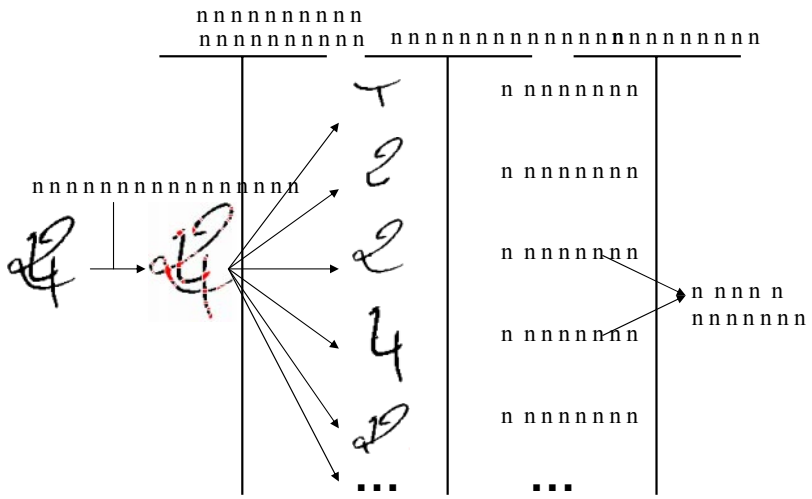


Figure 14: General scheme of the method

areas which include the singular areas associated to them. As there are potentially 2^n possible jointures (for n regular areas), we have to introduce heuristics, based on hypothesis, in order to reduce the number of subgraph candidates.

A first "strong" hypothesis is the connectivity of the candidate: none regular area of the candidate is isolated from the others. We produce only the connected candidates using the algorithm of the connected subgraphs tree [Cui04].

Candidates classification

In order to determine which candidates can probably correspond to the two symbols of the initial image, we try to recognize each produced candidate. We extract features (elliptic Fourier descriptors or Zernike moments) from the image of the candidate and give them to a classifier based on a RBF neural network. This classifier associates a score to each possible class. This score can be interpreted as a confidence index on the belonging of the symbol to the corresponding class.

Candidates selection

We select the pair of candidates which can better correspond to the initial symbols using 2 criteria: a criterion of complementarity and another criterion based on the score given by the classifier.

[Cui04] B. Cuissart, *Plus grande structure commune à deux graphes : méthode de calcul et intérêt dans un contexte SAR*, PdD Thesis, Université de Caen/Basse-Normandie, December 2004.

For the complementarity criterion, we check if the two candidates of the pair share a minimum of pixels, and if only a minimum of pixels from the initial image are not included in one of theses candidates.

For all pairs obtained using the complementarity criterion, we calculate a score multiplying the score of the 2 candidates. The pair with the higher score is considerer as the pair of candidates which more probably correspond to the initial symbols.

Optimization

Because of the high number of candidates (even with the hypothesis of connectivity), we have to use other hypothesis in order to reduce it:

- minimal area: the symbol candidates must have a surface upper than a minimal value,
- strict complementarity: no regular areas are shared by the two candidates, and no regular areas are unused by them. We consider that the two symbols share only regular areas,

Test results

100 pairs of touching digit were produced by a random intersection of two digit images extracted from the IRONOFF database [VGLKB99]. 59.3% of these examples were correctly segmented (sometimes with a false recognition of one candidate by the classifier).

34 pairs of touching musical symbols were extracted of printed music parts, considering only 5 symbol classes (5 kinds of musical accidentals). 91% of these examples were correctly segmented (sometimes with a false recognition of one candidate).

Conclusion

After validating these encouraging results on large test database recently, the future work will focus on using an *a priori* knowledge in order to reduce the number of candidates in a generic way, with a more complete collaboration between the segmentation phase and the collaboration phase.

7 Contracts and Grants with Industry

7.1 EVODIA

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

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

[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.

- 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. In 2006, Evodia has been also prizewinner of the 8th national competition in the "creation" category. 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.

The Evodia company uses three licences:

- DocRead, see section 5.3,
- the image processing library, see section 5.4,
- RESIF, see section 5.1.

Contact: `yvan.ride@evodia.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 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 6.5): 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 5.7).

7.6 Access to Handwritten Archives Documents

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

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

A long-term collaboration with the *Conseil Général des Yvelines* started in 2002 with this research project which 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 [11] 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. Within the research project a validation has been done on 105,000 pages of military forms. They have an automatic access by handwritten last names and 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.

After an industrial transfer to the Evodia company 430,000 pages of military forms have now an automatic access by handwritten names and the platform for annotations has been deployed on other kind of documents.

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 has been validated on 85,000 documents (section 6.4).

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).

8.2 International initiatives

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

8.3 Visiting scientist

Bidyut B. Chaudhuri, professor at ISI (Indian Statistical Institute in Kolkata) spent one month in our team, in the context of cooperation on on-line handwriting recognition and 10 th IWFHR.

9 Dissemination

9.1 Leadership within scientific community

9.1.1 Chairman

- G.Lorette is general co-chairman of the 10^{th} IWFHR 2006, (International Workshop on Frontiers in Handwriting Recognition), Oct. 2006, `http://www.irisa.fr/iwfhr10`
- G.Lorette is general co-chairman of the 18^{th} ICPR 2006 (International Conference on Pattern Recognition), Aug. 2006, `http://www.comp.hkbu.edu.hk/~icpr06/`

9.1.2 Program Committee

- É. Anquetil, B. Coüasnon and G. Lorette are members of the program committee of CIFED 2006 (Colloque International Francophone sur l'Écrit et le Document), Sep. 2006.
- B. Coüasnon is a member of the program committee of:
  - ICDAR 2007 (International Conference on Document Analysis and Recognition), Sep. 2007,
  - GREC 2007 (International Workshop on Graphics Recognition), Sep. 2007
  - AXMEDIS 2006 (International Conference on Automated Production of Cross Media Content for Multi-channel Distribution), Dec. 2006.

9.1.3 Organizing Comittee

- É. Anquetil and G. Lorette are in the organizing comittee of: IWFHR'10 (International Workshop on Frontiers in Handwriting Recognition), Oct. 2006, `http://www.irisa.fr/iwfhr10`.

9.1.4 Reviewing

- G. Lorette is a reviewer of IEEE ISCIP-06 (International Conference on Signal and Image processing, Dec. 2006, `http://www.icsip.org`.

9.1.5 Member of scientific society

- G. Lorette is a member of the scientific council of PSI (Université-INSA de Rouen) and of the E3I computer laboratory (E3I-Université de Tours).
- É. Anquetil, B. Coüasnon, and G. Lorette take part in the animation structure " SA 5.2-Écrit " of the Pôle 2 : ICC (Interaction coopération et communication), Axe 5: " Communication orale, écrite et visuelle " du GDR-PRC CNRS I^3 (Information, Interaction, Intelligence).

- J. Camillerapp, G. Lorette, É. Anquetil, and B. Coüasnon take part in the activities of the society GRCE : " Groupe de Recherche en Communication Écrite ".

9.2 University education

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

- H. 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.
- G. 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, awards

- B. Coüasnon was invited for a presentation to the seminar INRIA IST 2006 : "Pérenniser le document numérique", Amboise, France, Oct. 2006 [12].
- S. Macé and É. Anquetil received the best paper prize at the Sixth International Conference on "Computer Aided Design of User Interfaces", CADUI'2006 held in Bucarest, Romania, June 2006 for the paper "A Generic Approach for Pen-based User Interface Development" [15]

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.

Articles in referred journals and book chapters

[10] E. Anquetil, G. Lorette, "New Advances and New Challenges in On-line Handwriting Recognition & Electronic Ink Management", *in : Digital Document Processing: Major Directions and Recent Advances(Advances in Pattern Recognition)*, B. Chaudhuri (editor), Springer-Verlag, November 2006, to be published.

[11] B. Coüasnon, J. Camillerapp, I. Leplumey, "Access by Content to Handwritten Archive Documents: Generic Document Recognition Method and Platform for Annotations", *International Journal on Document Analysis and Recognition, IJDAR*, 2006, to appear.

[12] B. Coüasnon, "Archivage numérique des documents anciens : Nouveaux systèmes d'accès aux documents manuscrits numérisés", *in : Pérenniser le document numérique, Séminaire INRIA*, J. M. L. Calderan, B. Hidoine (editor), ADBS éditions, Oct 2006, p. 147–170.

[13] 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 8*, 2-3, June 2006, p. 111–122.

[14] S. Macé, E. Anquetil, B. Bossis, "Pen-Based Interaction for Intuitive Music Composition and Editing", *in : Intelligent Music Information Systems: Tools and Methodologies*, J. Shen, J. Shepherd, B. Cui, and L. Liu (editors), Idea Group, 2007, to be published.

[15] S. Macé, E. Anquetil, "A Generic Approach for Pen-based User Interface Development", *in : Computer-Aided Design of User Interfaces V, Proceedings of the 6th International Conference on Computer-Aided Design of User Interfaces (CADUI'06)*, G. Calvary, C. Pribeanu, G. Santucci, and J. Vanderdonckt (editors), Springer-Verlag, 2006, ch. 5, p. 57–70, to be published.

[16] H. Mouchère, E. Anquetil, N. Ragot, "Writer Style Adaptation in On-line Handwriting Recognizers by a Fuzzy Mechanism Approach : The ADAPT Method", *International Journal of Pattern Recognition and Artificial Intelligence (IJPRAI), to be published*, 2006.

[17] T. Paquet, G. Lorette, "Reconnaissance de l'écriture manuscrite", *in : Les documents écrits : De la numérisation à l'indexation par le contenu (Traité IC2, série Informatique et Systèmes d'Information)*, R. Mullot (editor), Hermès Science Lavoisier, 2006, to be published.

Publications in Conferences and Workshops

[18] F. Bouteruche, E. Anquetil, "Fuzzy Point of View Combination for Contextual Shape Recognition: Application to On-line Graphic Gesture Recognition", *in : Proceedings of the 18th International Conference on Pattern Recognition (ICPR'06)*, p. 1088–1091, 2006.

[19] F. Bouteruche, S. Macé, E. Anquetil, "Fuzzy Relative Positioning for On-Line Handwritten Stroke Analysis", *in : Proceedings of the 10th International Workshop on Frontier in Handwriting Recognition (IWFHR'06)*, p. 391–396, La Baule, France, October 2006.

[20] B. Coüasnon, "What can we learn from the processing of 165,000 forms from the 19th century?", *in : International Workshop on Document Image Analysis for Libraries (DIAL'06)*, p. 172–177, April 2006.

[21] A. Lemaitre, J. Camillerapp, "Text Line Extraction in Handwritten Document with Kalman Filter Applied on Low Resolution Image", *in : Document Image Analysis for Libraries (DIAL'06)*, p. 38–45, April 2006.

[22] A. Lemaitre, B. Coüasnon, I. Leplumey, "Using a neighbourhood graph based on Voronoï tessellation with DMOS, a generic method for structured document recognition", *in : Graphics Recognition, Sixth IAPR International Workshop GREC 2005, Revised Selected Papers*, *LNCS 3926*, p. 267–278, 2006.

[23] S. Macé, E. Anquetil, "A Generic Approach for Eager Interpretation of On-line Handwritten Structured Documents", *in : Proceedings of the International Conference on Pattern Recognition (ICPR'06)*, p. 1106–1109, 2006.

[24] S. Macé, E. Anquetil, "Reconnaissance à la volée de documents structurés manuscrits en-ligne", *in : Actes du Colloque International Francophone sur l'Ecrit et le Document (CIFED'06), to be published*, p. 175–180, 2006.

[25] I. Martinat, B. Coüasnon, "A Minimal and Sufficient Way of Introducing External Knowledge for Table Recognition in Archival Documents", *in : Sixth IAPR International Workshop on Graphics Recognition,GREC 2005, Revised Selected Papers*, *LNCS 3926*, p. 206–217, 2006.

[26] H. Mouchère, E. Anquetil, N. Ragot, "Etude et gestion des types de rejet pour l'optimisation de classifieurs", *in : 15eme congrès francophone Reconnaissance des Formes et Intelligence Artificielle (RFIA 2006)*, Presses Universitaire François-Rabelais, Tours, France, Janvier 2006.

[27] H. Mouchère, E. Anquetil, "Generalization Capacity of Handwritten Outlier Symbols Rejection with Neural Network", *in : Proceedings of the 10th International Workshop on Frontier in Handwriting Recognition (IWFHR'06)*, p. 187–192, La Baule, France, October 2006.

[28] H. Mouchère, E. Anquetil, "Synthèse de caractères manuscrits en-ligne pour la reconnaissance de l'écriture", *in : Actes du Colloque International Francophone sur l'Ecrit et le Document (CIFED'06)*, p. 187–192, 2006.

[29] H. Mouchère, E. Anquetil, "A Unified Strategy to Deal with Different Natures of Reject", *in : Proceedings of the International Conference on Pattern Recognition (ICPR'06)*, p. 792–795, 2006.

[30] S. Quiniou, E. Anquetil, "A Priori and A Posteriori Integration and Combination of Language Models in an On-line Handwritten Sentence Recognition System", *in : Proceedings of the 10th International Workshop on Frontiers in Handwriting Recognition (IWFHR'06)*, p. 403–408, La Baule, France, October 2006.

[31] L. Rousseau, E. Anquetil, J. Camillerapp, "What knowledge about handwritten letters can be used to recover their drawing order?", *in : Proceedings of the 10th International Workshop on Frontiers in Handwriting recognition (IWFHR'06)*, p. 355–360, La Baule, France, October 2006.