Project-Team Imadoc

Recognition and interpretation of images and documents

Rennes

Activity Report
2007

Contents

1 Team

Team Leader

Jean Camillerapp [Professor Insa]

Administrative Assistants

Stéphanie Lemaile [TR Inria]

Staff members (Insa)

Eric Anquetil [Associate Professor]

Bertrand Coüasnon [Associate Professor]

Ivan Leplumey [Associate Professor]

Yann Ricquebourg [Associate Professor]

Harold Mouchère [Lecturer from 01 Oct 2007]

Staff members (Université of Rennes 1)

Guy Lorette [Professor]

Charles Quéguiner [Associate Professor]

Hélène Richy [Associate Professor]

Laetitia Rousseau [Lecturer to 31 Aug 2007]

Solen Quiniou [Lecturer from 01 Oct 2007]

Ph.D. Students

François Bouteruche [Insa research grant to 30 Sep 2007]

Isaac Martinat [Insa research grant to 30 Sep 2007]

Harold Mouchère [CNRS/Région research grant to 30 Sep 2007]

Solen Quiniou [Menrt research grant to 30 Sep 2007]

Aurélie Lemaitre [Menrt research grant]

Sébastien Macé [Région/Insa research grant]

Christophe Renaudin [Menrt research grant]

Abdullah Almaksour [Insa research grant from 01 Oct 2007]

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 concerns 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 is 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 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 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

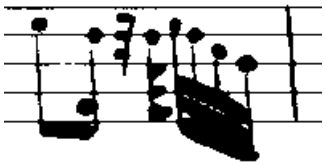
3.1 Early vision

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



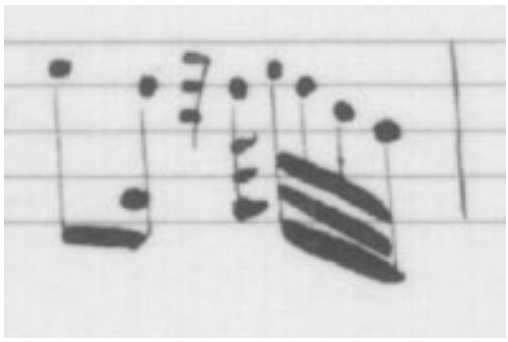
(a) Threshold of 210 : allowing the detection of thin lines

(b) Threshold of 195 : the vertical elements remain

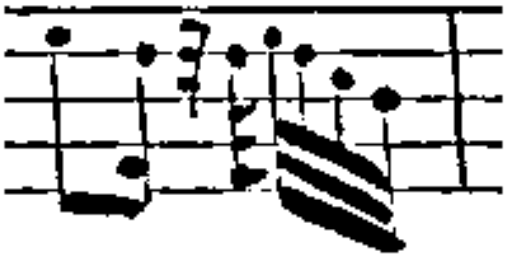
(c) Threshold of 110 : the bars of the notes are separated

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.

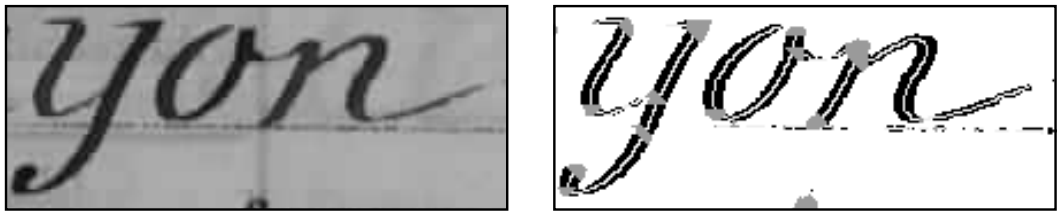


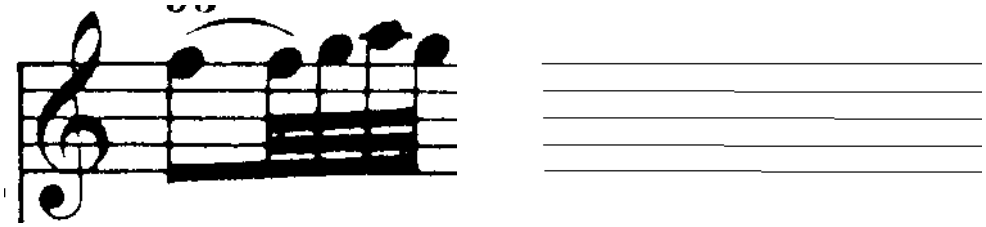
Figure 3: Processing of an image from an old document. The skeleton appears in white in the medium of the line and singular areas are in dark gray.

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 [6].

[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 (figure 4) with the predicted run-length when linear structures cross each other or overlap with other shapes.

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 to determine threshold and thus to obtain a good choice.

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).

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: Eric 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, Eric 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;
- JournRead: a prototype for newspaper structure recognition;
- 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);
- NatuRead: a software for recognition of naturalization decree registers from 1883 to 1930. This software has been applied on 85,088 pages of the *Centre Historique des Archives Nationales*;
- LettRead: a software for extracting structure of mail documents. It has been applied on 1150 images provided by the French project RIMES. Results are presented in part 6.4;
- BanglaRead: a software for extracting headline in Bangla script. This extraction represent a pre-processing tool for handwritting recognition. This work has been realized in collaboration with University of Kolkata and applied on 1922 words from 26 writers.

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.

This library 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], Eric Anquetil.

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

The FDT library is a classification software for extracting and exploiting fuzzy decision trees. Fuzzy sets are automatically extracted by fuzzy clustering algorithms in one dimension or more. Depending on the clustering algorithm used, they can either describe the classification problem based on a discrimination point of view or on an intrinsic point of view. The resulting classifiers can be more accurate than other traditional classifiers such as radial basis function 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 define 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.

This platform for archives document retrieval can deal with textual and geometric annotations at the same level. Moreover, it 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.

[4]Today associate professor at University of Tours

[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.

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.

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, Eric 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

[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.

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: our functionalities and ergonomic choices are validated by professional musicians.

Figure 6 presents a screenshot of the system.

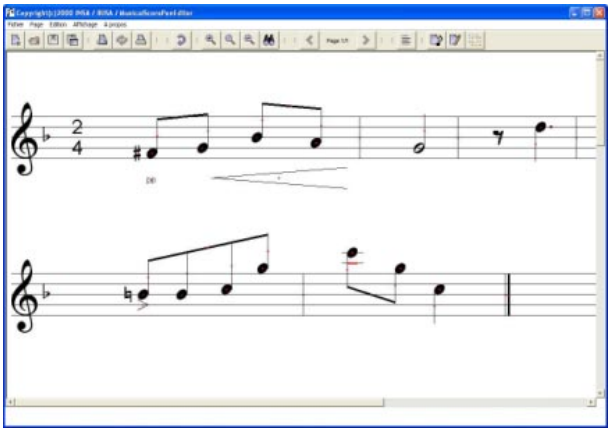


Figure 6: Screenshot of the pen-based musical score editor.

5.8 Script&Go Electricité

Participant: Sébastien Macé, Eric Anquetil.

Keywords: Handwriting recognition, electric diagram, pen-based interface.

The IMADOC research team has developed *DALI*, a framework for the design of pen-based software for structured document composition (see section 6.9). A part of this framework has been transferred to the Evodia society and resulted in the *Script&Go Electricité* software.

Script&Go Electricité in a pen-based software for the hand-drawn composition and editing of electric diagrams. It enables a user to compose electric diagrams by drawing their components on the sensitive screen of a tablet PC (figure 7 presents a user composing such a document); the software interprets the hand-drawn strokes progressively. The user thus gets an immediate feedback. We have shown in [22, 23] that *Script&Go Electricité* offers a very significant gain of time on the composition of the diagrams.

Script&Go Electricité allows composing a very large panel of electric components, such as inputs/outputs, connexions, switches, circuit breakers, disconnecting switches, earths, generators, transformers, fuses, etc. Composition and editing are designed to be as intuitive and as fast as possible. The user can select elements by surrounding them and then move them or delete them with the pen. The software also offers annotation functionalities thanks to dedicated tracings.

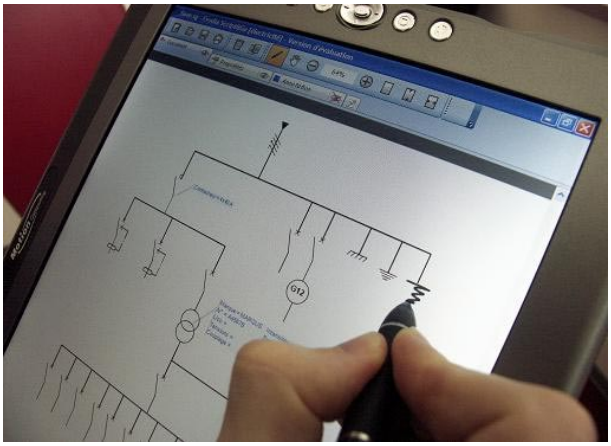


Figure 7: *Script&Go Electricité* enables a user to compose electric diagrams by drawing their elements on the screen of a tablet PC.

6 New Results

6.1 Recovery of the drawing order

Participants: Laëtitia Rousseau, Jean Camillerapp, Eric Anquetil.

A way to do off-line handwriting recognition is to recover the drawing order in 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 [30] to recover the drawing order consists in using handwriting knowledge. It concerns lowercase cursive words.
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

Search for starting and ending points
Several candidates are selected as starting and ending points. First, the graph is splitted up. Secondly, several localizations of starting and ending points 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 an algorithm presented by Kato and Yasuhara in [KY00] and

[KY00] Y. Kato, M. Yasuhara, "Recovery of drawing order from single-stroke handwritting images", *IEEE Trans.on PAMI 22*, sept 2000, p. 938–949.

we have extended [RAC04].

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 : 29.1%).
2 - Loops have to be drawn in the good direction (paths eliminated : 20%).

Then 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 path 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 words including both on-line and off-line signals. The on-line recognition system used to recognize these words is RESIFMots (cf. section 5.1). Each writer has written 250 English words. The dictionary contains 15,000 words. 8 writers have been used for learning and 13 for tests.

Table 1 contains recognition rates for the real on-line signals (On) and for the on-line signals constructed from the image (Off).

Signal	On	Off
TOP 1	80.9%	64.8%
TOP 10	94.8%	84.1%

Table 1: On-line and Off-line recognition rates

Future works will focus on the 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

[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.

archival documents rulings can be broken, curved, and ink bleeds through paper, so flip side rulings are visible.

To recognize these difficult documents a system needs an external knowledge. Thus we propose a recognition system [24] using a language to describe tables. This language is composed of two parts. The first one is a logical part where the user describes the row and column hierarchy. The second one is a physical part where the user describes the row and column separators, and optionnally he can also define the number and/or size of columns and/or rows. The advantage of this language is to describe tables with different levels of precision. On the one hand, the description can be very general. In this case, documents with important differences can be recognized with the same description but documents can not contain noise. For example, for a general description, a multi-row hierarchy can be described without specifying the number of rows for each level. On the other hand, the description given by the user can be very precise. In this case, very damaged documents can be recognized but for the same description, variations between documents can not be important. For example, for a precise description, the numbers of rows and columns can be specified. For a more precise description, sizes can also be given for some columns and some rows. The user can change easily and quickly from a general description to a precise one in adding or modifying some specifications with different precisions. He also can specify a general and precise description, for example the description can be precise for the columns where the number is fixed, and general for the rows where the number is unfixed.

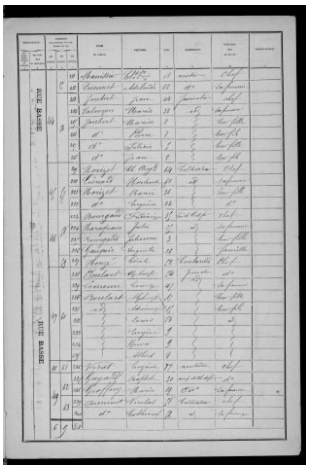
From the images, the system extracts a set of rulings and from this set and it has to recognize a table in function of the user description. This set is not totally representative from the specification, some rulings are not a part of the structure to be detected like flip side rulings and rulings from the handwritting. Some rulings can be missing or broken. To link the user specification and the set of extracted rulings we use some specific intersections we defined as final intersections, they are the intersections with one or two extremities.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. Intersections are detected with a tolerance which is automatically increased for the recognition of noisy documents when user description is precise. Using final intersections and our description language we show in [24] that our system can recognize very different table structures with a general description and very damaged documents with a precise description.

In [24] we shown on some images that our system can recognize census tables from different years with different structures using a same general description (figure 8). On this figure the 1881 table contains 8 columns whereas the 1911 table contains 10 columns and the row number for each level of hierarchy is also different between these two documents. We shown only the first level of hierarchy but the system recognize the different levels of row hierarchy.

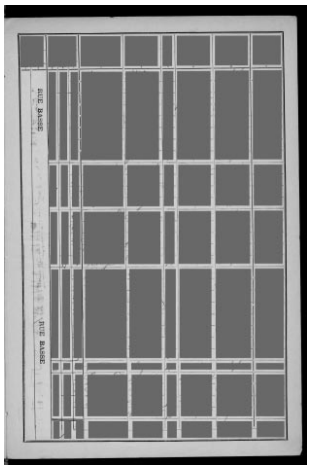
This work has been done in cooperation with the *Archives départementales des Yvelines* in France, with the support of the *Conseil Général des Yvelines* (section 7.6).

6.3 Cooperation between multiresolution images for document recognition

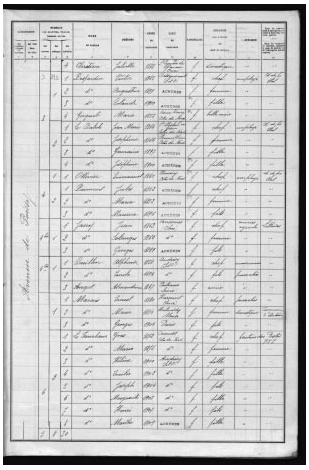
Participants: Aurélie Lemaitre, Jean Camillerapp, Bertrand Coüasnon.



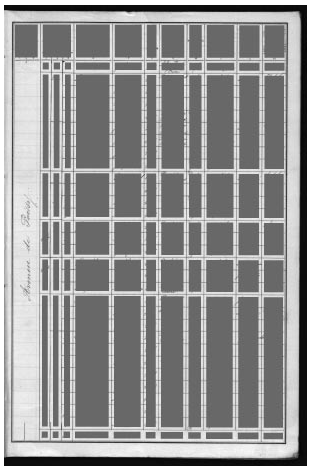
(a) 1881 : 8 columns



(b) level 1 of row hierarchy



(c) 1911 : 10 columns



(d) level 1 of row hierarchy

Figure 8: Census Tables of 1881 and 1911 and the recognized structures with a same general description, column number is unfixed like row number.

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 informations extracted at low and high resolution. This is possible thanks to a specific operator that has been introduced in DMOS method (see

section 3.2): operator `FOCUSING` that makes it possible to change the level of resolution of the studied image during the analysis, being guided by the symbolic knowledge. This mechanism is described in [19].

We set up a new cooperation between digital level (features) and symbolic level (knowledge): wedging. This mechanism is decomposed as follow:

1. Detect a line segment at low resolution.
2. Focus attention at high resolution.
3. Study digital level (black pixels) to wedge the line on existing points.

This mechanism of wedging makes it possible to find accurately the position of a line, that is not visible at high resolution, but found at low resolution. This mechanism has been used to improve results obtained for headline detection in Bangla script [21].

6.4 Application of perceptive vision for RIMES contest

Participants: Aurélie Lemaitre, Jean Camillerapp, Bertrand Coüasnon.

We have applied our implementation of the perceptive vision(see part 6.3) to French mail document recognition. We introduced our work in the context of the national French project RIMES. RIMES stands for "Reconnaissance et Indexation de données Manuscrites et de fac similES", which means "handwritten data and fax recognition and indexing". This project is financed by the French Ministries of Research and Defence. Its goal is to build a large handwritten document database, to define criteria and metrics for the evaluation campaign using those data, and to drive this campaign. Nine French research teams took part in this project as participants on different tasks linked with document recognition and retrieving.

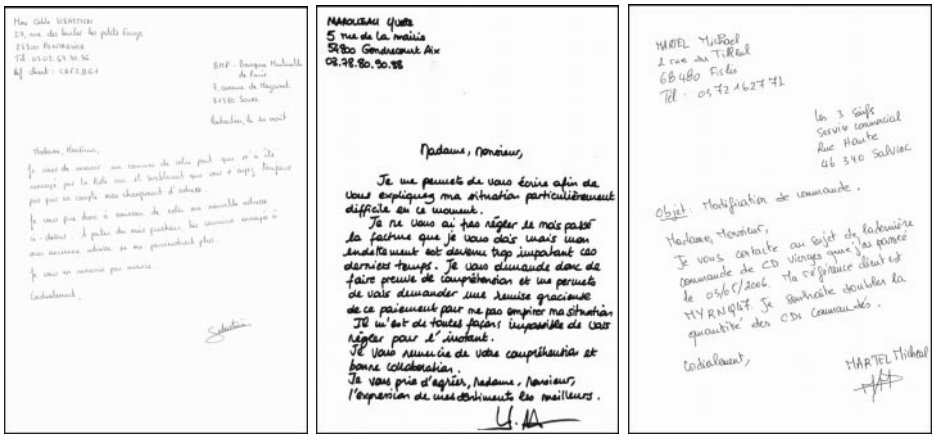
Our work aims at extracting the logical structure of French handwritten mail documents, and detecting elements like addressee details, sender details, date, subject, text body, signature, opening. It consists in two joined tasks: the segmentation of documents into blocks and the labeling of such blocks.We are faced with very variable documents: some examples are given in the figure 9.

Using DMOS method to describe this kind of document makes it possible to try successive page segmentations until labeling succeeds for each block. Thus, we can deal with unconstrained low structured handwritten mail documents.

In the metric proposed by RIMES project, we count each pixel of the image that has been wrongly labeled, balanced by its grey-level. Indeed, a white uncorrect pixel has a small impact, whereas a dark uncorrect pixel is a big mistake.

We applied our method on the 100 handwritten mail documents of RIMES contest. The score given by RIMES metric is 9.27% of error, which correspond of the percentage of pixels wrongly labeled. We then applied our system on a larger base (1150 images). The whole method is presented in [20]. Results are given in table 2.

We obtain a global recognition rate (recall) of 91.7% for the base of 1150 images, which correspond of the percentage of pixels wrongly labeled, out of the 254,207,489 black pixels we have to label. The global precision is 92.6%. The third column gives the importance, in



(a) (b) (c)

Figure 9: Example of handwritten mail documents

Class	Involved pixels	Whole base 1150 images	
		Recall	Precision
Text body	61.5%	97.2%	96.9%
Sender details	15.1%	92.0%	92.1%
Addressee details	9.0%	84.1%	83.9%
Signature	4.2%	88.9%	90.7%
Subject	4.0%	66.4%	72.1%
Date, Place	3.2%	54.0%	78.4%
Opening	2.9%	81.1%	76.0%
Global	-	91.7%	92.6%

Table 2: Results on 1150 handwritten mail documents

number of pixels, of the classes. "Text body" represents the main part, 61.5% of pixels. We can see that recognition rate is globally proportional to the importance of the class. "Subject" recognition is difficult because it can be easily confused with several other classes: "Date, Place"', "Sender details" or "Opening". Improving these three classes would automatically improve "Subject" recognition rate.

6.5 Using the relative spatial context of stroke in 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. The previous approach we proposed [BA06] to this problem uses 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 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.

Our new works consider the joint use of spatial context and shape as a full-fledged source of knowlegde whereas our former works consider it as a point of view. As a consequence, our hierarchical points of view are now composed of three levels. The global system relies on up to six points of view since there is six possible orders to use the sources of knowledge. There is no joint point of view anymore.

The main challenge of this work is to automatically generate the six hierarchical 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. There are 18 handwritten gestures to recognize.

The proposed method achieves a 95,51% recognition rate whereas a RBFN achieves a 92.72% recognition rate. We succeed to reduce the error rate by 38% compared to the RBFN using all the features.

[BA06] 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.

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 7.4).

6.6 Handwriting character synthesis for on-line writer adaptation

Participants: Harold Mouchère, Eric 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 [MAR05a,MAR05b] [16] to improve the recognition rate by designing a technique for on-line adaptation to the writer style. In a more recent work [MA06] [25, 18] we propose to use the synthesis of on-line characters to improve the adaptation speed and quality. In these studies we compare different synthesis strategies.

To synthetise handwriting recognition we use different reshaping procedures. We have compare classical *"Image Distortions"* 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 the *"On-line Distortions"* . We propose to change the speed and the curvature of the strokes as shown by the figure 11. In [25, 18] we study the Analogical Proportion in collaboration with Sabri Bayoudh and Laurent Miclet from the team Cordial to generate new characters.

Character Generation by Analogical Proportion

Analogy is a way of reasoning which has been studied throughout the history of philosophy and has been widely used in Artificial Intelligence and Linguistics [Lep98,Lep03]. We are interested here in a special case of analogy: the Analogical Proportion (AP) between four objects a, b, c and d in the same universe. Having four objects in AP is usually expressed as follows:

[MAR05a] H. Mouchère, E. Anquetil, N. Ragot, "On-line Writer Adaptation for Handwriting Recognition using Fuzzy Inference Systems", *in : Proceedings of the 8th International Conference on Document Analysis and Recognition (ICDAR)*, B. Werner (editor), *2*, IEEE Computer Society, p. 1075–1079, Seoul, Korea, August 2005.

[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.

[MA06] 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.

[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.

[Lep98] Y. Lepage, "Solving Analogies on Words: an Algorithm", *in : Proc. of COLING-ACL'98*, *1*, p. 728–735, 1998.

[Lep03] Y. Lepage, *De l'analogie rendant compte de la commutation en linguistique*, PdD Thesis, HdR Joseph Fournier Grenoble, May 2003.

"a *is to* b *as* c *is to* d". Depending on what are the objects, AP can have very different interpretations. When one of the objects is unknown, as in "wolf *is to* leaf *as* wolves *is to* x", finding a satisfying x is called *solving an analogical equation*. x would be leaves if we consider either a semantic interpretation of analogy as well as if the interpretation is on the combinatorics of the sequences of letters.

We basically represent characters by their Freeman code sequences. Let $\Sigma = \{1, 2, .., 16\}$ be the Freeman code alphabet, with an AP composed of equations like: $3 : 5 :: 12 : 14$. To solve an analogical equation on sequences, we align four sequences of different length, the empty character $\backsim$ being added to Σ. A Dynamic Programing (DP) or the A^* algorithm is used to find the better alignement.

Sequence Segmentation

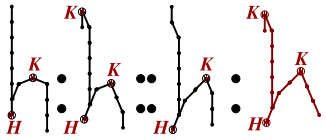
Aligning sequences composed only of Freeman code could lead to match strokes of different types and therefore to a non pertinent generation of the fourth stroke. In order to generate a coherent sequence we choose to use anchorage points représented in Σ by capital letters. Hence, resolution by AP will first resolve analogies on anchorage points then resolve analogies on sequences between anchorage points as shown in the Figure 10. The anchorage points are chosen in an appropriate way to conduct a logical modeling of letters with respect to the most stable strokes of each letter class [AL97].

Example

In the example of Figure 10, we have chosen the character "h" and used three handwritten examples h_1, h_2 and h_3. We generate by the algorithm DP a sequence x_0, representing a new character "h" with the least analogical dissimilarity.

$$
\begin{aligned}
h_1 &= 9 \backsim 9 \backsim 99999 \backsim H\ 1\ 2 \backsim 4 K\ 6\ 9\ 9\ 9 \\
h_2 &= 1\ K \backsim 8\ 9\ 9\ 9\ 9\ 9\ 10 H \backsim 2\ 2\ 4 K \backsim 8\ 8\ 9 \\
h_3 &= \backsim \backsim\ 9\ 8\ 9\ 9\ 9\ 9\ 9\ 10 H\ 2\ 2\ 3\ 3 K\ 8\ 9\ 9 \backsim \\
x &= 1\ K \backsim 8\ 9\ 9\ 9\ 9\ 9\ 10 H\ 2\ 2\ 3\ 3 K\ 8\ 8\ 8 \backsim
\end{aligned}
$$

(a)



(b)

Figure 10: (a) Resolution on Freeman direction sequences by AP. (b) the corresponding characters representation.

Figure 11 shows examples of synthetics handwriting characters generated by Image Distortions, On-line Distortions and Analogy.

Experiments

Twelve different writers have made 40 times the 26 lowercase letters (1040 characters) inputted on a PDA. Each writer database is randomly split in four parts with 10 characters per class.

[AL97] E. Anquetil, G. Lorette, "Perceptual Model of Handwriting Drawing Application to the Handwriting Segmentation Problem", *in : Proc. of the 4th Int. Conf. on Document Analysis and Recognition*, p. 112–117, 1997.

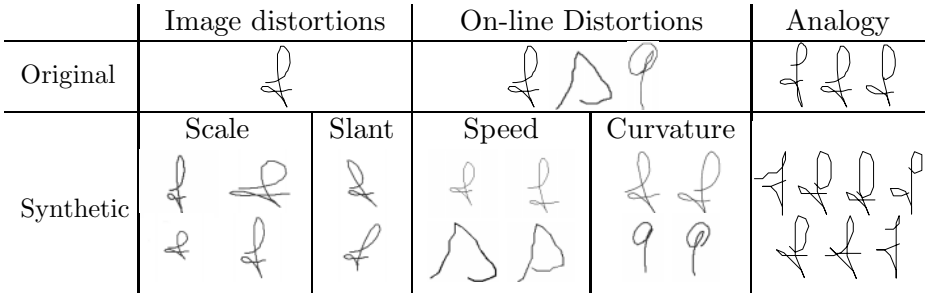


Figure 11: Examples of synthetic characters generated by the three approaches.

We use them in a 4-fold stratified cross validation: one fourth for $D10$ database (260 data) and three fourth for $D30$ database (780 data). The experimentations are made of two phases in which simple writer-dependent systems based on Fuzzy Inference Systems [16] are learned.

Firstly two writer-dependent classifiers are learned for each writer on his $D10$ database and evaluated on his $D30$ database, and inversely. The two mean recognition rates are the Reference Rates $RR10$ and $RR30$, i.e. the recognition rate achievable without character synthesis.

Secondly the handwriting generation strategies are tested. For a given writer, two, three, four, six, eight or ten characters per class are randomly chosen in his $D10$ database. Then 300 synthetic characters are generated per class to make a synthetic learning database. A writer-dependent classifier is learned with this base and tested on the $D30$ database of the writer.

We study three different strategies for the generation of synthetic learning databases: the strategy *"Image Distortions"*, the strategy *"On-line & Image Distortions"* and the *"Analogy and Distortions"* strategy Table 3 sums up the experiment results.

Nb. of used characters	2	3	4	6	8	10
Image Distortions	76.1 ± 3.3	82.5 ±2.4	85.8 ±2.0	89.4 ±1.7	91.5 ±1.6	92.7 ±1.3
On-line & Image Distortions	84.4 ±2.6	88.0 ±2.1	90.3 ±1.7	92.3 ±1.6	93.4 ±1.2	94.2 ±1.0
Analogy and Distortions	84.9 ±2.6	89.3 ±2.1	91.3 ±1.4	93.5 ±1.1	94.5 ±1.0	95.2 ±0.9

Table 3: Writer Dependent Mean (WDM) performance rate (%) and Writer Dependent Deviation (WDD) (%) for different synthetic handwriting generation strategies, the reference rate $RR10$ is 82.3 % $RR30$ is 94.5 %.

Two main conclusions can be deduced from these results. Firstly with only two original characters the "Analogy and Distortions" allows better writer-dependent recognition rate than the reference $RR10$ and eight to be better than $RR30$. Secondly Table 3 shows that the three generation approaches (image distortions, on-line distortions and analogy) are complementary. Indeed the three strategies generate 300 data and we note that the richer are the distortions the better are WDM rates.

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

6.7 Using confusion networks for on-line handwritten sentence recognition

Participants: Solen Quiniou, Eric Anquetil, Guy Lorette.

Most handwritten sentence recognition systems use word graphs to represent alternative sentence hypotheses. The standard Maximum A Posteriori (MAP) decoding is then performed to find the sentence with the highest probability given information from the recognition system and a language model. Whereas the MAP approach aims at minimizing the sentence error rate, the most commonly used metric to evaluate recognition system performances is word error rate: there is thus a mismatch between the approach used for the recognition and the objective being the maximization of correctly recognized words. A new approach, which aims at minimizing the word error rate, was introduced for speech recognition: it relies on *confusion networks* [Man00].

These networks are used to represent a set of alternative sentence hypotheses, based on *confusion sets* which correspond to confusions between word hypotheses at a considered position of the sentence to recognize (these sets correspond to the edges between two nodes, on figure 12). This representation relies on *word posterior probabilities*. The posterior probability of a word is the sum of the probabilities of all graph paths this word belongs to. These word posterior probabilities can then be used to retrieve the best sentence hypothesis or as confidence measures on the words.

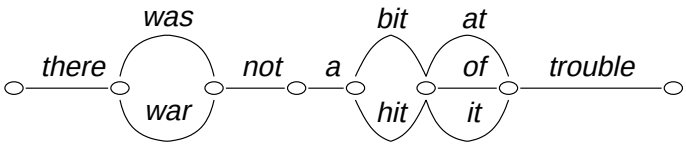


Figure 12: Example of a confusion network.

We integrated confusion networks in our handwriting recognition system to improve its performance. To our knowledge, confusion networks have not been used yet in the field of handwritten sentence recognition. This confusion network is jointly used with the MAP decoding (already present in our recognition system [QA06]), to detect potential recognition errors. The word posterior probabilities as well as other sources of knowledge are then used to correct these errors. Moreover, a rejection strategy is introduced to reject words that cannot be corrected by the presented approach. This rejection strategy allows the highlight of misrecognized words in an input interface or enables an additional recognition step on

[Man00] L. Mangu, *Finding Consensus in Speech Recognition*, PdD Thesis, Johns Hopkins University, 2000.

[QA06] 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.

these rejected words. The whole detection and correction has first been used on manually segmented sentences [26]. Then, we extended it to the processing of automatically segmented sentences [27]. In this latter case, we had to handle several segmentation hypotheses, thus dealing with possible under- and over-segmentation problems.

Experiments carried out on 320 handwritten sentences showed the interest of the proposed error detection and correction approach. Indeed, its use allowed a 31.3 % word error rate relative reduction when sentences are manually segmented and a 60 % reduction when words are automatically extracted.

6.8 Automatic word extraction of on-line sentences

Participants: Solen Quiniou, François Bouteruche, Eric Anquetil, Guy Lorette.

In [QA06], we have presented an on-line handwritten sentence recognition system which needs a manual segmentation of sentences. In [28], we have proposed a method to automatically extract words from the considered sentences. This task is done before the recognition of the sentence and consists in two steps: the extraction of the words according to the best segmentation and then the generation of additional segmentation hypotheses.

During the first step of the segmentation task, the extraction of the words is based on the characterization of the gaps between two consecutive strokes of the handwritten sentence. Three types of gaps are considered: intra-word gaps, inter-word gaps and inter-line gaps. This gap characterization relies on the computation of the distance between the two considered strokes: the considered distance is the x-distance Δx_{ref}^{new} between the most-on-the-right point P_{ref}^{mtr} of the reference stroke S_{ref} and the most-on-the-left point P_{new}^{mtl} of the newly written stroke S_{new} (see figure 13). To avoid skew and slant problems, we only considered points between the lower and the upper baselines.

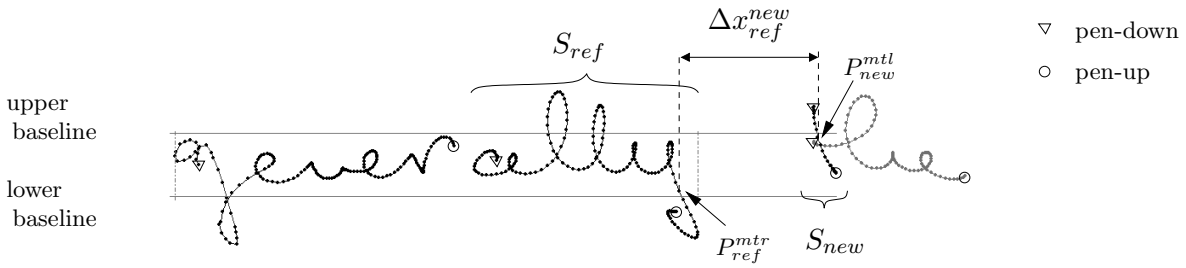


Figure 13: Example of distance Δx_{ref}^{new} computation.

A Radial Basis Function Network (RBFN) is then used to classify the inter-stroke gaps, using the distance Δx_{ref}^{new}, the maximum and median distance Δx_{ref}^{new} as well as the distance between the top of the S_{new} bounding box and the lower baseline. The three outputs of the RBFN are the score obtained for each gap type and the inter-stroke gap type is then the type associated with the highest score.

The second step of the extraction task aims at dealing with over- and under-segmentation problems. We then proposed a method which allows the generation of additional extraction hypotheses by also limiting the total number of segmentation hypotheses. A confidence index is thus added to the RBFN result, evaluate the reliability of the first answer. A reconsideration threshold is then learned on the difference between the two best scores. This index is used to detect potential under- and over-segmentation and also to generate additional segmentation hypotheses. A word graph is finally used to store all the segmentation hypotheses, as shown by figure 14 where additional segmentation hypotheses are represented by dotted edges.

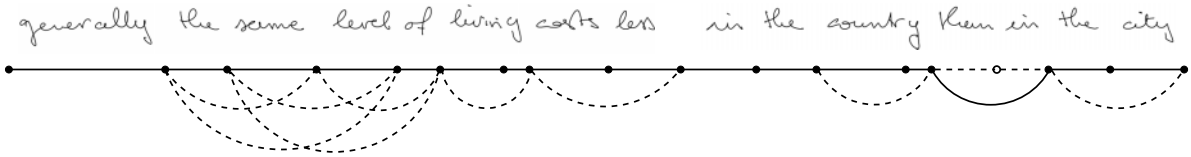


Figure 14: Example of a word graph with additional extraction hypotheses.

Experiments were carried out on 395 sentences (6 038 words) written by 15 writers. The extraction strategy which reconsiders 20 % of the inter-stroke gap classification results allows the augmentation of the word extraction rate which rises from 90.1 % (without additional hypotheses generation) to 97.8 %. The word recognition rate also increases from 79.1 % to 85.4 %, which gets closer to the 87.9 % rate achieved with a manual segmentation

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

Participants: Sébastien Macé, Eric 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 (see for instance section 5.1). 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. More precisely, our goal is to design pen-based software for structured document composition. In order to take advantage of the interaction with the user, we interpret the user strokes *on the fly*, directly during the composition of the document.

We propose a new generic method, called *DALI*, for the design of pen-based software for incremental document composition. The main component of this method is a new formalism for the description of documents and their composition which can be adapted to various domains. It is called *Context-Driven Constraint Multiset Grammars* (*CD-CMG*). CD-CMG are based on *Constraint Multiset Grammars*, but they also model new natures of knowledge:

- in order to model the incremental composition of structured documents, CD-CMG model not only a static description of a document, but also dynamic information about the way it can be drawn. For that purpose, their first originality is to model the relative

positioning of the document elements ; this *global vision* of the document is exploited to filter the possible interpretations of a hand-drawn stroke depending on its context, before having a *local vision* of this element in order to interpret its shape.

- CD-CMG make it possible to couple a structural recognition with a statistical recognition of the document elements. They also exploit the fuzzy logic framework for such purpose, which is well adapted to take into account the imprecision of hand-drawing: it makes it possible to evaluate the degree to which an interpretation is likely and facilitates the decision making process. This is the second main originality of the formalism and has been presented more in detail in [22].

Like presented last year, the presented methodology has been used to design several pen-based editor prototypes, for instance for musical score (see section 5.7) and various other musical notations [15], graphs or UML class diagram composition and editing. This year, we have worked more specifically, in collaboration with the Evodia society (see 7.1), on the design of *Script&Go Electricité*, a pen-based software for the composition and the editing of electrical diagrams. The goal was to evaluate the method in a real applicative context, with real industrial problematics. Results of these works have been published in [22, 23]. Figure 15 presents two screenshots of *Script&Go Electricité*. More information about this software are given in section 5.8.

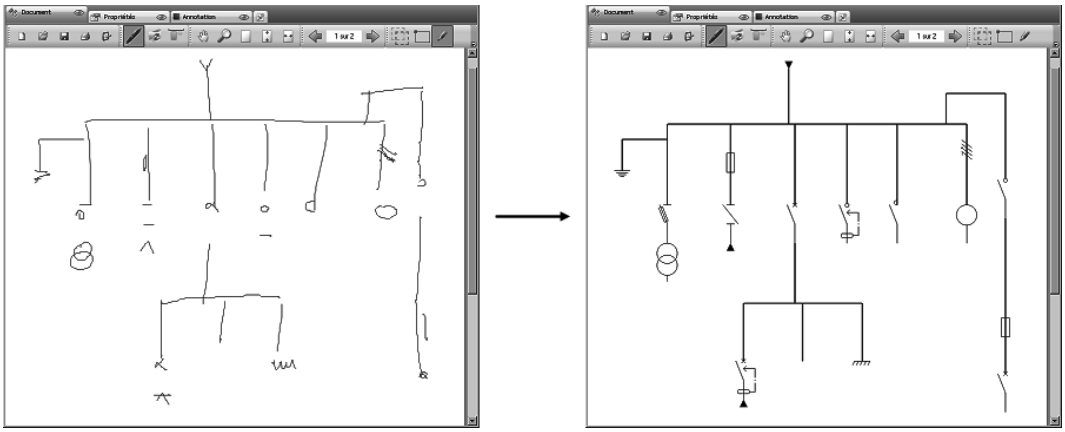


Figure 15: Screenshots of the *Script&Go Electricité* pen-based software: on the left the hand-drawn strokes, on the right the corresponding recognized document.

6.10 A General Method of Segmentation-Recognition Collaboration Applied to Pairs of Touching and Overlapping Symbols

Participants: Christophe Renaudin, Yann Ricquebourg, Jean Camillerapp.

We have developped a general method to segment and recognize pairs of touching and overlapping symbols. The method is based on the evaluation of several segmentation candidates produced by grouping elements of an over-segmentation. A pure general method would

involve a high number of segmentation candidates. From a general scheme (Figure 16), and according to the kind of application (digits, musical symbols, ...), we introduce an *a priori* knowledge, through heuristics, on the one hand to give the specific features of the problem, and on the other hand to improve the efficiency and the velocity of the method.

Our method was more precisely presented at the ICDAR 2007 conference [29].

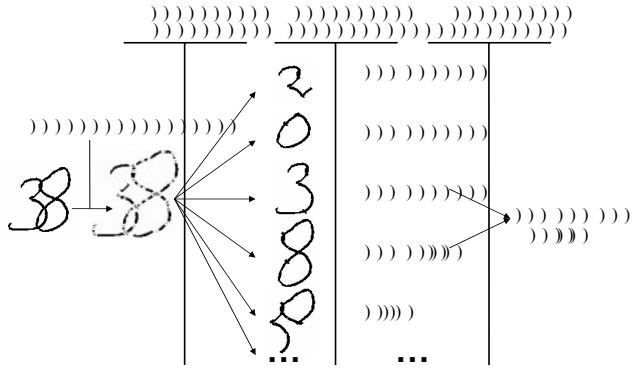


Figure 16: General scheme of the method

We tested it with images representing pairs of overlapping digits. Each image was automatically created by the random intersection of two single digit images randomly chosen in the IRONOFF database [VGLKB99].

We executed tests with automatic evaluation of the method result on a large number of examples (50,000) to obtain quantitative results for the efficiency of the method. For each example, the classes of the best candidate pair given by the method are compared to the ground truth, i.e. to the classes of the two initial digits of the example. If the candidate classes correspond to the initial classes, the example is ranked in the *success* category. When only one or none of the two candidates has a class corresponding to the class of one of the initial digits, the example is ranked in the *failure* category. When no pair of candidates is produced, the example is ranked in the *reject* category. Table 4 presents the obtained results.

Table 4: Results of the tests with automatic evaluation on 50,000 examples

Success	Failure	Reject
76.60%	20.04%	3.36%

We also executed tests with human evaluation on a smaller number of examples (1,000) to have a qualitative interpretation of the results. The results are detailed in [29].

[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.

7 Contracts and Grants with Industry

7.1 EVODIA

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

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

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

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

Participants: Eric Anquetil, Sébastien Macé.

- Company: EVODIA
- Contract: INSA

The *DALI* technology from the IMADOC team (see section 6.9) has been transfered to the Evodia company. It has been used to design the software Script&Go, a pen-based electric diagram editor (see section 5.8), which has been licensed to the *Apave* company.

7.3 THALES/EVODIA

Participants: Eric Anquetil.

- Company: EVODIA
- Contract: INSA

The *RESIF* technology from the IMADOC team (see section 5.1) has been transfered to the Evodia company. This year, Evodia has licensed RESIF to *Thales* company. RESIF has been embedded in a new experimental software for an enhanced avionic cockpit. The idea is to use sensitive screen to directly input parameters using handwriting recognition.

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

Participants: Eric 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.

7.5 Evaluation of Pen-based Interface Ergonomic Quality.

Participants: Eric Anquetil, François Bouteruche.

- Partner: 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.

7.6 Access to Handwritten Archives Documents

Participants: Bertrand Coüasnon, Eric 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 first research project which ended in October 2007. A new research project just started in October 2007 for 4 years (section 7.7). The first project was about making handwritten archives documents accessible to public. In collaboration with the *Archives des Yvelines*, Imadoc worked 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 [14] 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 Assisted Transcription of Handwritten Words in Archives Documents

Participants: Bertrand Coüasnon, Eric Anquetil, Jean Camillerapp.

- Partner: *Conseil Général des Yvelines, Archives des Yvelines*

- Contract: INSA 7030, INRIA 2702

This research project started in October 2007 and will end in September 2011. It is the second project of a long-term collaboration with the *Conseil Général des Yvelines* started in 2002 (the first one is described in section 7.6).

The objective is to work on automatic transcription of handwritten cursive words in archival documents, with the help of manual validation and interraction.

This project will involve one year of postdoc, two years of engineer and one PhD.

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. A platform with fast leaf-through capability on all these naturalization decrees is now available at the *Centre Historique des Archives Nationales*.

8 Other Grants and Activities

8.1 National initiatives

8.1.1 ANR MDCA Navidomass

Imadoc is involved in a 3 years ANR MDCA (Masses de Données - Connaissances Ambiantes) project Navidomass (NAVigation In Document MASSes), wich started in January 2007. The main research goal of this project is to work in a collaborative framework on the Analysis of Old Documents. This goal consists in developing Pattern Recognition and Image Analysis techniques that allow extracting knowledge from documents and converting them to Digital Libraries containing the scanned pages enriched with semantical information.

Partners of this project are: L3i (Université La Rochelle), Loria (Nancy), LI (Université de Tours), CRIP5 (Université de Paris 5), LITIS (Université de Rouen) and Irisa-Insa (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

Mohamed Cheriet, professor at ETS (Université du Québec) spent two day in our team, in the context of the PhD Thesis of S. Quiniou and to prepare future cooperation.

9 Dissemination

9.1 Leadership within scientific community

9.1.1 Program Committee

- E. Anquetil is member of the program committee of:
  - ICDAR 2009 (International Conference on Document Analysis and Recognition), Jul. 2009.
  - ICFHR 2008, (International Conference on Frontiers in Handwriting Recognition), Aug. 2008,
- E. Anquetil, B. Coüasnon and G. Lorette are members of the program committee of CIFED 2008 (Colloque International Francophone sur l'Écrit et le Document), Oct. 2008.
- 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.

9.1.2 Reviewing

- G. Lorette is a reviewer of IEEE ISCIP-06 (International Conference on Signal and Image processing, Dec. 2006, `http://www.icsip.org`.
- G. Lorette is a reviewer of IJIT-DM(International Journal of Information Technology and Decision Making, Dec. 2007.
- E. Anquetil is a reviewer of IEEE PAMI(IEEE Transactions on Pattern Analysis and Machine Intelligence).
- E. Anquetil is a reviewer of Pattern Recognition Letters.
- J. Camillerapp is a reviewer of Pattern Recognition Letters.

9.1.3 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).

- E. Anquetil is a member of the administration council of the society GRCE : " Groupe de Recherche en Communication Écrite ".

- E. 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, E. 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-PROFESSIONAL MITIC at l'IFSIC, University of Rennes 1.

- E. Anquetil and B. Coüasnon give lectures at MASTER-RESEARCH *d'informatique* of University of Rennes 1/INSA.

9.3 Participation in conferences, seminars, invitations, awards

- B. Coüasnon was invited for a course at the Ecole nationales des Chartes on : "Automatic Access to Old Documents", Paris, France, Dec. 2007.

10 Bibliography

Major publications by the team in recent years

[1] E. ANQUETIL, G. LORETTE, "Automatic generation of hierarchical fuzzy classification systems based on explicit fuzzy rules deduced from possibilistic clustering: application to on-line handwritten character recognition", *in : Proceedings of the International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems, (IPMU'96)*, p. 259–264, Grenade, Spain, July 1996.

[2] 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.

[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] 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.

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

[7] G. Lorette, J.-P. Crettez, "Reconnaissance de l'écriture manuscrite", *in : Traité informatique*, Techniques de l'Ingénieur, 1998, ch. H 1358, p. 1–15.

[8] 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.

[9] 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.

[10] 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.

Doctoral dissertations and "Habilitation" theses

[11] H. Mouchère, *Etude des mécanismes d'adaptation et de rejet pour l'optimisation de classifieurs : Application à la reconnaissance de l'écriture manuscrite en-ligne*, PhD Thesis, Institut National des Sciences Appliquées de Rennes (INSA), december 2007.

[12] S. Quiniou, *Intégration de connaisances linguistiques pour la reconnaissance de textes manuscrits en-ligne*, PhD Thesis, Institut National des Sciences Appliquées de Rennes (INSA), december 2007.

[13] L. Rousseau, *Reconnaissance d'écriture manuscrite hors-ligne par reconstruction de l'ordre du tracé en vue de l'indexation de document d'archives*, PhD Thesis, Institut National des Sciences Appliquées de Rennes (INSA), june 2007.

Articles in referred journals and book chapters

[14] 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 9*, 2-4, 2007, p. 223–242, `http://springerlink.com/content/5843461264501u81/`.

[15] S. Macé, E. Anquetil, B. Bossis, *Pen-Based Interaction for Intuitive Music Composition and Editing*, Idea Group, 2007, ch. 12, p. 261–288.

[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) 21*, 1, 2007, p. 99–116, `http://www.worldscinet.com/ijprai/21/2101/S0218001407005326.html`.

[17] H. Richy, S. Després, *Métadonnées, ontologies et document numérique*, H7155v2, Techniques de l'Ingénieur, 2007, à paraître.

Publications in Conferences and Workshops

[18] S. Bayoudh, L. Miclet, H. Mouchère, E. Anquetil, "Learning a classifier with very few examples: knowledg based and analogy generation of new exemples for character recognition.", *in : 18th European Conference on Machine Learning (ECML)*, september 2007.

[19] A. Lemaitre, J. Camillerapp, B. Coüasnon, "Contribution of Multiresolution Description for Archive Document Structure Recognition", *in : ICDAR'07*, p. 247–251, Curitiba, Brazil, September 2007.

[20] A. Lemaitre, J. Camillerapp, B. Coüasnon, "A generic method for structure recognition of handwritten mail documents", *in : Document Recognition and Retrieval (DRR XV)*, 2008. to be published.

[21] A. Lemaitre, B. B. Chaudhuri, B. Coüasnon, "Perceptive Vision for Headline Localisation in Bangla Handwritten Text Recognition", *in : ICDAR'07*, p. 614–618, Curitiba, Brazil, September 2007.

[22] S. Macé, E. Anquetil, "Context-Driven Constraint Multiset Grammars with Incremental Parsing for On-line Structured Document Interpretation", *in : Proceedings of the 9th International Conference on Document Analysis and Recognition (ICDAR'07)*, p. 442–446, Curitiba, Brazil, September 2007.

[23] S. Macé, E. Anquetil, "Design of a Pen-Based Electric Diagram Editor Based on Context-Driven Constraint Multiset Grammars", *in : Human-Computer Interaction: Interaction Platforms and Techniques, Proceedings of the HCI International 2007 Conference, Part II, LNCS 4551*, J. A. Jacko (editor), Springer, p. 418–428, Beijing, P.R. China, July 2007.

[24] I. Martinat, B. Coüasnon, J. Camillerapp, "An Adaptative Recognition System using a Table Description Language for Hierarchical Table Structures in Archival Documents", *in : Seventh IAPR International Workshop on Graphics Recognition, GREC 2007*, p. 1–11, Curitiba, Brazil, September 2007.

[25] H. Mouchère, S. Bayoudh, E. Anquetil, L. Miclet, "Synthetic On-line Handwriting Generation by Distortions and Analogy", *in : 13th Conference of the International Graphonomics Society (IGS2007)*, p. 10–13, november 2007.

[26] S. Quiniou, E. Anquetil, "Use of a Confusion Network to Detect and Correct Errors in an On-line Handwritten Sentence Recognition System", *in : 9th International Conference on Document Analysis and Recognition (ICDAR'07)*, p. 382–386, Curitiba, Brazil, September 2007.

[27] S. Quiniou, E. Anquetil, "Utilisation de réseaux de confusion pour la reconnaissance de phrases manuscrites en-ligne", *in : 16ème Congrès Francophone de Reconnaissance des Formes et d'Intelligence Artificielle (RFIA'08)*, Amiens, France, Janvier 2008. à paraître.

[28] S. Quiniou, F. Bouteruche, E. Anquetil, "Word Extraction for the Recognition of On-Line Handwritten Sentences", *in : 13th Conference of the International Graphonomics Society (IGS'07)*, p. 52–55, Melbourne, Australia, November 2007. to be published.

[29] C. Renaudin, Y. Ricquebourg, J. Camillerapp, "A General Method of Segmentation-Recognition Collaboration Applied to Pairs of Touching and Overlapping Symbols", *in : Proceedings of the 9th International Conference on Document Analysis and Recognition*, p. 659–663, Curitiba, Brazil, September 2007.

[30] L. Rousseau, E. Anquetil, J. Camillerapp, "Word Extraction for the Recognition of On-Line Handwritten Sentences", *in : 13th Conference of the International Graphonomics Society (IGS2007)*, p. 44–47, Melbourne, Australia, November 2007. to be published.

Miscellaneous

[31] F. Bouteruche, S. Macé, E. Anquetil, "Reconnaissance de tracés manuscrits", *in : Revue électronique nationale de vulgarisation scientifique : interstices*, 2007, `http://www.interstices.info/display.jsp?id=c_17397&qs=id\`.