INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Project-Team IMADOC

Interprétation et Reconnaissance d'Images et de Documents

Rennes

THEME 3B

Activity Report

2003

Contents

1 Team

Head of project-team

Jean Camillerapp [Professor Insa]

Administrative assistant

Myriam David [TR Inria]

Staff member (Inria)

Jacques André [Research Director]

Bertrand Coüasnon [Research scientist (secondment)]

Staff member (Insa)

Éric Anquetil [Associate Professor]

Ivan Leplumey [Associate Professor]

Yann Ricquebourg [Associate Professor]

Staff member (Université of Rennes 1)

Guy Lorette [Professor]

Charles Quéguiner [Associate Professor]

Hélène Richy [Associate Professor]

Nicolas Ragot [Lecturer from 01 Oct 2003]

Ph. D. student

Nicolas Ragot [University research grant up to 30 Sep 2003]

Sabine Carbonnel [Menrt research grant]

Laetitia Rousseau [Insa research grant from 01 Oct 2003]

Technical staff

Laurent Pasquer [Contract département des Yvelines to 30 Sep 2003]

Grégory Maitrallain [Contract département des Yvelines]

Fabrice Boyer [Contract Insa from 01 Apr 2003 to 30 Sep 2003]

2 Overall objectives

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

3 Scientific foundations

3.1 Early vision

Key words: early vision, binarization, skeletonization, Kalman filtering.

Adaptative binarization

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

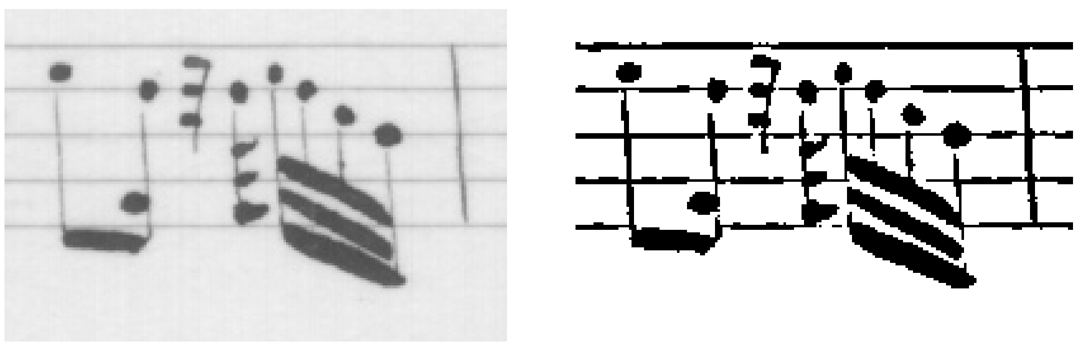
(a) Threshold of 210 : allowing the detection of thin lines

(b) Threshold of 195 : the vertical elements remain

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

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

A digitalization in gray levels allows to implement algorithms of adaptative binarization, often based on cooperation between edge detection and region growing. Those give a local 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 objects characteristics are sensitive to space sampling noise. So increasing space resolution, improves the relevance of parameters related to visual indices. With 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 this parameters.

Gray level image skeletonization

Skeletonization algorithms have been developed to thin images of lines. But in documents images, it is necessary to distinguish in the layout (figure 3) the area really skeletizable (also called regular area) and the singular areas, principally localised in crossings in which the concept of median axis does not have practical interest [10].

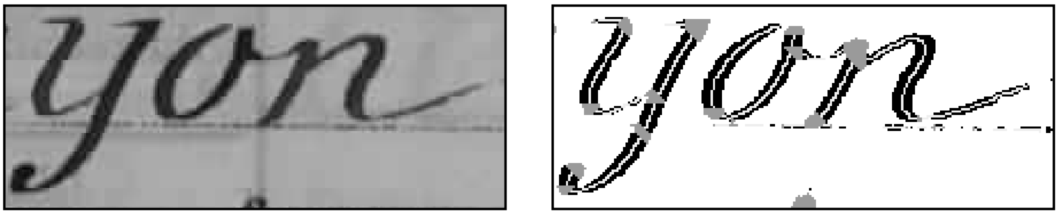


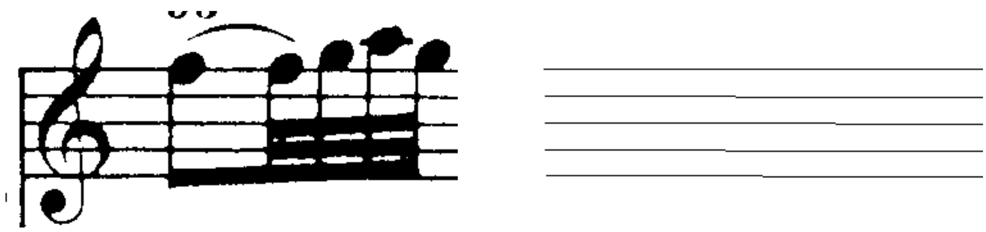
Figure 3: Processing of an image from an old document. The skeleton appears in white in the medium of the line and singular areas are in dark gray.

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

Segmentation with Kalman filters

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

The Kalman filters is a technique of parameters identification from an ordered series of measures. In the case of lines, the model is reduced to the thickness of the line, its slope and the equivalent of the ordinate at zero point. Measure results from the position and the thickness of the black run-length[1] orthogonal to the drawing.



(a) Original image

(b) Extracted stave lines

Figure 4: Detection of stave lines with Kalman filters

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

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

a measure to a segment and on the other hand the probability to continue an assumption disturbed by the presence of another object [4].

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

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

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

3.2 Dmos, a Generic Method for Structured Document Recognition

Key words: syntactical recognition, grammar, definite clause grammar, segmentation, structured document recognition, musical scores, mathematical formulae, table strucrture, form.

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

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

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

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

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

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

EPF Language

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

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

Position Operator (encapsulated by `AT`):

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

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

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

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

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

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

With this syntax it is possible, for example, to describe a simplified beamed note (a beamed eighth note, with only 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 neccessery 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 specific properties needed for parsing bidimensional documents. Compares to classical monodimensional parser the main properties of the bidimensional parser we develop are:

- changing the parsed structure during parsing for contextual segmentation. The parsed structure can then be changed to be like it should be if there was no segmentation error;

- detecting of 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 and so everywhere in the parsed structure. The parser uses the position operators to select the next element to parse ;

- dealing with noise. We can consider that dealing with noise corresponds to find 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 developed a generic recognition method for structured documents. We implement 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 stucture of the document with an EPF grammar which is then simply compiled. In this way we have developed various recognition systems: on musical scores, on mathematical formulae, on recursive table structures...

3.3 Pattern recognition by fuzzy inference systems

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

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

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

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

Fuzzy inference is then defined as below :

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

where

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

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

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

[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, it is able to absorb the variability of the handwriting shapes, and can cope with limited computing and memory resources. RESIF software has successfully been integrated on a Smart Phone device with limited resources (cf. section 7.1). Recently, a new classification approach, named Mélidis, has been elaborated. This new system can be considered as an extension of RESIF. The main goals are to obtain a more generic recognition system able to deal with different kinds of classification problems without needed *a priori* knowledge and without loss of transparency so that optimization can be done afterwards by experts (cf. section 6.6).

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 developed (Dmos method) allows a quite fast adaptation to new kind of document. Documents we already worked on are:

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

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

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

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

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

- Retroconversion of paper documents to avoid a manual input to get a usable electronic version. It can for example be used for a new edition, or to produce a Braille document;
- Kind of document identification, to make for example an automatic management of faxes;

- Automatic production of indexes and annotations for an automatic access by content to documents (section 6.3);
- Detection of specific areas in a document to improve its access (cf. section 6.3).

4.2 Handwriting recognition systems

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

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

In *on-line recogntion* a huge market has arise due to the recent economic developpement of the following devices:

- pocket computers (Palm, PocketPC),
- tabletPCs (computers without keybord 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 and 7.1).

4.3 Pen-based interaction

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

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

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

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

Most applications will provide the users with a multimodal interface that enable the choice among various modes for issuing commands.

5 Software

5.1 RESIF : Handwriting Recognition by Hierarchical Fuzzy Inference Systems

Contributed by: Éric Anquetil.

Key words: Handwriting Recognition, smartphone, fuzzy logic.

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

- RESIFCar is the software specialized to recognize handwritten characters: latin alphabet, digits and special symbols.
- RESIFMot is the software for unconstrained cursive handwritten words recognition.
- RESIFApp is the automatic learning process that generates from a handwritten character database the hierarchical fuzzy models used by the recognition systems: RESIFCar and RESIFMot.

RESIFCar and RESIFApp are yet in their fourth version. In 2001, through an industrial collaboration with PurpleLabs company [2], RESIFCar has been successfully integrated into mobile devices (smartphones) which can be characterized by their limited computing and memory resources. Theses mobile phones are today commercialized in Europe (cf. section 7.1).

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

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

Contributed by: Bertrand Coüasnon.

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

Key words: recognition, structured document, musical scores, mathematical formulae, table structures, forms, archives.

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

This generator allows us a fast adaptation to a new type of document. Indeed, it is only necessary to define a new grammar in EPF, which describe 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 produce by compilation.

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

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

5.3 Library to extract visual indices

Contributed by: Jean Camillerapp.

Key words: Early vision, adaptative binarization, vectorisation, Kalman filtering, squeletonization, multi-resolution..

It is the library on which the DocRead software described in the section 5.2 is based.

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

Using Kalman filters 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 the best adapted to the recognition of all the structures. This is why the library also provides possibilities of multi-resolution processing.

6 New results

6.1 A platform for image document annotation

Contributed by: Bertrand Coüasnon, Grégory Maitrallain, Ivan Leplumey, Fabrice Boyer.

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

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

We choose to develop the platform with a classical architecture: a web server (Apache) with a servlet container (TomCat); the Java servlet access to a relational database (PostGreSQL) to store annotations and send them to the client: a java applet running in a web browser (figure 5).

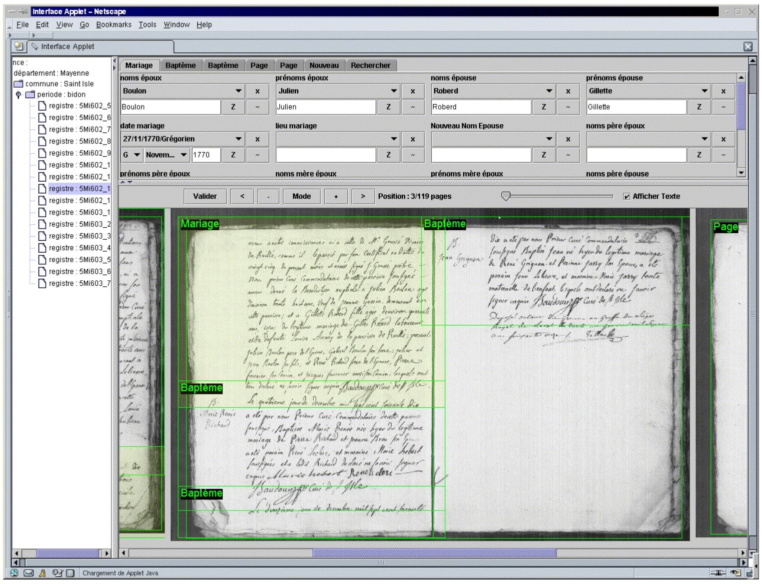


Figure 5: One interface to consult, annotate, retrieve Archives documents.

With the platform, on a web browser, a user can leaf through images of Archives documents. When a page is displayed, all the associated annotations are presented on the interface:

geometric annotations are drawn on the image, the textual annotations are presented in tabs for the nodes of the structure of annotations (a marriage certificate ...) and in field boxes for the leaves (name, date ...)(figure 5). The reader can consult annotations, add or modify annotation (if he has the right to), but is limited by the allowed annotation structure given by the DTD configuration file, according to the kind of document. The system can also store various interpretations if readers do not agree.

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

We present in the section 6.3 examples of automatic annotations which are then managed by this platform on various kinds of Archives documents. We show the interest of automatic annotations and the complementary of the automatic and manual annotations.

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 (figure 6).

Figure 6: A pen-based interface on the platform for annotations

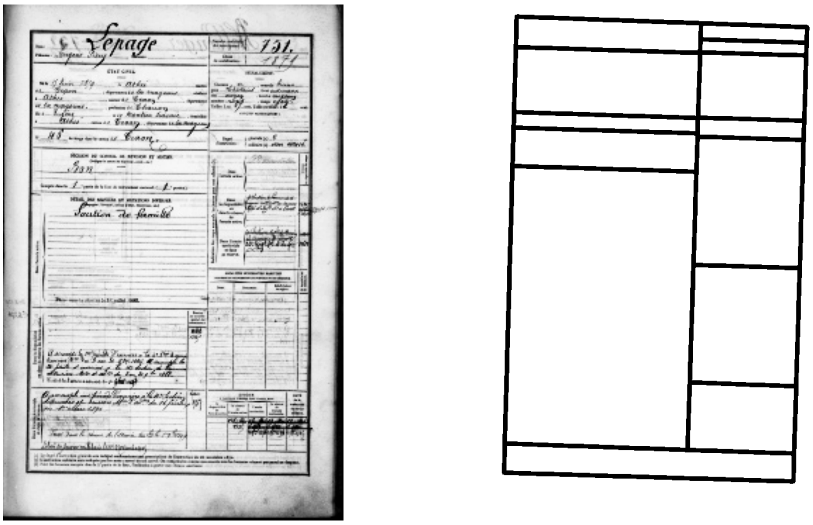
6.2 Indexing old printed forms with handwritten last name

Contributed by: Laurent Pasquer, Jean Camillerapp, Bertrand Coüasnon.

Digitizing old documents is important to protect originals and to view them on Internet. However, it gives nothing more to readers for retrieval of the document they are looking for. Handwriting recognition in old documents is, most of time, extremely difficult or impossible to do. Nevertheles on some kind of old documents it is possible to think about automatic indexing of handwritting.

In structured documents as printed forms, indexation of some squares would be a very uesful tool to select documents. Military forms of incorporation of the 19th century, of which there are million copies, are an example of indexable documents by their handwritten contents. Indeed these forms have a square *name* in which the last name is well handwritted.

DMOS method (cf 3.2) makes it possible to locate in the image the various square included in the form and in particular the square *name* (figure 7). After that this square only is processed.



(a) A military form (b) Localized squares

Figure 7: Automatic extraction of the structure of a military form

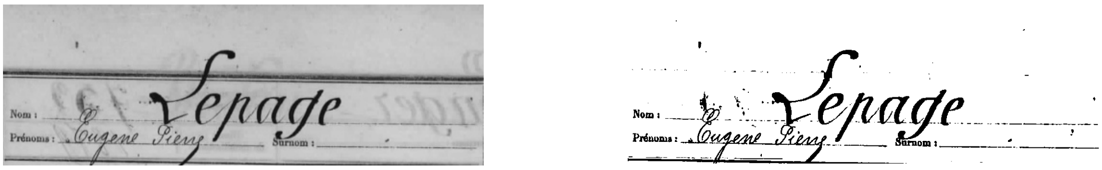
Last names are written by using an alphabet quitte standardized: the slanting roundhand. In this handwritting style letters are systematically decomposed using a very limited vocabulary of forms: the *graphem*. So to index images of military forms we chose to represent them with a segmentation in *graphem* of the last name.

For that, images are first binarized and then segmented into regular area and singular area using methods descibed in section 3.1. Then regular area are merged to obtain the layout of the handwritting.

Each layout, represented by a dozen of characteristics, becomes the input of a classifier. This one assigns this layout to a class of *graphem*. It is this list of names of classes which constitutes the index of the image.

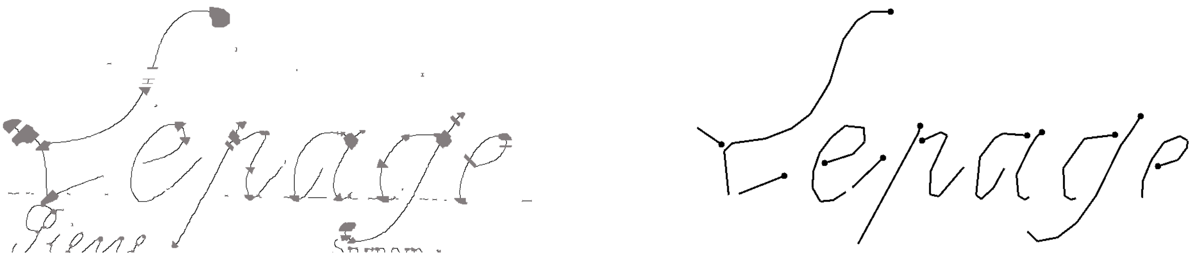
To seek the images which correspond to a patronym, an user formulates his request with the string of characters which spell the name. Then each character of the string is systematically translated in *graphem*.

The request is then compared with the index of all images using an edit distance. So it is possible to order the list of the images to present to the user the nearest images to this request (figure 10).



(a) Original image (b) Binarized image without frames lines

Figure 8: Binarization of the square *name*



(a) Regular and singular areas (b) Graphems

Figure 9: From image to graphem

This research, described more in detail in [21], was carried out with a grant of the department of Yvelines (section 7.3).

6.3 Automatic and collective annotations on archives documents

Contributed by: Bertrand Coüasnon, Jean Camillerapp, Ivan Leplumey, Laurent Pasquer.

6.3.1 Register of births, marriages and deaths

Automatic Annotations Those documents are really difficult to automatically annotate, due to the weak structure and the poor quality of the handwritten text. The documents are scanned by double pages. We defined a grammar in EPF describing the notion of page. With the DMOS method we have been able to produce a recognition system which detects the position of each page and produces automatic annotations. A new test has been done this year on 4,008 double pages: 99.3% have been correctly detected with 0.7% of reject.

With these page annotations, a reader can leaf through a register page by page with a zoom automatically adapted on the page, making the reading more comfortable.

Collective Annotations On those pages, collective annotations can be added: the type of certificate (birth, marriage ...); for a birth, annotations can be for example: name and surname of the child, place of birth, name of the mother, name of the father ... The position of the certificate can be defined by the reader or just be associated to the automatic annotation of page. Of course there is no obligation of filling all the fields.

Applications A demo of the platform on civil status registers is available at the address: `http://imadoc-ar.irisa.fr`. This demo presents only two districts with the automatic annotations on the position of pages and the collective annotations, on 4,000 images (8,000 pages). This demo is also available in the reading rooms of the *Archives départementales de la Mayenne* and the *Archives départementales d'Ille-et-Vilaine* (figure 5).

6.3.2 Register of Military Forms

Automatic Annotations These documents are made of quite damaged military enrollment form of the 19th century. Various problems make them difficult to recognize: the size of the cells change from year to year, there is a lot of pasted sheets of paper which hide the form structure, we can find stamps on it, ink bleeding through the paper ... We have then defined a specific EPF grammar which can take into account those difficulties.

From this specific (and small) EPF grammar, we automatically produced a new recognition system to detect the form structure (figure 7). We have successfully tested last year this recognition system on 60,223 forms from the *Archives départementales de la Mayenne*. This year the system has been tested on 28,731 forms from the *Archives départementales des Yvelines*. Those military form are more damaged than those from the *Archives départementales de la Mayenne*. Despite this, 27,874 forms have been correctly recognized, their structure with the cells positions have been correclty detect (97.02%). Each cell produces an automatic annotation: a geometric annotation (the polygon of the cell) and a textual annotation (the name of the cell).

The automatic indexing of handwritten last names in those military forms is possible with the help of cells locations (see section 6.2).

Collective Annotations By changing the DTD file on the platform, it is possible to specify the allowed annotations on these military forms. For example, the cell containing some birth information or the cell containing a physical description of the person, could be collectively annoted. The user will have to select the cell (an automatic annotation) to zoom on it and to associate to it some textual annotations. All those annotations could then be used for a future query by another reader.

Applications The 60,000 images (automatically cropped to remove protected information by using the geometric annotations automatically produced) are publically available on the *Archives départementales de la Mayenne* web site (`http://www.cg53.fr/Fr/Archives/` follow *Archives en ligne* then *Conscrits de la Mayenne*).

The platform with an automatic access by handwritten last names on those military forms will be available in January 2004, at the *Archives départementales des Yvelines* in the reading

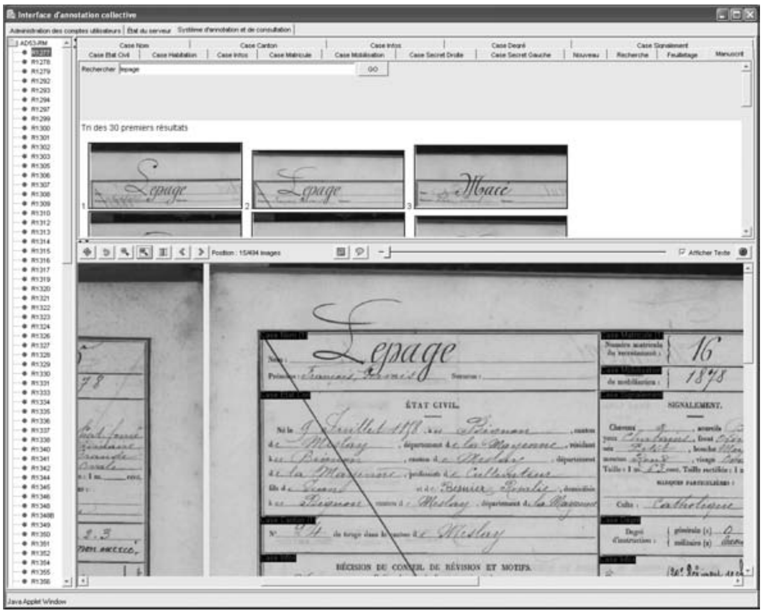


Figure 10: Access by handwritten names in military forms. The request is "lepage" among 494 pages (names). Answers are presented in the upper part (two "lepage" are automatically selected). The reader can then annotate some cells if he wants to.

room as well as on Internet. Readers, after retrieving the right document by a request on handwritten names will be able to add collective annotation on it. At the beginning 28,000 pages will be open to public (figure 10).

6.3.3 Naturalization Decrees

Automatic Annotations These documents are from the end of the 19th century and the beginning of the 20th century. They are unique documents which are for some people the only one which can justify their French nationality. A decree is usually made of around ten pages. They are fully handwritten or sometimes fully typewritten. They are organized in two columns with paragraphs where each paragraph concerns one person. His name is usually the first name in the paragraph. To retrieve the decree concerning one person is very tedious: the reader needs to leaf through all the pages of all the decrees.

Compared to the military forms, the structure is very weak because only made of the organization in paragraphs of the handwritten text. Due to the genericity of the DMOS method, we have been able to define an EPF grammar describing the organization of decrees in handwritten text-line, in paragraphs and columns, only with the help of the connected components detected in the image.

From this description, by compilation, a recognition system have been produced, which is able to detect the position of the name and the file number. A new test has been done this year on 1,126 images: 7,525 names or file numbers have been correctly detect (their positions)

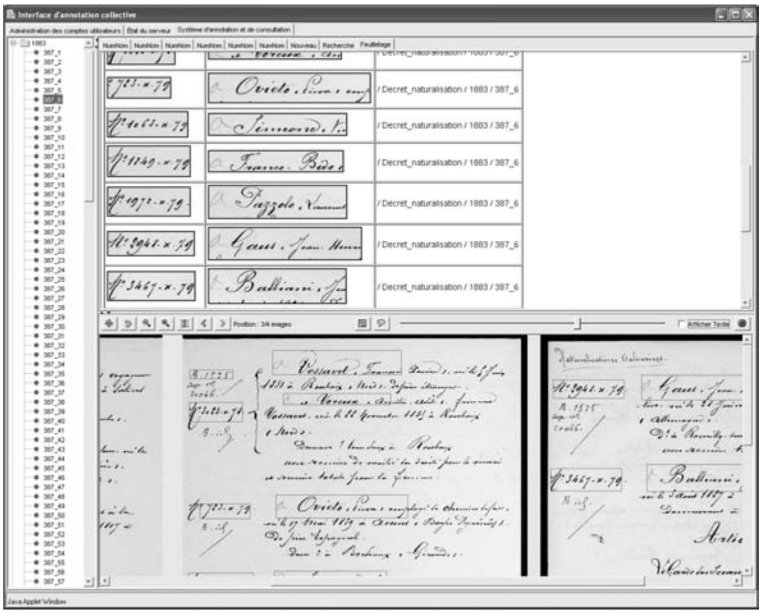


Figure 11: Fast leaf-through using automatic annotations. Names are presented in a table in the upper part. The reader can then access to the complete page and annotate it if he wants to.

on 7,568 (99.43%). This rate has been done with 575 false detection.

These positions are the automatic annotations which are added in the platform. With these annotations the platform can present a table with only images of the file number and the name. This allows much faster leafing through a decree.

Collective Annotations When a reader found the name he was looking for, the platform presents the original page with all the existing annotations. The reader can then leaf through the original pages and if he wants, add some collective annotation like the name or the date of birth ...

Applications This work on naturalization decrees has been done in cooperation with the french national Archives (*Centre Historique des Archives Nationales*). A fast leaf-through capability using automatic annotations has been also included in the platform. A reader can then select the name he his looking for and has a direct access to the full document ready to add collective annotation (figure 11).

6.4 ILib: a feature extraction library

Contributed by: Bertrand Coüasnon, Grégory Maitrallain, Yann Ricquebourg.

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

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

To achieve this goal, we have been working at the definition and the progressive implementation of a portable C++ API. First concerning "classical" features often cited as required by common recognition systems (as surface, gravity center, curvature...), the library tends now to integrates more original or promising characteristic functions as Zernike moments, (concerning plain pixel images) see Figure 12, or elliptic Fourier descriptors (concerning contour curves) see Figure 13. Both these functions have interesting hierarchized information properties, thus leading to a progressive separation of major factors and details, associated with a high power of information concentration since only a few descriptors leads to a quite total representation of the characterized prototypes.

Besides, we are developing a graphical user interface (see Figure 14), for the test and validation of those feature extraction functions. The GUI intends to be as generic as possible since it has been based on the common API we defined in the framework of the library.

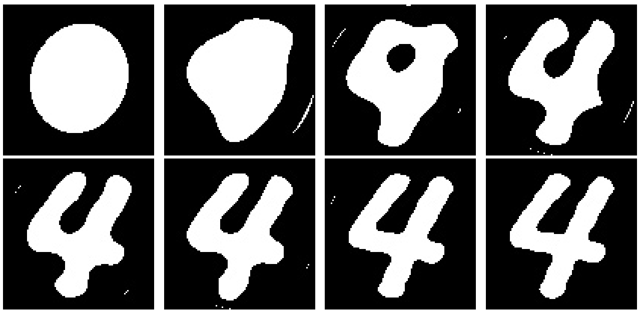


Figure 12: Progressive synthesis of the image of a digit 4, using its Zernike moments of order up to 2, 6, 8, 10, 12, 14, 18 and 20

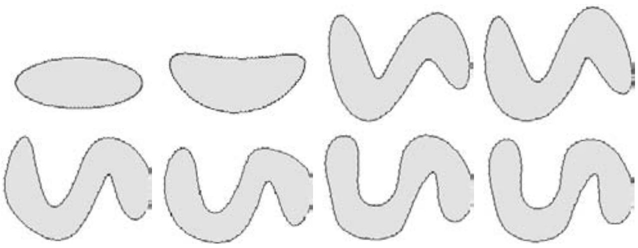


Figure 13: Progressive synthesis of a contour curve, using its elliptic Fourrier coefficients of order up to 1, 2, 3, 4, 5, 6, 7 and 8

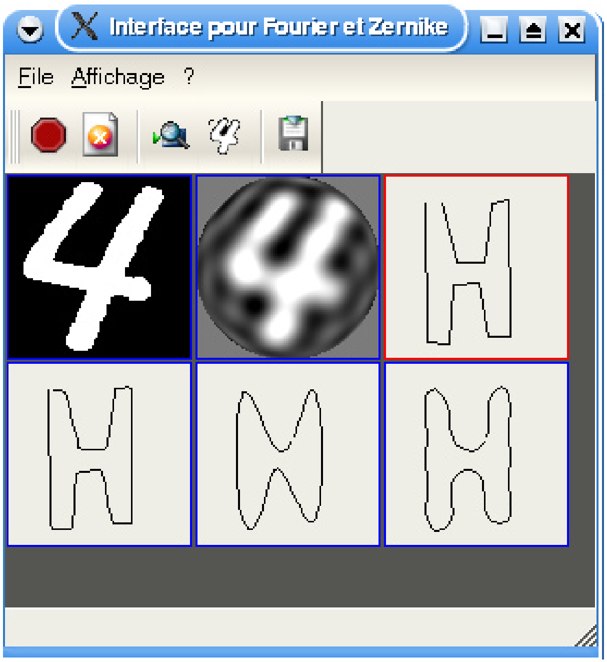


Figure 14: The test and validation graphical interface showing a Zernike moments application on a digit 4 and a elliptic Fourier coefficients computation on a letter H contour curve

6.5 Generic method of symbol segmentation

Contributed by: Bertrand Coüasnon, Yann Ricquebourg.

For the recognition of strongly structured documents (musical partitions, mathematical formulas, tabular forms...), the DMOS system developed in the Imadoc team consists of a grammatical formalism and an associated analyzer (see section 3.2). This analyzer makes it possible to introduce the context (modelled by the descriptive grammar of the type of document) formulating an *a priori* on the symbol to be found in the phase of segmentation (allowing, for example, to separate two connected symbols and thus to recognize them). This approach is based on the segmentation of the processed image which provides the lexical "terminals" of the grammatical analysis. However in the existing system it is necessary to define a specific segmentation/recognition algorithm for each kind of symbol. This hampers its generalization and its adaptation to new documents.

The prime objectives of this work are thus to develop a generic method of segmentation being able to adapt to any symbol, which moreover enables to deal with difficult cases where the initial segmentation did not isolate correctly the symbols to be recognized (over or under-segmentation). We explore the possibilities of basing this adaptation on a traditional modeling by statistical training: an analysis in principal components, providing a general model and its variation modes.

The main advantages of this technique, validated in nearby contexts (recognition of small face images, of road signs, etc.) are:

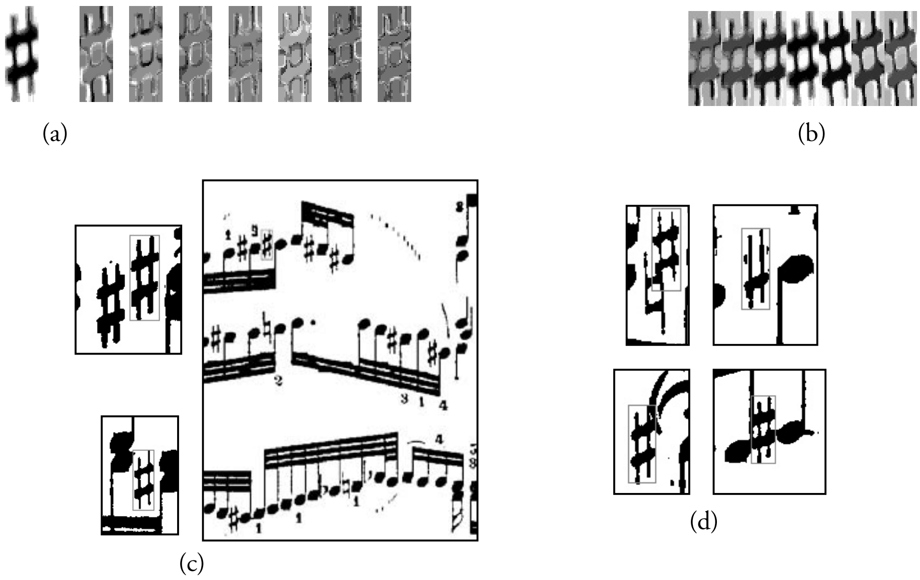


Figure 15: Results of detection/resegmentation using the analysis in principal components of a training set of sharps. (a) Resulting model and some associated eigen vectors. (b) Illustration of the deformation amplitude allowed by the first eigen vector. (c) Detection/recognition in three disturbed environments. (d) Detection/recognition in spite of difficulties (alteration, connexity)

- to be generic. The training is indeed based on a population of data vectors which can straightforwardly be the pixels of symbol images but also any numerical characteristics suitable for the problem.

- to produce, with the eigen vectors computed in the statistical transformation (figure 15a), a set of discriminating characteristics hierarchized by degree of representativeness (from the common factors to the rare details). The experiments usually shows that one can then reduce the set to a few first eigen vectors. This reduce the CPU cost of the recognition phase, while preserving a very good cover of the population to be characterized.

- to allow a re-segmentation: associated with an algorithm of optimization, one can extract the zone from the image corresponding best to the learned models (figure 15c). Therefore, such a technique can cope with the delicate cases where a simple connexity component segmentation is not satisfying to lead to a recognition (because resulting in a over or under-segmentation of the objects to identify). One can change the segmentation and propose another one better suited to the recognized part (figure 15d).

We recently focused on a optimized new implementation of this algorithm, whose performances are interesting in spite of the heavy cost of the optimum search. Sufficiently optimized, this mechanism will advantageously be integrated into the DMOS analysis chain, conferring then the possibility to re-examine the initial segmentation and to handle recognition failure areas.

6.6 MÉLIDIS: Pattern recognition by intrinsic/discriminant dual modeling based on a hierarchical organization of fuzzy inference systems

Contributed by: Nicolas Ragot, Éric Anquetil, Guy Lorette.

In previous works, the classifier ResifCar [2] was specifically elaborated for on-line cursive character recognition using knowledge about the different possible allographic shapes and the way they are written. This system is based on an explicit modeling by hierarchical fuzzy rules. Indeed, during the design, a great care has been taken to elaborate a “transparent” (understandable) system which could be improved after the learning phase by an expert. Thanks to these transparency properties, ResifCar can actually achieve good recognition performances for the latin letters, digits and some usual symbols. Moreover, thanks to the compactness of the modeling process, it has successfully been integrated on a Smart Phone device with limited resources. But the strength of the system is also its main drawback: its dedicated architecture based on *a priori* knowledge on cursive handwriting makes its adaptation difficult to other classification tasks.

The Mélidis system is a new classifier which can be considered as an extension of Resifcar. The main goals for the elaboration of this system were to obtain a generic recognition system able to deal with different kinds of classification problems without need *a priori* knowledge and without loss of transparency so that *a posteriori* optimization can be done by experts. The system should also be as compact as possible for integration on devices with limited resources such as PDA or Smart Phone.

The resulting classifier is based on an original hybrid architecture composed of two complementary modeling levels. The first level models each class to be recognized according to its intrinsic characteristics. Using this kind of modeling, the system is able to find automatically stable subclasses that compose the initial classes. The intrinsic models (description of the subclasses by fuzzy prototypes) are generated by an adapted fuzzy clustering algorithm like the Possibilistic C-Mean (PCM) [Kri94] on each class separately.

The aim of the second level of the hybrid system is to operate a progressive discrimination of the classes. To do this, decision boundaries between symbols of a given class and others that are confusing according to the intrinsic level (this not necessarily whole classes) are searched. This is done by specific Fuzzy Decision Trees (FDT) dedicated to discrimination.

For recognition of unknown shapes, both levels are formalized homogeneously by fuzzy inference systems that are merged for final decision.

To evaluate this generic hybrid system and its aptitude to adapt itself to different classification problems, several experiments were done on classical benchmarks and on on-line

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

character recognition (digits and lower-case letters). In all cases, results are close to other approaches among the best (including a Support Vector Machine) which is promising since all the capacities of the system were not exploited. Moreover, the compactness of the Mélidis system is far more important than that of a Support Vector Machine. Thus, integration on low resources devices should be easier.
This work has also pointed out several interesting perspectives. One of them is: how to take advantage of the two kind of knowledge (intrinsic and discriminant) and their formalization by fuzzy inference systems to integrate efficient rejection management and writers adaptation processes? Some preliminary studies are presently carried out.

6.7 Lexical Post-Processing Optimization for Handwritten Word Recognition

Contributed by: Sabine Carbonnel, Éric Anquetil.

We study lexical post-processing optimization for on-line handwritten word recognition. The aim of this work is to explore the combination of different lexical post-processing approaches in order to optimize the recognition rate, the recognition time and memory requirements of an on-line handwritten recognizer system. The present method focuses on the following tasks: a lexicon organization with word filtering, based on holistic word features to deal with large vocabulary (creation of static sub-lexicon compressed in a trie structure); a dedicated string matching algorithm for on-line handwriting (to compensate the recognition and the segmentation errors); and a specific exploration strategy of the results provided by the analytical word recognition process.

In the following paragraph we present the principles of a lexical postprocessing. Two approaches are presented and we explain for each one the training phase (or organization in sub-lexicons). The exploitation of the lexical knowledge is presented at the following paragraph. In the last paragraph an experimental comparison between the two approaches is exposed.

Principle of a lexical post-processing
The two presented approaches of lexical postprocessing are composed of two phases:

- the training and the indexing phase of the sub-lexicons (phase of organization);
- the exploitation phase of the organized lexicon (postprocessing phase) For the two approaches, the postprocessing phase is based on the comparison (using an edit distance) between the recognition hypotheses and the words of a relevant sub-lexicon wich is selected between all the other (lexicon reduction).

The choices concerning the lexicon organization, the index and the way of carrying out the reduction are very significant because they have a high influence on the postprocessing quality (speed, recognition rate and memory requierement).

Lexical organization
Within the context of a postprocessing with lexicon, the reduction consists in selecting a relevant lexicon using various criteria. The objective is to save time while exploring only a part of the dictionary. We present two approaches of organization of the lexicon [23].

In the first approach, the words are grouped together according to their visual similarity (generic shape) and their size. The reduction step then consists in selecting the sub-lexicons having exactly the same generic silhouette and a similar size to those of the recognition hypotheses. In the second approach, vectors of characteristics are extracted from the word and are gathered in fuzzy partitions (indexed by their average centers or vectors) using an unsupervised classification algorithm. The reduction is carried out by the calculation of the distances between the hypotheses vectors and the index of the sub-lexicons: the nearest sub-lexicons are selected. Compare hypotheses and words of the selected sub-lexicons allow to determine the most coherent word.

Post-processing

The lexical postprocessing is carried out in two stages:

- reduction of the lexicon (with the selection of a relevant sub-lexicon);
- comparison of the recognition hypotheses and the sub-lexicon words with a specific edit distance [22] in order to determine the word of the lexicon the nearest to the handwritten word (correction of segmentation and recognition errors).

We combine two edit distances to find a good compromise between the recognition rate and the post-processing time. The combination of a modified distance adapted to handwriting, and a simple basic distance permit to explore the list of hypotheses.

Results

The experimental results illustrated the interest of the study of the lexical postprocessing optimization in particular during the treatment of large lexicons. We note for the first approach (generic shape) a significant profit over the computing times and for the second (not supervised classification of words) a reduced memory requierement. The second approach offers optimization possibilities, in particular while working on the choice of a better representation space of the words.

7 Contracts and Grants involving industry

7.1 Integration into a Smart Phone Device of RESIFCar : an On-line Handwritten Character Recognition System.

Contributed by: Eric Anquetil.

- Company : PurpleLabs
- Software licence : INSA 3014

The collaboration with PurpleLabs company has started in 2001. PurpleLabs offers fully integrated solutions for handset designs (http://www.purplelabs.com). The ultimate goal of this collaboration is to implement an accurate handwriting recognizer into mobile devices with limited computing and memory resources. RESIFCar software (cf. section 5.1), developed by IMADOC Project, has been optimized to achieve this challenge. The hierarchical fuzzy modeling of RESIFCar offered a compact and robust knowledge representation. Moreover the decision process that is based on adapted fuzzy inference systems has allowed a reduction of the computing requirements without involving a decreasing in the performance.

Today several mobile phones design by PurpleLabs integrate RESIFCar software so that user can communicate directly with his smart phone by writing on the tactile screen of the mobile device using his personal handwriting.

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

Contributed by: Eric Anquetil, Sabine Carbonel.

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

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

- The first axe consists in the study of the integration possibilities of the on-line handwriting recognition systems RESIFCar and RESIFMot (cf. section 5.1) into smartphone device;
- The second axe focuses on the task of lexicon post-processing (cf. section 6.7) to deal with large word vocabulary in the context of limited resources. The aim of this work is to explore the combination of different approaches to reduce time computing and memory resources involved in lexicon post-processing.

7.3 Access to Handwritten Archives Documents

Contributed by: Bertrand Coüasnon, Eric Anquetil, Laurent Pasquer, Jean Camillerapp, Ivan Leplumey, Grégory Maitrallain, Fabrice Boyer.

- Partner : *Conseil Général des Yvelines, Archives des Yvelines*
- Contract : INSA 3009, INRIA 1 02 C 0602 00 31406 01 1
- Période : July 2002 to October 2007.

This research project is about making handwritten archives documents accessible to public. In collaboration with the *Archives des Yvelines*, Imadoc is working on three axes:

- document structure recognition and handwriting recognition to automatically produce annotations. These annotations offer the possibility to access by handwritten content to archives documents;

- a web platform to manage images, automatic annotations and collective annotations made by readers;
- a pen-based interface build on this platform to offer a new way of interacting with digital documents and paper documents.

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

8 Other grants

8.1 National initiatives

8.1.1 ACI Madonne (Ministry Grant)

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

8.1.2 RTP

Participation of Imadoc to the *Réseaux Thématiques Pluridisciplinaires* :

- RTP-DOC (STIC) :
  - AS Numérisation et valorisation des collections : Jacques André, Bertrand Coüasnon
  - AS Modèle(s) de publication sur le web : participation of Jacques André and managing of the *AS* by Irisa.

8.2 International initiatives

Guy Lorette is in charge of international relations at IFSIC, and member of the international relations group of the Rennes 1 University. It is also french deleguate at *Governing Board* of the IAPR *(International Association of Pattern Recognition)* and member their *Constitutions and Bylaws Committee.*

Guy Lorette was reviewer for a PhD dissertation at HKBU (Hong-Kong).

9 Dissemination

9.1 Leadership within scientific community

9.1.1 Editorial board

Members of the team are in editorial board of:

- *Document numérique*,

- *Les Cahiers GUTenberg*,
- Encyclopaedia *Techniques de l'ingénieur – Informatique* ;
- *Typography series* from Addison-Wesley.
- *Pattern Recognition*,
- *IJDAR–International Journal of Document Analysis and Recognition.*

9.1.2 Programme comittee

Members of the team are in Programme comittee of:

- ICDAR2003 (International Conference on Document Analysis and Recognition),
- EuroTex 2003,
- WEDELMUSIC 2003 (International Conference on Web Delivering of Music),
- WEDELMUSIC 2004 (International Conference on Web Delivering of Music),
- CHIR 2004 (Colloque Histoire de l'Informatique),
- CIFED 2004 (Conférence Internationale Francophone sur l'Ecrit et le Document),
- IWFHR'9 (International Workshop on Frontiers in Handwritting Recognition),
- IWFHR'10 (International Workshop on Frontiers in Handwritting Recognition) member of the Organising Committee.

9.1.3 Member of scientific society

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

Guy Lorette is :

- member of the scientific council of PSI (Université-INSA de Rouen) and of E3I computer laboratory (E3I-Université de Tours);
- vice president in load of international relations of GRCE;
- delegated at l'IAPR by l'AFRIF : ¡¡ Association Française de Reconnaissance et d'Interprétation de Formes ¿¿, de son CA ;
- member of the ASTI society.

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

Jean Camillerapp, Guy Lorette, Eric Anquetil et Bertrand Coüasnon take part in the activities of the society GRCE : ¡¡ Groupe de Recherche en Communication Écrite ¿¿.

9.1.4 Forum list

Jacques André has the responsability of the list *typographie* (`sympa@irisa.fr`), stored at URL : `https://www.irisa.fr/wws/info/typographie`.

9.2 University education

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

- Guy Lorette give lectures at DEA d'informatique and DESS MITIC at l'IFSIC, University of Rennes 1.
- Hélène Richy is in charge of the DESS MITIC (Méthodes Informatiques et Technologies de l'Information et de la Communication) at l'IFSIC, University of Rennes 1 and give lectures in this cursus.

9.3 Participation to conferences, seminars, invitations

- Jacques André presented invited paper at DIDAPRO (June 2003) and ATALA (Oct 2003)
- Bertrand Coüasnon presented an invited paper at MUSICNETWORK (Dec 2002)

10 Bibliography

Reference works and articles by the team

[1] J. ANDRÉ, M.-A. CHABIN (editors), *Les documents anciens*, Numéro spécial de *Document numérique*, vol. 3, n°1-2, Hermès, 1999.

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

[3] B. COÜASNON, J. CAMILLERAPP, "A way to separate knowledge from program in structured document analysis: application to optical music recognition", *in: Proceedings of the 3rd International Conference on Document Analysis and Recognition, (ICDAR'95)*, p. 1092–1097, Montréal, Canada, August 1995.

[4] V. P. D'ANDECY, J. CAMILLERAPP, I. LEPLUMEY, "Analyse de partitions musicales", *Traitement du Signal 12(6)*, 1996, p. 653–661.

[5] G. LORETTE, E. ANQUETIL, "Théorie des catastrophes, géométrie différentielle et segmentation de l'écriture cursive", *in: Actes du 4ème Colloque National sur l'Écrit et le Document, (CNED'96)*, p. 1–6, Nantes, France, July 1996.

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

[7] G. Lorette, Y. Lecourtier, "Is recognition and interpretation of handwritten texts a scene analysis problem?", *in: Proceedings of the 3rd International Workshop in Frontiers in Handwriting Recognition , (IWFHR3)*, p. 184–196, Buffalo, USA, May 1993.

[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, Éric Anquetil, G. Lorette, "Coherent Knowledge Source Integration through Perceptual Cycle Framework for Handwriting Recognition", *in: Advances in Handwriting Recognition*, S.-W. Lee (editor), World Scientific, 1999, ch. 5 in PART2 : On-Line Handwriting Recognition, p. 59–68.

[10] J. Pettier, J. Camillerapp, "Script representation by a generalized skeleton", *in: Proceedings of the 2nd International Conference on Document Analysis and Recognition, (ICDAR'93)*, p. 850–853, Tsukuba, Japan, October 1993.

Books and monographs

[11] B. Coüasnon, J.-P. Dalbéra, H. Emptoz (editors), *Numérisation et patrimoine*, *Numéro spécial de Document Numérique*, *7*, 3-4, Hermes-Lavoisier, 2003.

Doctoral dissertations and "habilitation" theses

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

Articles and book chapters

[13] J. André, M. Bovanni, T. Bouche, "La fonte Fourier", *Lettre GUTenberg*, 25, June 2003.

[14] J. André, "Numérisation et codage des caractères de livres anciens", *Document numérique 7*, 3-4, 2003, p. 127–142.

[15] B. Coüasnon, J. Camillerapp, "Accès par le contenu aux documents manuscrits d'archives numérisés", *Document Numérique*, p. 61–84.

[16] 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.

[17] 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.

Communications in congresses, etc.

[18] J. André, "Towards an inventory of ancient types and the related standardised encoding", *in: Proceedings of EuroTeX proceedings*, p. 165–169, Brest, June 2003.

[19] J. André, "Unicode et la ponctuation", *in: Actes de la journée "Typographie, ponctuation et analyse syntaxique"*, p. 1–11, ATALA, Paris, November 2003.

[20] J. André, "Histoire d'Œ, histoire d'@ – rumeurs, typographie et enseignement", *in: Actes des Journées francophones de didactique des progiciels*, p. 1–18, Hermes-Lavoisier, Paris, June 2004. to appear.

[21] J. Camillerapp, L. Pasquer, B. Coüasnon, "Indexation automatique de formulaires anciens par reconnaissance du patronyme manuscrit", *in: RFIA, Reconnaissance des Formes et Intelligence Artificielle*, Toulouse, France, 2004. to appear.

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

[23] S. Carbonnel, E. Anquetil, "Modélisation et intégration de connaissances lexicales pour le post-traitement de l'écriture manuscrite en-ligne", *in: Reconnaissance de formes et intelligence artificielle (RFIA'04)*, Toulouse, France, 2004. to appear.

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

[25] B. Coüasnon, I. Leplumey, "A Generic System for Making Archives Documents Accessible to Public", *in: ICDAR, International Conference on Document Analysis and Recognition*, Edinburgh, UK, August 2003.

[26] B. Coüasnon, "Dealing with noise in DMOS, a Generic Method for Structured Document Recognition", *in: IAPR International Workshop on Graphics Recognition*, Barcelona, Spain, July 2003.

[27] N. Ragot, E. Anquetil, "A Generic Hybrid Classifier Based on Hierarchical Fuzzy Modeling: Experiments on On-Line Handwritten Character Recognition", *in: IEEE International Conference on Document Analysis and Recognition (ICDAR'03)*, *2*, p. 963–967, Edinburgh, August 2003.

Miscellaneous

[28] B. Coüasnon, "Consultation et annotation collective de documents anciens", *in: Rapport de fin de projet Mégalis*, January 2003.