Project-Team IntuiDoc

*Intuitive user interaction for document*

*Rennes*

*Activity Report*

*2012*
Contents

1 Team 5

2 Overall Objectives 5

3 Scientific Topics 6
  3.1 Bidimensional visual languages integrating the user-interaction concept 6
  3.2 Combining points of view for image interpretation 6
  3.3 Incremental learning and evolving fuzzy classifiers 7
  3.4 Pen- and Gesture-Based Interaction 7

4 Application Domains 8
  4.1 Paper document analysis: batch or interactive interpretation 8
  4.2 Evolving pen- and touch- based interaction 9
  4.3 Handwriting, hand-drawn symbol, and gesture recognition systems 9

5 Software 10
  5.1 RESIF: Handwriting recognition by hierarchical fuzzy inference systems 10
  5.2 EVOLVE++ / EVOLVE TOUCH: Evolving recognition engine 10
  5.3 DALI: a framework for the design of pen-based document sketching systems 10
  5.4 IMISKETCH: interactive off-line sketches recognition 11
  5.5 DocRead: an automatic generator of recognition systems on structured documents 11
  5.6 Precoce: Library to extract visual indices 12
  5.7 LIMO: an isolated handwriting word recognizer 12
  5.8 iLib: a feature extraction library 13

6 New Results 13
  6.1 Heterogeneous Baseline Feature Set (HBF49) for the representation of hand-drawn symbols 13
  6.2 Evolving Classification System for Handwritten Gesture Recognition 14
  6.3 Evolve-Touch: Personalized gesture-based interaction for multitouch surfaces 14
  6.4 Architectural floor plan recognition 15
  6.5 Gestures’ learning on interactive tablets 16
  6.6 Collection of a new set of gestures obtained by real application 17
  6.7 Use of perceptive vision for on-line flowchart recognition 17
  6.8 Detection of text lines with a particle filter 18
  6.9 Contextual and assisted framework for document interpretation 18
  6.10 Experiments on lexicon size variations 19

7 Contracts and Grants with Industry 20
  7.1 SCRIPT and GO 20
  7.2 MobiSketch ANR project 20
  7.3 Evolve project: European Regional Development Fund (FEDER) 21
  7.4 Cassidian - PEA DGA project 21

8 Other Grants and Activities 21
  8.1 National initiatives 21
     8.1.1 Experiments on interactive gestural commands using a digital pen or touch screen 21
     8.1.2 Igidoc - Brittany region Grant (ARED) 22
  8.2 International initiatives 22
1 Team

Permanent Members

Eric Anquetil [Team Leader, Professor, Insa, HDR]
Jean Camillerapp [Emeritus Professor, Insa, HDR]
Bertrand Coùasnon [Associate Professor, Insa, HDR]
Ivan Leplume [Associate Professor, Insa]
Aurélie Lemaitre [Associate Professor, Uni. Rennes 2]
Yann Ricquebourg [Associate Professor, Insa]

Fixed-term positions

Abdullah Almaksour [Rennes1 Lecturer until aug 2012, Insa Res. Eng. sept to nov 2012]
Manuel Bouillon [Insa PhD student, Menrt grant from oct 2012]
Cérés Carton [Insa PhD student, Brittany region/INSA grant from dec 2012]
Joseph Chazalon [Insa PhD student, Rennes1 Lecturer until aug, Insa Res.Eng. from sep 2012]
Adrien Delaye [Research Engineer, Insa until feb 2012]
Achraf Ghorbel [Insa PhD student, Insa grant until aug, Lecturer Rennes1 from sep 2012]
Peiyu Li [Insa PhD student, Menrt grant]
Baptiste Poirriez [Insa Research Engineer from apr 2012]
Ney Renau-Ferrer [Uni. Rennes 2 Lecturer until aug 2012]
Grégoire Richard [Insa Research Engineer from mar 2012]

Administrative Assistant

Nadia Saint-Pierre [Inria]

2 Overall Objectives

The IntuiDoc project is an evolution from the previous Imadoc team.

The project IntuiDoc carries out research on the Human-Document Interaction by associating the issues of fields of Pattern Recognition and Human-Machine Interaction. The objective is to find solutions to design an efficient, robust and intuitive Human-Document communication based on a continuum between the document under its paper format and the same document as its digital interpreted form (figure I).

The project lies to the domain convergence of the analysis and recognition of complex handwritten documents (manuscript, printed paper, form, graph, sketches, composite documents, etc.), and of the pen- and gesture-based interaction for touch or pen capable devices such as smartphones, tablets and multitouch surfaces.
IntuiDoc aims to investigate new approaches to analyze, recognize, compose or interpret complex documents by introducing an interactive process. The originality is to explicitly integrate the user in the process of analysis and decision making, both in the recognition of complex documents and in the "on-the-fly" interpretation of on-line hand-drawn sketches. This strategy aims to address the limitations of current approaches that are based on non-interactive treatments. The concept is to strengthen the decision processes based on implicit or explicit correction of a final user to avoid the propagation of errors of interpretation throughout the analysis.

Figure 1: Continuum from paper document to digital interpreted document

3 Scientific Topics

3.1 Bidimensional visual languages integrating the user-interaction concept

The knowledge associated with the structure of the documents are modeled using bidimensional grammars and visual languages, and by studying new approaches based on constraint multiset grammars. The aim is to design generic methods for structured document analysis and composition.

The introduction of the user in structured document recognition process requires to model this interaction to be able to describe what are the possible interactions for the user, in association with the structural modeling of the document (bidimensional grammars). With the introduction of the user in the analysis process we need to control the requests for the user. If the user interaction occurs on one isolated document, the interaction can be synchronous. On the other hand, during the treatment of a large collections of document, the challenge consists for the analyzer to collect requests and postpone the interaction with the user, to build an asynchronous interaction. To elaborate evolving systems for structured document recognition we explore grammatical inference. This objective, that is already a real challenge for mono-dimensional grammars, is very complex for bidimensional grammars. Our strategy is to perform this inference with the help of user interaction and by focusing inference for the physical structure analysis.

3.2 Combining points of view for image interpretation

Combining several ways of interpreting the content of a document can improve its recognition [5]. Thus we study some mechanisms of knowledge fusion to combine the results various document analysis
techniques that are usually studied separately \[10\]. The knowledge fusion must be as flexible as possible, and if necessary in an asynchronous way.

First, we study different levels of analysis of the image: the analysis of multiresolution images enables to inspire from the human perceptive vision that detects salient objects in a document without specific knowledge.

Secondly, we explore some low-level image processing techniques to extract some local primitives: line segment extraction with Kalman filtering, use of particle filter for text line extraction, word spotting using interest point detector, texture analysis... We also introduce some knowledge coming from other processes, such as the result of commercial OCR. All these kinds of contents have to be combined, depending on the studied kind of document, to process the best results.

At last, the originality of our work is to combine a structural analysis with the introduction of statistical data. This combination enables to exploit the expression power of a structural analysis while benefiting from the large range of statistical approaches.

3.3 Incremental learning and evolving fuzzy classifiers

To develop a robust and contextual recognition of the elements which form a printed or handwritten document, we design hybrid recognition methods (statistical / structural) which rely in particular on the theory of fuzzy logic to manage the inaccuracy of handwritten strokes.

Traditionally, a classification system is trained using a learning dataset under the supervision of an expert that controls and optimizes the learning process. The system performance is fundamentally related to the learning algorithm and the learning dataset. The classification system is delivered to the final user to be used in real applicative contexts. Typically, no learning algorithms are available at the user side. The main weakness in the above-mentioned conception paradigm is that the knowledge base is constrained by the learning dataset available on the expert side and cannot be extended by the data provided on the user side. These drawbacks increase the need for new type of classification systems that can learn, adapt and evolve in a lifelong continuous manner. For example, in the framework of on the fly composition of documents, it is interesting to allow user to choose its own set of gestures to assign them to different symbols or commands. In the context of interactive document recognition, it is essential to learn from the user interactions, the unknown symbols by integrating dynamically new classes of these symbols in the recognition system.

In evolving systems, incremental learning algorithms are used to learn from the data samples provided by the user after sending a validation or a correction signal in order to confirm or change the label suggested by the classifier. Contrary to the traditional paradigm, there is no separation between the learning phase and the operation phase in evolving classification systems. One of the key features in evolving classifiers is that incoming samples may bring in new unseen classes that are learned by the classifier without destroying its knowledge base or forgetting the existing classes. IntuiDoc designs new incremental approaches for the learning of classification models based on first-order Takagi-Sugeno fuzzy inference systems \[1\]. This approach includes, on the one hand, the adaptation of consequences of the fuzzy rules using the recursive least-squares method, and, on the other hand, an incremental learning of the antecedent of these rules according to the evolution of data density in the input space.

3.4 Pen- and Gesture-Based Interaction

Accuracy and robustness of developed systems are key elements for the User Acceptance. To cope with these requirements, the recognition systems have to be adjustable during their use in the application. We design an evolving recognition engine with an online, incremental, and lifelong learning process.

Portable, touch or pen capable devices such as smartphones, tablets and multitouch surfaces become more and more ubiquitous. Such devices allow for natural interaction via handwriting and gestures. For this sort of usage we designed personalisable gesture recognition engine. We aim to provide the user
the possibility to define his own gesture library for various activities. For instance, we work on intuitive
mechanism to get user feedback on recognizer answers, which allows the latter to continuously enhance
its performance. We also design direct object manipulation such as rotation, zoom or translation...

Some complex applications need a lot of commands thus the definition of gesture commands and
the gesture memorization becomes an important task. Our objective is to obtain natural fluid gestures
and to help the user to learn as quickly as possible. The key point is to obtain a complete and
customizable set of gestural commands to interact with applications: this induces to be able to design
an auto-evolutional gesture recognition system and, in the same time, a framework to help the user to
memorize his gestural command set.

The main approaches of gesture learning help are based on Marking Menus which propose two ways
of utilization: a novice mode where the user has menus displayed to help him to finalize his gesture
and an expert mode where he only needs to draw the required gesture and the recognizer will try to
understand which command is invoked. All these approaches help the users to memorize gestures by
making them practice drawing. Obviously, the final form of gestures strongly depends on the menu
ergonomics.

In this way, we design the Customizable Gesture Menus which combine the advantages of marking
menus and personalization capability to give the user an optimal memorization help with customizable
set of gestures.

In this scientific area, it is vital to take into account the user and therefore the uses. This is why
IntuiDoc actively collaborates with the multidisciplinary research platform LOUSTIC to assist the
experiments on gestural commands and on the learning strategies which explicitly involves the final
user.

4 Application Domains

4.1 Paper document analysis: batch or interactive interpretation

Applications of our research in the context of paper document analysis are very wide. Indeed, the
generic approaches we have developed (the DMOS-P and IMISKETCH method) allow a quite fast
adaptation to new types of documents. Documents we have already worked on are:

- Architectural floor plan interpretation;
- Orchestra scores with polyphonic staves;
- Mathematical formulae;
- Table structures, forms with recognition of the hierarchical organization;
- Archives documents: more or less structured old forms [2].

With the help of handwriting recognition, those 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 indices and annotations for an automatic access by content to docu-
  ments;
- Detection of specific areas in a document to improve its access.
4.2 Evolving pen- and touch- based interaction

One target application is the use of online handwritten gesture classifiers to facilitate user interactions on pen- and touch- based interfaces like tablet, smartphones, whiteboards, multitouch surfaces, etc. The challenge is to develop applications for these devices with personalization capacity so that user can define his own set of gesture and add new gestures at any moment. Gestures, Symbols or letters can be drawn differently from one user to another, and users may want to add or remove gestures, as long as they use the application. Moreover, users would often change progressively the manner by which they draw gestures. Novice users start drawing carefully and slowly their gestures, while they do them in a more fluid and rapid manner as they become expert. The classifier hence needs to evolve and follow the changes in the data flow. If most users will use a common subset of gestures, each user will need some specific gestures classes for his own usage but that others won’t use. In addition, classifier usage may change with time, and the end user may need to add, remove or change gestures classes to fit his needs. That is why the classifier needs to be customisable by end users.

Several applications are developed and experimented using gesture commands, intuitive editing commands, handwritten letter or symbol recognition:

- touch- or pen- based composition: musical score, graph, architectural floor plan...
- annotating documents (photo, archive) for indexing (digital library),
- personalizing Gesture-Based Interaction for Touch-Sensitive Screens.

4.3 Handwriting, hand-drawn symbol, and gesture recognition systems

Recently, there has been a new increase in the applications of handwriting recognition [16] in the domains of automatic processing of paper documents (off-line recognition) as well as in the new modality of man-machine interaction (graphical gesture recognition), 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:

- smartphone (Iphone, Galaxy note...),
- tabletPCs (Surface, Ipda...),
- multitouch table (Microsoft PixelSense).

As future user-centered interaction systems will offer more flexibility and give more liberty to the final user(ability to choose his personal gestures and customize the system), the recognition methods will have to be extensible and to deal with new unpredictable symbols.

To date, the “ResifCar” and “Evolve Touch” software were embedded in such devices (cf. subsections 5.1 5.2).
5 Software

All the presented softwares have been deposit in APP. More details on those softwares can be found on Intuidoc web site (http://www.irisa.fr/intuidoc).

5.1 RESIF: Handwriting recognition by hierarchical fuzzy inference systems

Contact: 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 five version. Through industrial collaborations, RESIFCar has been successfully integrated into mobile devices (smartphones) which are characterized by their limited computing and memory resources.

ResifCar has been integrated in the educational software Toutaki of Evodia/Script&Go Company. This Tablet PC software helps the young children to learn how to write. Toutaki has been licensed to HITACHI Company to be embedded to their Electronic Whiteboards.

5.2 EVOLVE++ / EVOLVE TOUCH: Evolving recognition engine

Contact: Eric Anquetil

Keywords: Incremental recognition, Evolving system, Gestures and Symbols Recognition.

Evolve++ is an evolving recognition engine, that can be trained incrementally, starting from few data samples. Evolve++ is based on a fuzzy inference systems that learn incrementally and cope with class adding.

Evolve-Touch is a derived software based on Evolve++ for the application domain of graphical gesture recognition for multitouch devices. Evolve-Touch offer a complete framework to allow user to manage and customize his gesture sets for different application contexts in simple and user-friendly manner. An intuitive mechanism is adopted to get user feedback on recognizer answers, which allows the latter to continuously enhance its performance. To demonstrate the main features of Evolve-Touch system, a showcase application is presented in this video: http://youtu.be/qOzIY6uYF8. This work is supported by a European Regional Development Fund (FEDER), and protected by a European patent (N°1258803)

5.3 DALI: a framework for the design of pen-based document sketching systems

Contact: Eric Anquetil
Keywords: Sketch recognition, pen-based interaction, visual language theory, industrial transfer.

DALI is a framework for the interpretation of hand-drawn sketches drawn on tablet PCs. The first property of the DALI method is its genericity, which means that it can be used to design pen-based software to sketch various natures of documents. It is based on the visual language and grammar theory that makes it possible to model bidimensional symbols and documents [6, 7]. DALI interprets the user strokes on-the-fly, directly during the design of the document; it means that each time the user draws a stroke, the system analyses it and produces a visual feedback, showing how it is interpreted.

This way, the user is an actor of the interpretation process, because he can progressively correct the errors of the system. Thus, the interpretation process can rely on the information given by the user to better interpret the following strokes. The coupling of these two properties increases significantly the efficiency and the robustness of the sketch interpretation process.

The DALI method has been used to design several pen-based prototypes, for instance for the sketching of musical scores, electrical sketches, UML class diagrams, architectural floor plans, etc.

It has been transferred to the Script&Go society, which led to the design of Script&Go Electrical Sketches for electrical sketches and Script&Go Plans for architectural floor plan sketching. These softwares are today commercialized and used daily by hundreds of technicians in France. Script&Go Electrical Sketches has been rewarded with the "Trophées de l’innovation" 2008 for uses, applications and communicating solutions for enterprises", in the category named "Solutions Métiers".

5.4 IMISKETCH: interactive off-line sketches recognition

Contact: Eric Anquetil

Keywords: interactive off-line Recognition, sketches, 2D architectural floor plan.

IMISketch is a new generic method for interactive interpretation of image of sketches (structured document). The goal is the mapping of technical paper document to numerical ones. IMISketch has been used to deal with off-line handwritten 2D architectural floor plan recognition [11].

5.5 DocRead: an automatic generator of recognition systems on structured documents

Contact: 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-P method [3]. 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.

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 more than 480,000 pages of the Archives de la Mayenne and Archives des Yvelines;

• 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.

• BanglaRead: a software for extracting headline in Bangla script. This extraction represent a pre-processing tool for handwriting recognition. This work has been realized in collaboration with University of Kolkata and applied on 1922 words from 26 writers.

• FlowRead: a software for on-line flow-chart segmentation and structure recognition.

5.6 Precoce: Library to extract visual indices

Contact: Jean Camillerapp

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

It is the library on which the DocRead software described in the section 5.5 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 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.

5.7 LIMO : an isolated handwriting word recognizer

Contact: Bertrand Coüasnon

Keywords: Offline Handwriting Recognition, HMM, SVM, rejection.

The LIMO software purposes is to realize isolated handwriting word recognition. It takes as input an image of the word and produces a list of N-best hypotheses. It can be used to recognize words belonging to a finite dictionary or an open dictionary (with n-gram language model). The recognizer has a two stages architecture:

• A HMM-based recognition that generates a list of N-best hypotheses,
• A SVM-based verification to rescore the N-best hypotheses using SVM character classifiers and to select the best one.

The recognizer also has an integrated rejection feature which combines the HMM and SVM recognition scores to accept or reject the analyzed sample.

5.8 iLib: a feature extraction library

Contact: Yann Ricquebourg

Keywords: Feature extraction.

Concerning the studies of efficient classification and recognition methods, the 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.

Those main purposes have been achieved and the library now proposes in a generic and stable way “classical” features often cited as required by common recognition systems (as surface, gravity center, curvature...), as well as more original or promising characteristic functions as Zernike moments, (concerning plain pixel images), elliptic Fourier descriptors (concerning contour curves), or holes and hollows determination based on chanfrein distance and convex envelope.

Figure 2: Some of the iLib functionalities: progressive decomposition using elliptic Fourier descriptors, using Zernike moments, holes and hollows determination, zoning

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

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

6 New Results

6.1 Heterogeneous Baseline Feature Set (HBF49) for the representation of hand-drawn symbols

Participants: Adrien Delaye, Eric Anquetil.
As the rise of pen-enabled interfaces is accompanied with an increased number of techniques for recognition of pen-based input, recent trends in symbol recognition show an escalation in systems complexity (number of features, classifiers combination) or the over-specialization of systems to specific datasets or applications. We introduce in [12] the Heterogeneous Baseline Feature Set (HBF49), a unique set of features for the representation of hand-drawn symbols to be used by symbol recognition systems. This set of 49 simple features is able to handle a large diversity of symbols in various experimental contexts, in particular the HBF49 is very well suited to be used to recognize user-defined gestural commands [6.5] [7.3]. We demonstrated that using off-the-shelf statistical classifiers, the HBF49 representation performs comparably or better than state-of-the-art results reported on eight databases of hand-drawn symbols. An original effort has been made for guaranteeing transparency of features design and reproducibility of experiments to enable and spur re-use of this feature set for classification system comparison among the machine learning community.

6.2 Evolving Classification System for Handwritten Gesture Recognition

Participants: Manuel Bouillon, Eric Anquetil, Abdullah Almaksour.

Evolving classification systems have appeared in the last decade to meet the need for recognizers that work in changing environments. They use incremental learning to adapt to the data flow and cope with class adding (or removal) at runtime. In this context, we study the decremental learning of such an evolving recognizer, based on a fuzzy inference systems. The aim of decremental learning is twofold. First, to maintain the system learning capacity over time, and second, to forget obsolete data and focus the learning process on current environment.

The target application of this work is the use of on-line handwriting gesture classifiers to facilitate interactions with computers using pen-based interfaces like smartphones, tablet computers, whiteboards, etc. Gestures can be drawn differently from one user to another, and users may want to add or remove gestures, as long as they use the application. Moreover, user would often change progressively the manner by which he draws gestures. A novice user starts drawing carefully and slowly his gestures, while he does them in more fluid manner when he becomes expert. To cope with these requirements, forgetting must be used to increase system reactivity and performance in such a dynamic environment.

We extend our evolving classification system Evolve by integrating decremental learning techniques. Our baseline Evolve is an on-line recognizer based on an evolving fuzzy inference system. We have developed two new approaches to introduce decremental learning in the optimization of fuzzy rules conclusions, both relying on a sliding window of data samples. The first one uses this window to unlearn completely old data – so-called downdating – whereas the second strategy uses the window to cut down old data weight in the optimization process – so-called (directional) forgetting.

We have shown that forgetting is necessary to maintain the system learning capacity over time and to correctly adapt to concept drifts, which makes decremental learning essential for the life-time use of an evolving classification system. These results have lead to a conference paper [17].

6.3 Evolve-Touch: Personalized gesture-based interaction for multitouch surfaces

Participants: Abdullah Almaksour, Eric Anquetil, Grégoire Richard.

We present a novel gesture-based interaction system for multitouch surfaces. The main feature of the presented system is its personalization capacity so that user can define his own set of gesture and add new gestures at any moment. The system is based on Evolve (see section 6.2), an evolving gesture recognition engine that can learn incrementally, starting from few data samples. We present a complete mechanism to allow user to manage his gesture sets for different application contexts in simple and
user-friendly manner. An intuitive mechanism is adopted to get user feedback on recognizer answers, which allows the latter to continuously enhance its performance. Moreover, we integrate into the global system a multi-touch object manipulation modality with a simple mode switching. We developed an application, named Evolve-showcase (video available at http://www.youtube.com/user/Intuidoc), to demonstrate the main features of Evolve-Touch. This work is supported by a European Regional Development Fund (FEDER), and protected by a European patent (Ref 1258803, see section 7.3).

6.4 Archituctural floor plan recognition

Participants: Achraf Ghorbel, Eric Anquetil, Aurélie Lemaitre.

In the context of the MobiSketch ANR research project (see section 7.2), we are working on the recognition of technical paper documents to digital ones. The principle is to offer a complete and homogeneous solution to unify paper document recognition and pen-based sketch interpretation (for instance: with Tablet PC). Our current work is to develop a new generic method (referred as IMISketch) for an interactive interpretation of sketches to avoid a fastidious verification phase [13].

We have proposed an optimization strategy to minimize the combinatory [23] [22]. We have proposed, also, a new hybrid strategy for exploration in which the recognition process alternates between a breadth-first and depth-first exploration. The strategy is totally driven by the grammatical description of the document [24].

The figure 4 illustrates an example of architectural floor plan interpretation.

Future work will focus on extending the experimental results on large image databases containing printed and vectored document architectural plans... We will also validate the criteria of acceptability and usability of the system by doing usage tests that will be conducted in collaboration with experts from the laboratory uses Loustic (http://www.loustic.net/rennes).

Figure 3: (a) gesture control panel (b)(c) online user feedback mechanism (d)(e) simultaneous scaling-rotation action (f) object suppression as scaling special case.
6.5 Gestures’ learning on interactive tablets

Participants: Peiyu Li, Grégory Richard, Eric Anquetil, Ney Renau-Ferrer.

Since the end of last year, we observed constraints and shortages of Continuous Marking Menus (CMM), we created a new concept: Semi-Customizable Gesture Commands (SCGC). This new type of commands allow user to associat his own gesture to predominant command of a family. And gestures for other commands are generated automatically (See RA2011, Section 6.4). To see how this approach works to help the learning of gestures and its efficiency, we made some experimentation. 9 users passed 3 different test each of them. Each test contained 18 gestures to be memorized. For test1, all gestures were predefined. For test2, user needed to choose himself all gestures. For test3, user chose one gesture for a predominant command per family (SCGC approach) and the rest was automatically generated. This experimentation proved that it is easier for user to learn gestures defined by themselves than predefined gestures (test2>test1). And it is easier for user to learn SCGC commands then simple self-defined gestures (test3>test2). These results are proved to be significant by tests ANOVA. These results led to publication [25].

In next step, we changed the interface of the test application for a better presentation. To make the SCGC commands easier to be manipulated and to be memorized, we made some variations on SCGC commands’ form. User needed to define the root gesture of a family and we added automatically 3 different final directions (down, right and left) to create gestures for the 3 commands of this family. The new method is illustrated on (figure 5). We also made new experimentation for this variation, with the coorporation of test platform Loustic. The result of this work is presented in publication [26].

![Figure 5: Semi-Customizable Gestural Commands (new approach)](image-url)
6.6 Collection of a new set of gestures obtained by real application

**Participants:** Ney Renau-Ferrer, PeiYu Li, Eric Anquetil, Adrien Delaye.

We created a simulated application to test *Semi-Customizable Gesture Commands (SCGC)* in real-life situation (see section 6.5). With this experimentation, we obtained a new database of realistic pen-based gestures for evaluation of recognition systems in pen-enabled interfaces (*ILGDB*). This database contains a large number of unconstrained user-defined gestures (figure 6). Thanks to its diversity, it can be served as a tool for benchmarking of gesture recognition systems. This result leads to the publication [27] and it is available at http://www.irisa.fr/intuidoc/ILGDB.html.

Figure 6: Gestures from ILGDB database

6.7 Use of perceptive vision for on-line flowchart recognition

**Participants:** Aurélie Lemaitre, Jean Camillerapp, Bertrand Coiasnon.

We have improved our work for handwritten on-line flowchart recognition (figure 7). It leads to an extended paper in LNCS [14].

Figure 7: Example of handwritten flowchart

We have proposed a grammatical description of this kind of document. This is particularly adapted to the bi-dimensional and structured properties of the flowcharts. In our implementation, we have shown that the existing off-line method, DMOS, could be used for the analysis of on-line documents.
We have proposed to combine two kinds of primitives: line segments and on-line strokes for the grammatical description. We have shown that it enables to simplify the problem of the text/symbol stroke separation, which is often met in the literature. We have validated our work on an open database.

The on-going work aims at showing that the use of several points of view of the same image can improve its recognition. It is similar to the mechanisms of perceptive vision used by the human vision. Our goal is to exploit the mechanisms of perceptive vision for the flowchart recognition.

### 6.8 Detection of text lines with a particle filter

**Participant:** Jean Camillerapp.

Detection of text lines is an important topic in the community of document processing. In the team, we have proposed a solution using perceptive vision \cite{LCC11}. It is based on detecting, by a Kalman filter, straight lines in images at very low resolution (20 dpi). This low resolution involves taking decisions with few pixels, which causes a certain sensitivity to noise.

Presently, we are still using perceptive vision, but by building a blurred image with a good resolution (100 to 150 dpi) (figure 8 (a)). This limitates the effects of quantization. Then, we localize lines in this image using a particle filters. A particle filter is a sophisticated model estimation technique based on simulation \cite{DJ08}. In the blurred image, the particles are spread from left to right with a vertical offset randomly chosen. The confidence of each particle is modified by the brightness of the image at its position. The more an image is darker, the more the confidence increases. Particles with a low confidence are periodically removed, while particles with a high confidence are duplicated (figure 8 (b)).

![Blurred image](image1)

![Propagation of particles](image2)

![Detected lines](image3)

Figure 8: Detection of text lines with a particle filter

The detected lines are formed by the average trajectory of a set of particles (figure 8 (c)). The results appear better than those obtained by the previous method. In the future work, we will focus on the choice of the initial position of a filter as well as the termination conditions.

### 6.9 Contextual and assisted framework for document interpretation

**Participants:** Joseph Chazalon, Bertrand Couësnon, Aurélie Lemaître.

*Fonds*, also called historical document collections, are important amounts of digitized documents which are difficult to interpret automatically: usual approaches require a lot of work during design, but do not manage to avoid producing many errors which have to be corrected after processing. To

\[18\]


\[DJ08\] A. Doucet, A. M. Johansen, “A Tutorial on Particle Filtering and Smoothing”, research report, Department of Statistic, University of British Columbia, USA, 2008.
cope with those limitations, our work aimed at improving the interpretation process by making use of information extracted from the fonds, or provided by human operators, while keeping a page by page processing.

During the past few years, we focused our efforts on the creation of a unifying framework which could enable the system designer to make use of both contextual and human-provided information during page interpretation. In our recent work, we managed to build a single foundation for various interaction forms, as presented in [19, 20].

Those new tools give the ability to easily transform an existing interpretation system and enable it to make use of contextual or human-provided information during page interpretation. Thanks to an extension of the description language of page structures used to generate interpretation programs, the system designer is now capable of expressing new properties about page contents, and, automatically, the implementation we proposed enables the generated system to:

- handle interpretation errors in the event of unexpected contents;
- ask for missing information;
- reintegrate and make use of external knowledge.

A global iterative mechanism permits to progressively brings contextual information, in an asynchronous way, to the later process, and improves its interpretation[21].

We were able to precisely formalize the method we proposed, and summed all our contributions in a thesis, whose defense is planned in January 2013.

This work is done in collaboration with the Archives Départementales des Yvelines, and funded by the Conseil Général des Yvelines.

6.10 Experiments on lexicon size variations

Participants: Yann Ricquebourg, Bertrand Coüasnon,

Handwriting recognition applications can take benefit of system quite tolerant and robust to situations slightly different from the ideal ones where the training was performed. In this case, having good rejection possibility is a strong advantage, especially in application cases where humans are integrated in the process, which become usual currently: proposing only a relevant subset of items rejected by the system is necessary to be reasonable.

We already had a system proposing a solution to reach good performance, by combining a first classifier, here HMM, with SVM to reinforce the global results and reach rejection possibilities. Then this amelioration of the results brought by the addition of SVM, and the good proportion of rejected items, have been confirmed by a large experimental campaign[RCG13] testing the recognition performance progressively:

- with larger lexicons (thus introducing ambiguities): the association of a second classifier helped to maintain a stable good performance all the same, compared to the ideal case with a lexicon just as needed.
- with smaller lexicons (thus introducing many unacceptable items limited by the lexicon): once gain the association of the second classifier helped to maintain a good performance since its rejection capabilities managed to refuse a lot of item to keep the authorized ones.

Figure 9: Lexicons relative performances (LPFR) for a constant Error Rate of 1% (left), 5% (center) and 10% (right) compared for systems S1 (HMM with simple threshold rejection) and S2 (HMM + SVM with multiple thresholds optimized from dynamic programming).

7 Contracts and Grants with Industry

7.1 SCRIPT and GO

Participant: Eric Anquetil.

- Partners: SCRIPT&GO company
- Contract: INSA

The IntuiDoc team has close links with the SCRIPT&GO company for transferring its industrial research results. This partnership is now based on more than six licensing agreements. They cover various technologies of the Intuiodoc team such as handwriting recognition, document analysis and on-the-fly diagrams and plans recognition. Since the creation of SCRIPT&GO companies, several technologies (Resif, Dali, Evolve) have been transferred indirectly, to companies such as Thales, Apave, Hitachi...

This partnership is also supported by several collaborative projects: in particular the ANR project "MobiSketch" (http://mobisketch.irisa.fr/) which is labeled by the "Images et Reseaux" cluster.

7.2 MobiSketch ANR project

Participants: Eric Anquetil, Achraf Ghorbel, Aurélie Lemaitre, Bertrand Coïssnon.

- Partners: SCRIPT&GO company, CRPCC Rennes 2 University
- Contract: INSA (leadership of the project)
- 42 months (2009-2013).

MobiSketch (http://mobisketch.irisa.fr/) is an ANR project (November 2009) from the "content and interaction" programme. It aims at developing generic pen-based software for structured document analysis, design and editing in a mobility context (on the field). The project deals with technical documents such as architectural floor plans. The principle is to offer a complete and homogenous solution from paper document recognition to pen-based document sketching with Table PC. The originality of the project is to explicitly integrate the user in a uniform assessment process to both recognize a paper document and hand-drawn design/edit a new document. To achieve this, we will focus both on new document analysis processes based on visual language and grammar theory but also on the usages to design a new "man-document" interaction based on pen or gesture interaction. The end of the MobiSketch project is fixed to May 2013. Mobisketch has achieved the design of IMISketch framework that has been used to deal with off-line handwritten 2D architectural floor plan recognition [11].
7.3 Evolve project: European Regional Development Fund (FEDER)

Participants: Eric Anquetil, Grégoire Richard.

- Contract: INSA (leadership of the project)
- Evolve-Touch is protected by a European patent (Ref 1258803).
- Industrial transfers: INSIDE software and SCRIPT&GO compagnies.

The goal of the European Regional Development Fund (FEDER) project "Evolve" is to accelerate the industrial transfer of "Evolve" technology [1]. This project achieves this year the design of Evolve-Touch framework, a derived software based on Evolve technology: the application domain is graphical gesture recognition for multitouch devices. Evolve-Touch offers a complete framework to allow user to manage and customize his gesture sets for different application contexts in simple and user-friendly manner. To demonstrate the main features of Evolve-Touch system, a showcase application is presented in this video: http://youtu.be/qOx4IY6uYf8. This work has been protected by a European patent (Ref 1258803). "Evolve-Touch" has been transferred to INSIDE software and SCRIPT&GO compagnies.

7.4 Cassidian - PEA DGA project

Participants: Bertrand Coüasnon, Aurélie Lemaitre, Baptiste Poirriez, Yann Ricquebourg.

- Partners: Cassidian (an EADS company)
- Contract: INSA
- 31 months (2012-2014)

This project is done in the context of a PEA (Programme d’Etude Amont - upstream study program) from the DGA (French Ministry of Defense) which involve Elda, LNE, and different research partners (A2iA, Irisa/Insa, Lip6, Litis) leaded by Cassidian, an EADS company, to build a prototype of an automatic document recognition system.

We work on the structure recognition of heterogeneous and multi-language documents, which can be handwritten, printed or mixed. The objective is to define a document structure recognition system able to deal with this large variety of documents, to extract and identify the different types of elements found on those documents. This is done by building a generic description of the document structure with the DMOS-P method. Two international competitions will be organized in this PEA, to evaluate the different modules of document recognition.

8 Other Grants and Activities

8.1 National initiatives

8.1.1 Experiments on interactive gestural commands using a digital pen or touch screen.

Participants: Eric Anquetil, PeiYu Li.

- Partner: LOUSTIC laboratory of Rennes (laboratoire d’observation des usages des technologies de l’information et de la communication)
This project has been labeled by the LOUSTIC laboratory of Rennes (laboratoire d’observation des usages des technologies de l’information et de la communication). The objective is to make several experiments on interactive gesture commands using a digital pen or touch screen. This new mode of interaction is designed to offer the user the ability to define at the same time by a graphic gesture: the type of command, the area targeted by the command and its parameters. Defining, modeling and recognizing these gesture commands is a challenge considering the following objectives: to integrate a large family of gestures within the same application, allowing the user to select and configure its control actions, make the gesture recognition engine self-evolving to suit the current user (incremental learning). To design efficient interactive gestural commands, it is necessary to place the user at the center of the development process by implementing several experimental tests.

8.1.2 Igidoc - Brittany region Grant (ARED)

Participants: Cérès Carton, Bertrand Coïasnon, Aurélie Lemaitre.

- Partners: Brittany region
- Contract: INSA
- 36 months (2012-2015)

This project funds (Brittany region/Insa grant) the Ph.D. of Cérès Carton on Interactive Grammatical Inference in Visual Grammars for Structured Documents Recognition. The objective is to simplify the definition of knowledge on a new kind of document, by integrating grammatical inference. It is a very challenging task in mono-dimensional grammars, and it is even more in bi-dimensional grammars like we have for document analysis. However, to reduce the complexity of this task, we will work on grammatical inference with a human interaction and we will in a first step, focus inference more on the physical structure, which is easier to do than on the logical part.

8.2 International initiatives

8.2.1 Québec University: Pen-based interfaces for collaborative environments

Participants: Eric Anquetil, PeiYu Li.

- Partner: ETS (École de Technologie Supérieure) from Québec University
- Contract: INSA research collaboration
- 2010-2012

This work is carried out in collaboration with Professor M. Cheriet of ETS, Director of the consortium Synchromedia (http://www.synchromedia.ca). The originality of this project is to explore in the context of collaborative work, human-computer interaction on pen-based or tactile devices. This year collaboration focuses on the test of our menus with gesture commands on interactive devices at Synchromedia. The gesture commands integrated are those used in an interactive interface for collaborative annotation of images developed by the laboratory Synchromedia. The PhD student PeiYu Li spent three months in the laboratory of ETS this year.
9 Dissemination

9.1 Leadership within scientific community

9.1.1 Program Chair and Committee

- E. Anquetil is Guest Editor of the Special Issue on Frontiers in Handwriting Processing of the Pattern Recognition Letters journal.
- E. Anquetil is member of the program committee of the International Conference on Frontiers in Handwriting Recognition (ICFHR 2012).
- E. Anquetil and B. Coüasnon are members of the program committee of the International Conference on Pattern Recognition (ICPR 2012).
- E. Anquetil and B. Coüasnon are members of the program committee of the ”Colloque International Francophone sur l’Écrit et le Document” (CIFED 2012).
- B. Coüasnon is Conference Co-Chair (with Richard Zanibbi, USA) of the international conference DRR 2013, 20th Document Recognition and Retrieval Conference, conference IS&T/SPIE of the Electronic Imaging Symposium in San Francisco.

9.1.2 Reviewing

- E. Anquetil is a reviewer in 2011 of:
  - IJDAR (International Journal on Document Analysis and Recognition);
  - RTS (Revue Traitement du Signal).
- A. Lemaitre and B. Coüasnon are reviewers of PRL (Pattern Recognition Letters) in 2012.

9.1.3 Member of scientific society

- E. Anquetil is a member of the executive committee of the society grce : “ Groupe de Recherche en Communication Écrite ”.
- E. Anquetil, B. Coüasnon, J. Camillerapp and A. Lemaitre, take part in the activities of the society grce : “ Groupe de Recherche en Communication Écrite ”.
- E. Anquetil is a member of the steering committee of LOUSTIC laboratory of Rennes (laboratoire d’observation des usages des technologies de l’information et de la communication).
- E. Anquetil is an elected member of the laboratory council of IRISA.

9.1.4 Participation to Phd defenses

- E. Anquetil was a reviewer for the Phd of Jinpeng Li, Symbol and Spatial Relation Knowledge Extraction Applied to On-Line Handwritten Scripts, Université de Nantes Angers Le Mans, Oct. 2012.
E. Anquetil was a reviewer for the Phd of Anthony Sorel, Gestion de la variabilité morphologique pour la reconnaissance de gestes naturels à partir de données 3D, Université de Rennes 2, Déc. 2012.

E. Anquetil was a member of the PhD committee of the Phd of Alice Hermann, Création et mise à jour guidées d’objets dans une base de connaissances, INSA de Rennes, Déc. 2012.

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.

- E. Anquetil and B. Coüasnon give lectures at MASTER-RESEARCH d’informatique of University of Rennes 1.
- E. Anquetil is in charge of the Colloquium (COLQ) and of the Internship defense and evaluation of the MASTER-RESEARCH d’informatique of University of Rennes 1.
- E. Anquetil is in charge of the module ”Analyse des documents et des flux audiovisuels pour l’indexation” (FAV) of the MASTER-RESEARCH d’informatique of University of Rennes 1.
- B. Coüasnon is in charge of the module on professionalization adapted to research (PROF) of the MASTER-RESEARCH d’informatique of University of Rennes 1.
- B. Coüasnon was invited for a course at MASTER-RESEARCH “New technologies applied to History” of the Ecole nationale des Chartes on : “Automatic Access to Old Documents”, Paris, France.
- A. Lemaitre was invited for a course at MASTER PRANET of Information and Communication of University of Rennes 2.

9.3 Invited talks
- B. Coüasnon was an invited speaker at the conference CIFED-CORIA 2012 in Bordeaux : ”Fusion of knowledge in document analysis. Examples on archives documents.”.

10 Bibliography
Major publications by the team in recent years


**Doctoral dissertations and “Habilitation” theses**


**Articles in referred journals and book chapters**


**Publications in Conferences and Workshops**


