Project-Team IntuiDoc

Intuitive user interaction for document

Rennes

Activity Report

2016

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 8
4.3 Handwriting, hand-drawn symbol, and gesture recognition systems 9

5 Software **9**
5.1 RESIF: Handwriting recognition by hierarchical fuzzy inference systems 10
5.2 EVOLVE++ / EVOLVE TOUCH: Evolving recognition engine 10
5.3 Varchitect: Windows Store application based on Evolve++/EvolveTouch 10
5.4 Vscript: Android tablet application based on Evolve++/EvolveTouch 11
5.5 DALI: a framework for the design of pen-based document sketching systems 11
5.6 IMISKETCH: interactive off-line sketches recognition 12
5.7 DocRead : an automatic generator of recognition systems on structured documents . . . 12
5.8 Precoce: Library to extract visual indices 13
5.9 LIMO : an isolated handwriting word recognizer 13
5.10 iLib: a feature extraction library 13

6 New Results **14**
6.1 Work on the early vision process 14
6.2 Interactive combination of deep learning and syntactical methods for contextual segmentation and structure learning in document recognition 15
6.2.1 Syntactical Method: DMOS-PI 16
6.2.2 Statistical Method: Deep Learning 17
6.2.3 Dataset Generation 17
6.2.4 Results 17
6.3 Recognition and interpretation of multi-touch gesture interactions 17
6.4 IntuiScript project: Handwriting Quality Analysis 19
6.4.1 Bloch letter and digit analysis: Multi-Criteria Handwriting Quality Analysis with Online Fuzzy Models 19
6.4.2 Preparation to the Cursive Writing 21
6.4.3 Handwriting Analysis of Cursive Letters 21
6.5 Transfer of hand-drawn pattern representation for skeleton-based gesture recognition . . 23

7 Contracts and Grants with Industry **25**
7.1 Script and Labs/ANR joint Laboratory: On-line handwriting and drawing recognition and interpretation for active Learning in e-education 25
7.2 Actif: Active Learning and Collaboration with Tablet Computer, Interactions and Feedbacks (e-Fran - National Innovative BPI Project) 25
7.3 Industrial collaboration contract with Excence company 26

7.4 Industrial software licensing with SCRIPT&GO company 26
7.5 IntuiScript: National Innovative BPI Project 26

8 Other Grants and Activities 27
8.1 National initiatives 27
8.1.1 IMPACT - Brittany and Pays de la Loire region Grant (ARED) 27
8.1.2 DMOS-PI Licensing - University of Nantes 28
8.2 International initiatives 28
8.2.1 Rochester Institute of Technology, USA 28

9 Dissemination 28
9.1 Leadership within scientific community 28
9.1.1 Program Chair and Committee 28
9.1.2 Reviewing 29
9.1.3 Member of scientific society 29
9.1.4 Participation to PhD and HDR defenses 30
9.2 University education 30

10 Bibliography 31

1 Team

Permanent Members

Eric Anquetil [Team Leader, Professor, Insa, HDR]

Nathalie Girard [Associate Professor, Uni. Rennes 1, from September 1st]

Jean Camillerapp [Emeritus Professor, Insa, HDR]

Bertrand Coüasnon [Associate Professor, Insa, HDR]

Ivan Leplumey [Associate Professor, Insa]

Aurélie Lemaitre [Associate Professor, Uni. Rennes 2, maternity leave from January 1st to August 1st]

Yann Ricquebourg [Associate Professor, Insa]

Fixed-term positions

Zhaoxin Chen [Insa PhD student, ARED Bretagne/Pays de la Loire]

Said-Yacine Boulahia [Insa PhD student, Government of Algeria grant]

Mickaël Renault [Insa Research Engineer]

Damien Simonnet [Insa Post-Doc]

Kwon-Young Thomas Choi [Insa PhD student, CD grant, from October 1st]

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 1).

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.

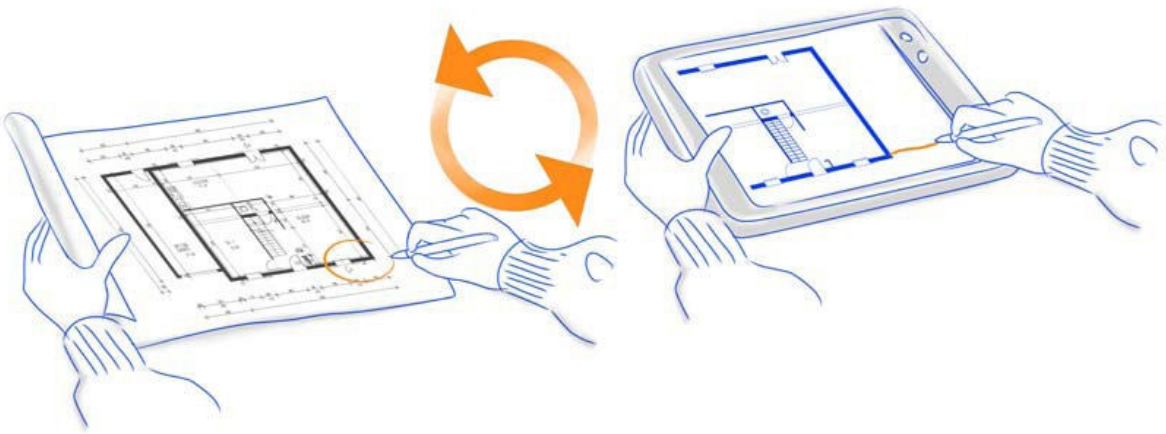


Figure 1: Continuum from paper document to digital interpreted document

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

Combinig several ways of interpreting the content of a document can improve its recognition [7]. Thus we study some mechanisms of knowledge fusion to combine the results various document analysis techniques that are usually studied separately. 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 objets 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 relies 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 [3].

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 documents;
- 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 developp 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 [10] 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, Ipad...*),
- 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 multi-touch 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. In 2014 we focused on bringing a qualitative evaluation of gestures. To demonstrate the main features of Evolve-Touch system, a showcase application is presented in this video : *http://youtu.be/qOx4IY6uYf8*. This work is supported by a *European Regional Development Fund* (FEDER), and protected by a *European and US patents* (N° 2995704 / 14/429,649). In 2014, EvolveTouch was made available for *Android*, *WinRT* and *iOS* tablet systems.

5.3 Varchitect: Windows Store application based on Evolve++/EvolveTouch

Contact: Eric Anquetil

Keywords: Incremental recognition, Evolving system, Gestures and Symbols Recognition, Interior design, Tablet, Windows Store.

Varchitect is a Windows Store application that was developped as part of the effort to port the *Evolve++/EvolveTouch* system to current tablet operating systems. It is available for free at http://apps.microsoft.com/windows/en-us/app/aa0889d0-2097-4a91-aa28-2a74df7e206c

With Varchitect, users can define their own set of gesture commands to insert furniture or architectural elements in a plan, and then design their interior with a stylus or fingers. The users can use a picture (taken from the tablet's builtin camera) as a base and set the scale of their plan to be sure everything fits. Plans made using Varchitect can be shared or printed.

This work is supported by a development fund from *SATT Ouest Valorisation*.

In 2015, Varchitect was downloaded more than 7000 times

5.4 Vscript: Android tablet application based on Evolve++/EvolveTouch

Contact: Eric Anquetil

Keywords: Incremental recognition, Evolving system, Gestures and Symbols Recognition, Learning, Handwriting, Tablet, Android..

Vscript is an Android application that was developped as part of the effort to port the *Evolve++/EvolveTouch* system to current tablet operating systems. It is available on the Android's Play Store at https://play.google.com/store/apps/details?id=fr.irisa.intuidoc.vscript

Vscript is a handwriting learning application for children. It is meant to be used on android tablets with a stylus, although touch input is supported. In this application the children are following a series of exercises of increasing challenge from identifying shapes for reproducing symbols (shapes, letters, numbers) and ultimately composing pictures and words. EvolveTouch is used to interpret the handwriting by recognizing symbols and giving indications of quality.

This work is supported by a development fund from *SATT Ouest Valorisation*.

5.5 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 [8, 9]. *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".

5.6 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 [6] .

5.7 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 [4]. 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;
- 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 handwritting recognition. This work has been realized in collaboration with University of Kolkata and applied on 1922 words from 26 writers.

- FlowRead: a sofwtare for on-line flow-chart segmentation and structure recognition.
- JournRead: a software for the recognition of old newspaper content structured in headlines, articles with title and author, graphics and tables. This software has been developed with a *SATT Ouest Valorisation* development fund;
- MaurdorRead: a software for the structure recognition of heterogeneous and multi-language documents, with handwritten, printed or mixed content. This software has been developed in the context of a PEA (Programme d'Etude Amont - upstream study program) from the DGA (French Ministry of Defense).

5.8 Precoce: Library to extract visual indices

Contact: Jean Camillerapp

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

It is the library on which the DocRead software described in the section 5.7 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.

5.9 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.

5.10 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.

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...).

Finally, motivated by our current perspective of experimenting CRFs on images, we are facing the difficulty to find implementations for general CRF use and able to process in the continuous domain of numerical data needed by image processing (whilst not theoretically compulsory, usual recommended implementations are restricted to symbolic data). To handle images, some implementations propose a workaround using a neuronal system to handle numeric data from images. But the CRF system is only on top of outputs of this subsystem, as a superior layer and as a posterior processing (RNNSharp from Microsoft, Hidden-Unit Conditional Random Fields, ...) Thus we work to add a fully numerical implementation of CRFs, from existing generic symbolic implementations (like CRF++ or Wapiti).

6 New Results

6.1 Work on the early vision process

Participant: Jean Camillerapp.

The following works aims at extracting in the images some primitives that may be used by DMOS method for the interpretation of the documents.

Music scores

In 2016 Kwon-Young Choi did his master's internship in the team on the construction of neural networks able to recognize and locate symbols. With this approach it is possible to overcome some of the difficulties linked to segmentation errors. He wanted to test his work in the field of musical scores. So we continued the work we carried out at the end of the year 2015 on the extraction of visual indices adapted to music. This made it possible to provide input data for this recognizer.

Extraction of clues from PDF files

We continued the work started in 2015 by working on a wider variety of documents.

After this experiment, we think that, for documents that mainly contain text, it is possible to extract from PDF files the elements useful for DMOS to retrieve the structure of a document and in particular the division into columns as well as the logical sequence of the columns.

However, it appeared that the tables that allow to translate the internal representations of characters into Unicode are not always present or may contain errors.

For graphic documents, there are many ways to express, in PDF, the same visual aspect. For example, a musical score can be described as a textual document with specific characters or, on the contrary, be completely drawn.

Therefore, it is difficult to process databases generated by different editors.

Extraction of text lines

In 2013 and 2014 we worked on detecting text lines using blurred images. In 2016 we have completely renewed this approach.

We kept the principle of the blurred image , but focusing first on the detection of the lower body of the text. This one is much more stable than the upper body because there are fewer descending jamb than ascending.

This new version has been made available to the IVC team of the IRCCYN in Nantes. (figure 2)

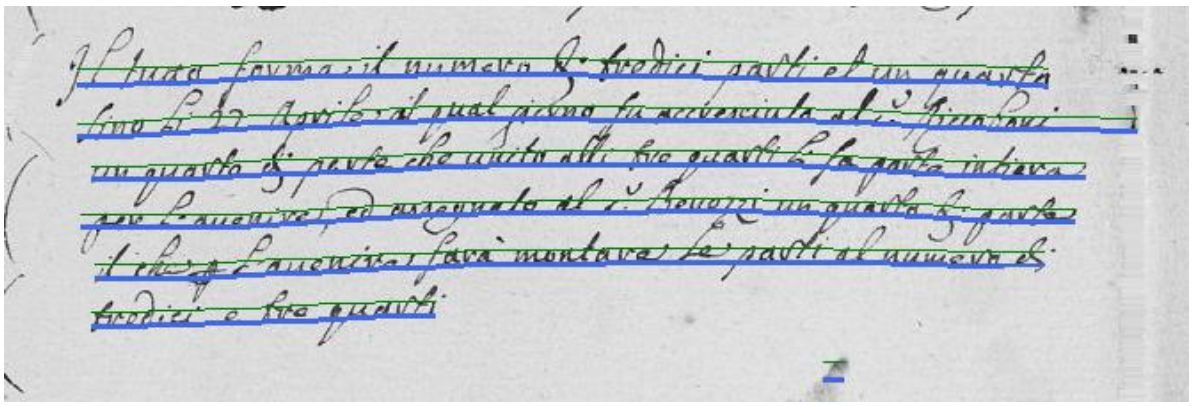


Figure 2: Extraction of handwritten text lines,blue is the base and green the top

We are currently working on an extension to multi-column. It is now able to process images that have been digitized with a slight rotation. (figure 3)

6.2 Interactive combination of deep learning and syntactical methods for contextual segmentation and structure learning in document recognition

Participants: Kwon-Young Choi, Bertrand Coüasnon, Yann Ricquebourg, Richard Zanibbi.

This work is done in the context of a collaboration with Richard Zanibbi from the Rochester Institute of Technology (see 8.2.1). In document image analysis, document recognition is often decomposed into steps like these: segment the image into multiple little components, recognize all these components, construct a logical representation by combining all of the logical components recognized. This workflow works as long as the two first steps are accurate and error free. But it is often that we encounter documents where the segmentation is not as trivial as to isolate its connected components. In fact, we can distinguish two problematic cases:

- The logical component is broken down into multiple parts, segmentation has to group multiple connected components to form a logical information;

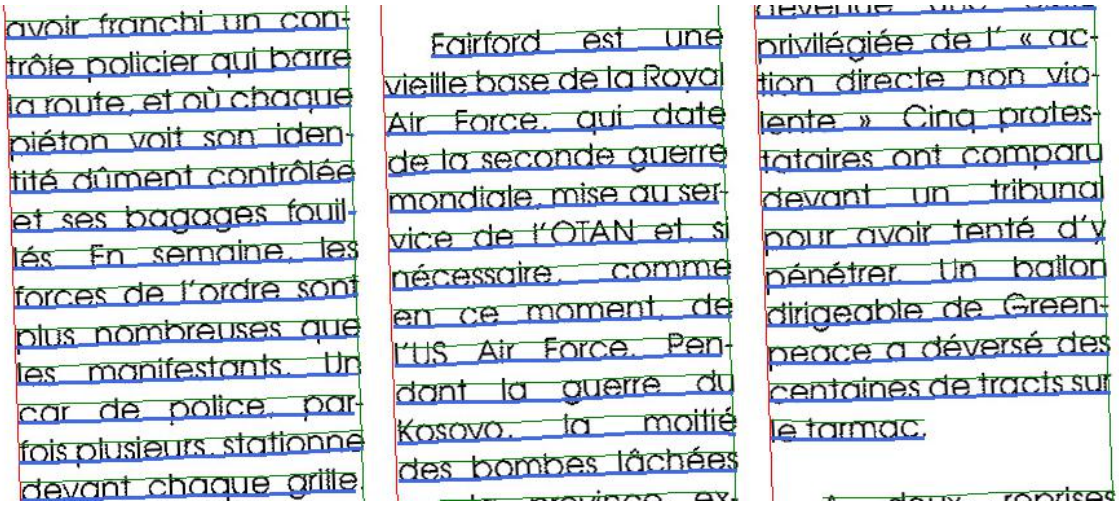


Figure 3: Detection of columns

- Multiple logical component are connected as one connected component, segmentation has to separate one connected component into multiple connected components.

These segmentation problems are commonly found because of the complex document structure, high document density, bad digitization like showed in figure 6.2. In order to resolve these segmentation problems we need both visual information and contextual knowledge, this is why we propose to use both syntactical methods and deep learning techniques in an hybrid model capable of object localization and recognition.

(a) Broken symbol (b) Touching symbols

Figure 4: Examples of segmentation problems

6.2.1 Syntactical Method: DMOS-PI

DMOS-P [4] is a generic multi-resolution document recognition method developed by the Intuidoc team. This method is capable of modeling complex document structure using syntactical rules in a bi-dimensional grammar and has been used to recognize a wide range of documents like musical scores, mathematical formulae, flow charts ... However, it has been established that trying to resolve segmentation problems using a-priori rules introduce too much complexity in the grammar and is very time consuming.

6.2.2 Statistical Method: Deep Learning

In another hand, Deep Learning techniques like convolutional or recurrent neural networks are very effective in order to produce accurate object detection models [FGMR10,RHGS15,SEZ+13]. But their use of contextual information is limited to a relative local scope, limiting their ability to recognize complex deep document structure. Therefore, we would like to bring complex contextual information to the network, both local and global, by using a syntactical method like DMOS-P. The model of neural network proposed would then be able to produce an accurate join localization and classification using both local and global context.

6.2.3 Dataset Generation

One of the downside of using deep learning techniques is that they need lost of training data in order to produce an accurate and generalizable model. However, the DMOS-PI system is well suited for this task as it is able to generate training data customized to the problem we want to resolve: join segmentation and classification.

6.2.4 Results

Music scores has first been used as a test case for this work. A prototype grammar has been used from the work of [CC94] and is able to construct a simple representation of the musical notation containing staves, notes with or without accidentals. This grammar used a simple music symbol classifier in order to generate training data of accidentals. The dataset generated is constituted of 2025 symbols of four different classes and we also generated a lot of reject data that are thumbnails containing no symbols. The dataset was entirely manually checked and ground-truth errors should now be really rare. A first model of convolutional neural network that is able to do a join localization and classification with rejection has been trained on this dataset. More models are planned to be experimented inspired from state of the art object detection models.

6.3 Recognition and interpretation of multi-touch gesture interactions

Participants: Zhaoxin Chen, Eric Anquetil, Christian Viard Gaudin (IRCCYN NANTES), Harold Mouchère (IRCCYN NANTES), Excence company.

This work is funded by the Impact ARED project:(Brittany and Pays de la Loire region/Insa grant - cf. section 8.1.1). This project is the result of the collaboration between IRISA/IntuiDoc and IRCCYN laboratories. The goal is to design a new process to recognize and interpret multi-touch interactions. It aims to extend the composition capability of complex structured documents (architectural plan, diagram, mathematical expressions ...) with multipoint interaction on touch-screen.

[FGMR10] P. F. Felzenszwalb, R. B. Girshick, D. McAllester, D. Ramanan, "Object Detection with Discriminatively Trained Part-Based Models", *IEEE Transactions on Pattern Analysis and Machine Intelligence 32*, 9, September 2010, p. 1627–1645.

[RHGS15] S. Ren, K. He, R. Girshick, J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks", *in: Advances in Neural Information Processing Systems 28*, C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama, and R. Garnett (editors), Curran Associates, Inc., 2015, p. 91–99, `http://papers.nips.cc/paper/5638-faster-r-cnn-towards-real-time-object-detection-with-region-proposal-networks.pdf`.

[SEZ+13] P. Sermanet, D. Eigen, X. Zhang, M. Mathieu, R. Fergus, Y. LeCun, "OverFeat: Integrated Recognition, Localization and Detection using Convolutional Networks", *arXiv:1312.6229 [cs]*, December 2013, arXiv: 1312.6229, `http://arxiv.org/abs/1312.6229`.

[CC94] B. Coüasnon, J. Camillerapp, "Using grammars to segment and recognize music scores", *International Association for Pattern Recognition Workshop on Document Analysis Systems*, 1994, p. 15–27, `ftp://ftp.idsa.prd.fr/local/IMADOC/couasnon/Articles/das94.ps`.

Multi-touch functionality extends the powerfulness of Human-Computer Interaction. Current multi-touch interactions are mainly restricted to perform direct object manipulations such as a zoom or a rotation. The objective of this work is to interpret a multi-touch gesture as a more complex indirect command, which could be used as a short cut command. Unlike the mono-touch gesture where strokes are always written in sequence, the strokes in multi-touch gesture may have complex synchronization or intersection relations. Two gestures may have the same appearance but contain different inner-stroke relations. A fundamental issue is the modeling of these relations between strokes as the key feature for multi-touch gesture recognition. Due to the fact that multi-touch gestures may have a variety number of simultaneous fingers' contacts, it is difficult to design a fixed number of feature set to characterize the dynamic spatial and temporal relations between contacts. Therefore, we propose to use graph (see Figure 5) to model the multi-touch gesture and integrate these spatial and temporal relations, and even shape information, into the graph.

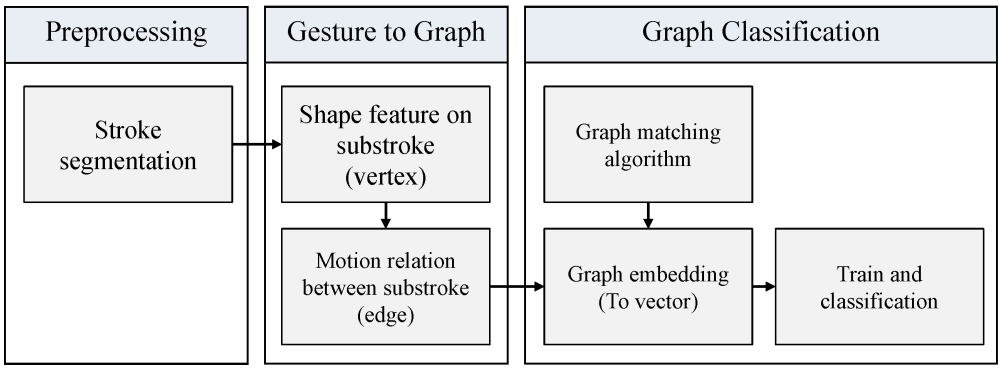


Figure 5: Framework of the graph modeling with motion based features.

This method is evaluated on a new multi-touch gesture database (MTGSet) containing, 6138 samples, 31 different kinds of gestures : multi-stroke, multi-touch and sequential multi-touch. We have made this dataset freely available to constitute a baseline benchmark for the multi-touch gesture recognition community.

We achieved 98.97% recognition rate on this dataset and prove that our recognition system can well capture the features for multi-touch gestures comparing to other static analyzing feature set. These results have been submitted to Pattern Recognition Journal [ZAVGM16]. To explore the possibility of using multi-touch gesture for both direct manipulation and indirect command in a same context, we studied the early recognition strategy aiming at recognize a gesture by its beginning part. We proposed a multi-classifier structure with reject option allowed classifiers. The experimental result shows that we achieved 82.38% accuracy rate with an average earliness of 55.89%. Meanwhile, the error rate is 5% lower than the system without reject option.

At the end of the year, in collaboration with the Excence company (cf. section 7.3), we collect a new online Multi-User Multi-Touch handwritten diagram DataBase (MUMTDB) for evaluating recognition systems under the multi-user situation(see Figure 6). The data is collected according to two predefined mind map scenarios which contains 9 classes of graphical symbols. Each scenario is completed by involving two users at the same time. Since the users are given freedom to draw the symbols as they want, the dataset contains a diversity of multi-stroke and even multi-touch symbols. It allows addressing new challenging problems regarding the recognition of simultaneous composition of structured documents.

[ZAVGM16] C. Zhaoxin, E. Anquetil, C. Viard-Gaudin, H. Mouchère, "Analyzing and recognizing multi-touch gestures using graph models", *Pattern Recognition*, Submitted in November 2016.

The dataset have been presented in the ICFHR International Conference [25] and is freely available on-line .

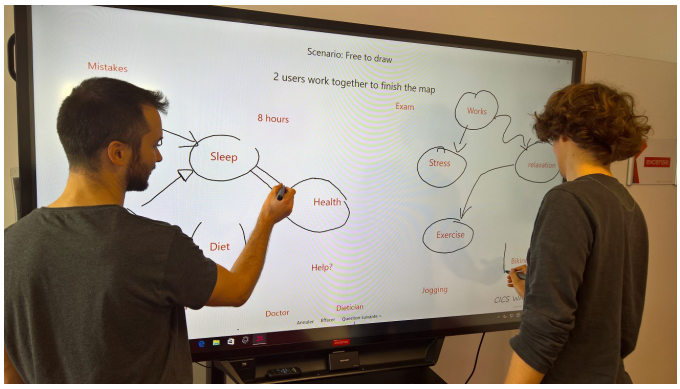


Figure 6: The diagram data acquisition procedure on a 80" touch screen. Two users are drawing the diagram together using styluses.

6.4 IntuiScript project: Handwriting Quality Analysis

Participants: Damien Simonnet, Eric Anquetil, Mickael Renault.

IntuiScript (*http://intuiscript.com/*) is a three years research project founded by the French government as part of innovative projects (BPI) in the e-education field It targets the introduction of innovative services and digital contents in the development of fundamental skills at school (see section 7.5). The main objective of the *IntuiScript* project is to offer an advanced digital writing learning experience at school by using tablet and tactile digital devices (with finger touch and stylus).

The *IntuiScript* project focuses on the improvement of previous modules (*i.e.,* block letters and cursive letters) and on the development of new modules (block digits, preparation to the cursive letter). Results have been published in the journal *National Handwriting Association* [18] and a research paper entitled *Multi-Criteria Handwriting Quality Analysis with Online Fuzzy Models* has been submitted to the journal *Pattern Recognition* [17]. Several presentation will be made in the International Symposium for Educational Literacy (SILE/ISEL) [20, 21]

6.4.1 Bloch letter and digit analysis: Multi-Criteria Handwriting Quality Analysis with Online Fuzzy Models

In the context of the development of a digital workbook providing feedback during the handwriting learning process, this work presents an approach to analyse handwriting quality with regards three different aspects: *shape*, *order* and *direction* and introduces a *multi-criteria* architecture with hierarchical dependencies (*e.g., order* and *direction* depend on *shape*) and a precise characterisation of each criteria with specific features.

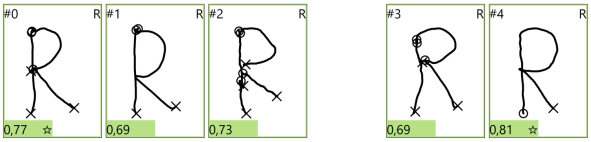
This work has shown significant improvements of the results on a dataset collected in four preschools with 171 children:

- *shape* : the analysis error for positive and negative gestures decreases of 33% and 27% (analyses rates on positive and negative samples are respectively 96% and 92%).

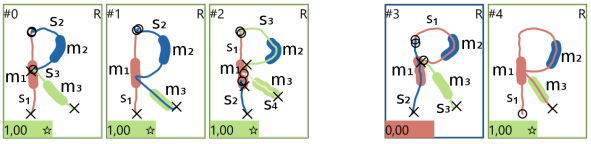
- *order* : the analysis error for positive and negative gestures decreases of 96% and 97% (analyses rates on positive and negative samples are respectively 98% and 99%).

- *direction* : the analysis error for positive and negative gestures decreases of 67% and 98% (analyses rates on positive and negative samples are respectively 98% and 99%).

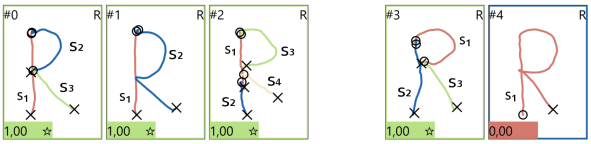
Qualitative results are presented in Figure 7 and show a continuous degradation of the *shape*, and the ability to detect specific errors (*e.g., order* and *direction*). Figure 8 illustrates three possible strategies for the teacher to customise the feedback returned to the children (*i.e., multi-criteria*). In Figure 8(a), it corresponds to an early learning stage, *shape*, *order* and *direction* must correspond to the model. As opposition, the strategy in Figure 8(b) corresponds to an advance learning stage, where children acquired there own writing style. The legibility is therefore the only evaluation criterion. Finally, in Figure 8(c) and Figure 8(d) respectively, there is no constraint of order and direction. These strategies are intermediate learning stages where a constraint about order or direction can be relaxed. This illustrates the versatility of the method to adapt to various pedagogical scenario to satisfy teacher expectations.



(a) Shape



(b) Order

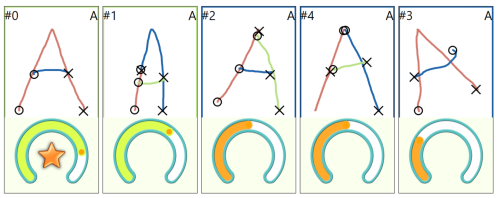


(c) Direction

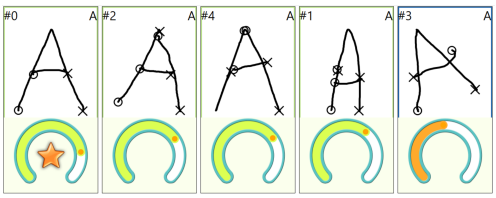
Figure 7: Qualitative results on the three criteria with positive shapes and where strokes are coloured from the first to the fourth with red, blue, green and brown (the end of each stroke being represented with an alpha colour rectangle). In (b) median strokes are represented with thicker lines.

In-class experiments conducted with the *global* classifier have shown that children get quickly familiar with the application and try to improve the evaluation score that is well understood. Moreover, the personalised feedback allows children to progress with autonomy at their own speed, and to stay concentrated during the whole session of 20 minutes which is difficult for young children.

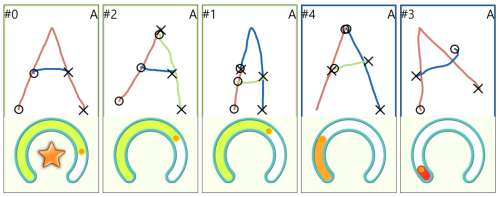
The new configuration of the feedback returned to the children (*i.e., multi-criteria* configuration) and the feedback characterising errors (*i.e., shape*, *order* and *direction*) will be used in the next in-class experiments.



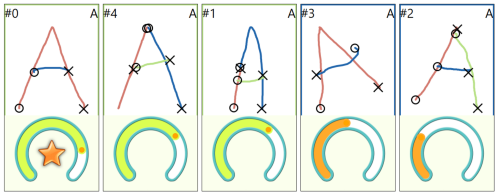
(a) Early learning stage: All Criteria ($\alpha_s = 0.7$, $\alpha_o = 0.15$, $\alpha_d = 0.15$).



(b) Advanced learning stage: Shape Only ($\alpha_s = 1.0$, $\alpha_o = 0.0$, $\alpha_d = 0.0$).



(c) Intermediate learning stage: No Order ($\alpha_s = 0.7$, $\alpha_o = 0.0$, $\alpha_d = 0.3$).



(d) Intermediate learning stage: No Direction ($\alpha_s = 0.7$, $\alpha_o = 0.3$, $\alpha_d = 0.0$).

Figure 8: Qualitative results of the *multi-criteria* classifier with the colour scale indicator returned to the children. It illustrates four possible strategies for the teacher to customise the feedback returned to the children (*i.e., multi-criteria*). In (a), it corresponds to an early learning stage, the shape, order and direction must correspond to the model. The strategy (b) only concentrates on the legibility. In (c) and (d) respectively, there is no constraint of order and direction.

6.4.2 Preparation to the Cursive Writing

The objective of this module is to work on the development of graphomotor skills by considering loops and elements present in cursive writing such as cusp. It evaluates the writing regularity, the respect of the line spacing with several sequences of exercises. These exercises contain reapeating pattern with various difficulties (loop, bridge), and spatial constraints such as checking points and obstacles as illustrated in Figure 9.

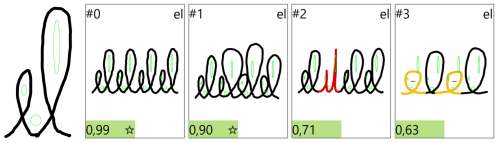
6.4.3 Handwriting Analysis of Cursive Letters

The objective of this module is to evaluate the quality of cursive handwriting production of a word or a subpart of a word compared to a reference model. It implies to deal with the segmentation of a word in letters, and to identify matching, missing and adding letters compared to the reference model.

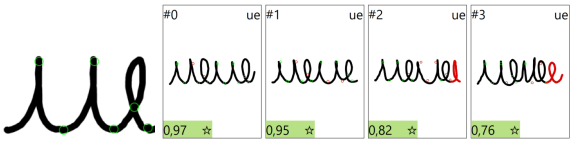
First, based on descending areas and singularity points (see Figure 10(b)), letter segments (see Figure 10(c)) are extracted, as in [AL97] (see section 5.1) by using the segmentation graph in Figure 10(d). Then, the global approach used for block letters is improved by adding features related to descending areas which are stable parts in the cursive writing. By opposition to the block writing analysis, only a global feedback is given to children because errors such as order and direction are less common with this type of writing. The final feedback returned to children is based on the quality of each letter drawn and its correspondence with the word asked (*i.e.,* matching, missing and adding letters).

The next steps of the handwriting cursive analysis engine will be related to the management of

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



(a) Sequence with loops (type 'el') with obstacles



(b) Sequence with loop and bridge (type 'ue') with obstacles and checkpoint

Figure 9: Example of feedback returned by the module preparation to the cursive writing for the template 'el' (a) and 'ue' (a). Obstacle are represented in (a) by green ellipses. Checkpoints corresponds to circles in (a) and are in green where the drawing passes through the checking point and red else. Red parts of the drawing correspond to incorrect drawing, and yellow parts to loops on which the size is disproportionate with regards to the model.

accents used in French (*i.e.,* acute, grave, circumflex and dieresis), and to specific feedback indicating the type of mistake (*e.g.,* incorrect letter shape, missing and adding letters).

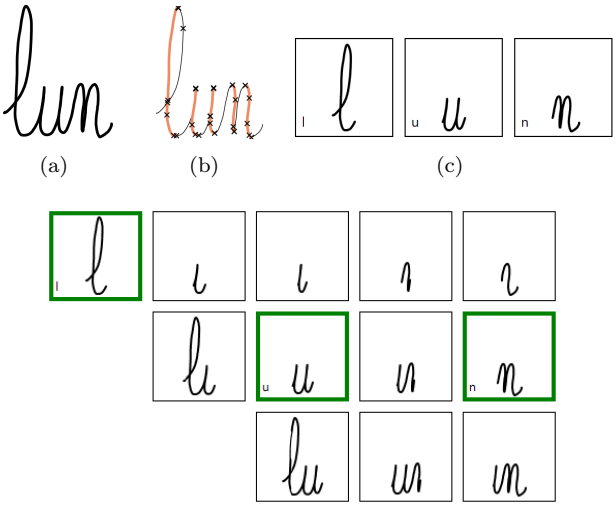


Figure 10: The original gesture (a) and the extraction of descending areas (lines in orange) and singularity points (cross) (b) that are used to extract the segmentation (c) from the segmentation graph (d).

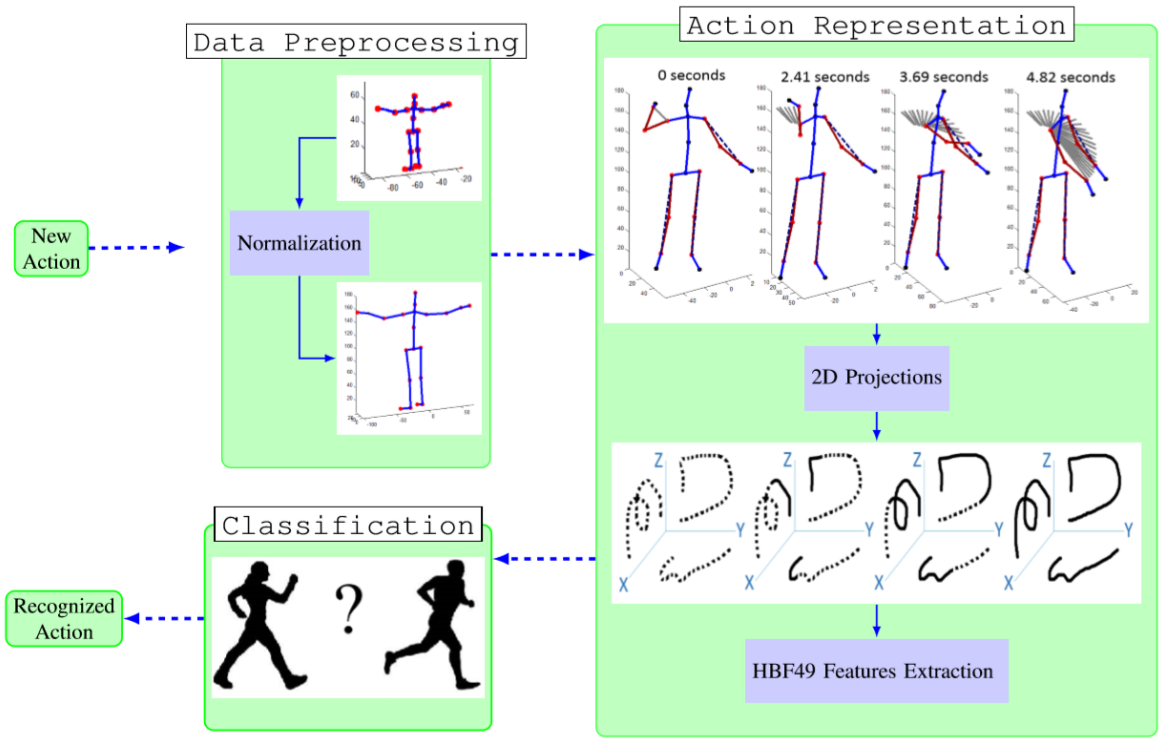


Figure 11: Illustration of the major steps constituting the 3DMM skeleton-based approach

6.5 Transfer of hand-drawn pattern representation for skeleton-based gesture recognition

Participants: Said-Yacine Boulahia, Eric Anquetil, Richard Kulpa, Franck Multon.

Human action recognition attracts increasing attention among researchers in computer vision. We are currently working on an approach to recognize human actions given sequences of 3D joint positions. Different from most previous skeleton-based approaches, our work does not consist of proposing a new set of features or an original learning method. The approach is rather based on the study of handwriting recognition work in order to explore its transposal to the recognition of 3D actions.

Several studies have been conducted during the last twenty years on the recognition of handwriting trajectories (writing, drawing, 2D gestures, etc.). It seems therefore promising to capture this knowledge and to transpose it to the 3D skeleton gestures which by nature are an extension of gestures produced on a 2D tablet.

In particular, we selected an efficient set of handwriting features, namely HBF49 [5], and proposed to model an action as a modified 2D writing symbol. This approach which we refer to as 3D Multistroke Mapping (3DMM) follows a straightforward recognition strategy since the aim is to measure the potential of such transposal. Figure 11 illustrates the major steps constituting the proposed action recognition approach.

The effectiveness of the 3DMM approach is evaluated by means of two classical classifiers, namely Support Vector Machine (SVM) and Multilayer Perceptron (MLP), on three challenging 3D action datasets captured by commodity depth cameras. To illustrate this performance we give in table 1 the

obtained results on one of these datasets.

Method	Recognition rate (%)
HMIJ + SVM [Ofli, 2014]	84.40
HMIJ + Nearest neighbour [Ofli, 2014]	80.73
SMIJ +Nearest neighbour [Ofli, 2014]	81.65
SMIJ +SVM [Ofli, 2014]	82.57
Dynamic Temporal Warping [Reyes, 2011]	82.08
MIJA/MIRM + LCSS [Pazhoumand, 2015]	85.23
Cov3DJ + SVM [Hussein, 2013]	95.41
BIPOD + SVM [Zhang, 2015]	96.70
HOD + SVM [Gowayyed, 2013]	97.27
3DMM approach + SVM + 100 best features	91,74
3DMM approach + MLP + 20 best features	92.66
3DMM approach + SVM + 400 best features	94,49
3DMM approach + MLP + 80 best features	94.49
HIF3D + SVM	**98.17**

Table 1: Comparisons between 3DMM, HIF3D and previous approaches on HDM05 database.

The experimental evaluations attest the validity of this transposal since the introduced approach achieves comparable results regarding more sophisticated algorithms. This approach is to be published in FG2017 [23].

These preliminary results encouraged us to go a step forward. In fact, instead of projecting the skeleton trajectories into 2D spaces, we proposed to extend the HBF49 handwriting features [5] to handle 3D skeleton trajectories. This resulted in a new set of 3D features that we called HIF3D as an acronym of Handwriting-Inspired Features for 3D skeleton-based action recognition [22]. An overall view of the process is resumed in Figure 12. The achieved performance is given in table 1.

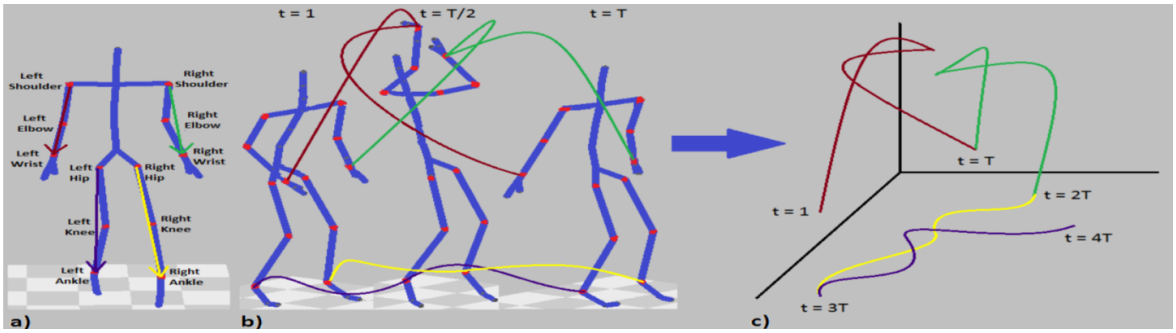


Figure 12: (a) Selected joints with the associated normalized vectors (collored arrows) and (b) illustration of the four morphology-independent trajectories and (c) the 3D multistroke pattern resulting from the trajectories assembling.

7 Contracts and Grants with Industry

7.1 Script and Labs/ANR joint Laboratory: On-line handwriting and drawing recognition and interpretation for active Learning in e-education

Participant: Eric Anquetil, Nathalie Girard.

- Partner: *Script&Go Company*
- 36 months (2017-2019).
- Contract: INSA

With the great success of the IntuiScript Project (BPI/PIA-2) (https://www-intuidoc.irisa.fr/projet-intuiscript/), we want to consolidate our implication in e-education to achieve innovative contribution based on pen-based tablet devices. In that way, we build a ANR joint laboratory (LabCom) between IntuiDoc and Script&Go Company on this promising topic for a total amount of 1 562K€(6.15 ETP) with 300K€of grant. This new structured partnership would have the goal to design new educational learning experience at school by using tablet and tactile digital devices: handwriting learning, generative drawing/sketching, on-line handwritten arithmetic expression, collaborative interaction using 2D gestures, document annotation...

7.2 Actif: Active Learning and Collaboration with Tablet Computer, Interactions and Feedbacks (e-Fran - National Innovative BPI Project)

Participant: Eric Anquetil, Nathalie Girard, Yann Ricquebourg.

- Partners: *LP3C*, *Loustic*, *Société Script&Go*, *Saooti*, *Région Bretagne*, *Espe*, *espace des sciences*
- 48 months (2017-2020).
- Contract: INSA

We have just won in September 2016 the project "ACTIVE" from the national call "e-fran / innovative national project ", for a total amount of 1 143 856€. This is a 4 years project (2017-2020), led by the LP3C of the University of Rennes 2. It carries on: "Active Learning and Collaboration with Tablet Computer, Interactions and Feedbacks". One of the key topics in this project we want to investigate is how to encourage new learning strategy based on "active learning" and especially on "generative drawing" using pen-based numerical devices. The concept is to ask students to create drawings while reading text for causing generative processing that leads to better learning outcomes. The goal is to improve the learning of student by considering "learning" as a generative activity especially with collaborative/cooperative interaction. In the scope of digital learning, the potential induces by pen-based devices (tablet, TNI and new interactive screen) is really interesting. The goal is to investigate how we can automatically generate intelligent "corrective" or "predictive" individual and collective feedbacks to users by interpreting their drawing process: summarizing, mapping, sketching... The IntuiDoc team will ensure all the design related to the analysis of graphic productions with one engineer, one postdoctoral researcher and one PhD student. They will work on the issues of "Generative Drawing" in the field of Geometry at secondary school.

7.3 Industrial collaboration contract with Excence company

Participant: Eric Anquetil, Zhaoxin Chen.

- Partners: *Excence*, *IRCCYN NANTES*
- 16 months (2015-2016).
- Contract: INSA

The contract with Excence company is based on the collaboration between IRISA/IntuiDoc and IRCCYN laboratories with the PhD student Chen Zhaoxin (see section 8.1.1). The PhD topic is to design a new process to recognize and interpret multi-touch interactions. This industrial collaboration is focused on a new dimension of this work: the multi-user interaction. The goal is to segment and recognize on the fly the multi-touch gestures perform by several users in the same time and on a same device such as interactive widescreen.

7.4 Industrial software licensing with SCRIPT&GO company

Participant: Eric Anquetil.

- Partners: *SCRIPT&GO company*
- since 2014
- 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 Intuidoc 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/) and the *IntuiScript* Project (http://intuiscript.com/) which are labelled by the "Images et Réseaux" cluster.

7.5 IntuiScript: National Innovative BPI Project

Participants: Eric Anquetil, Mickael Renault, Damien Simonnet.

- Partners: *SCRIPT&GO company*, *Microsoft*, *Academy of Brittany*, *Region of Brittany*, *LOUSTIC laboratory of Rennes (laboratoire d'obervation des usages des technologies de l'information et de la communication)*
- 32 months (2014-2017).
- 2 resources for IntuiDoc: 1 Research Engineer and 1 Post-Doc .
- Contract: INSA

IntuiScript (*http://intuiscript.com/*) is a three years research project founded by the French government as part of innovative projects (BPI) in the e-education field targeting the introduction of innovative services and digital contents in the development of fundamental skills at school.

The main objective of the *IntuiScript* project is to offer an advanced digital writing learning experience at school by using tablet and tactile digital devices (with finger touch and stylus). This project is structured around the conception of a digital workbook to help teachers and children from three to seven years old during the handwriting learning process, by giving on-line and off-line feedback. The former gives a personalised feedback to children to help them to learn from their mistakes with autonomy by presenting them adapted pedagogic content. The latter is a detailed analysis for teachers to evaluate the content of the digital workbook composed of the historic of all letters written by a children.

A user-centered design approach is used in the development of this digital workbook: modules are designed by educational experts followed by experiments in school to use feedback from children and teacher to improve the pedagogical approach of exercises performed on tactile digital devices.

The validation of this research project is based on experiments performed in school from half day to a month with a large number of French primary school students in Brittany. Currently, two sessions of experiments have been conducted, the first one occured in june 2015 in Rennes area, the second one which have opened the project to other schools from brittany (Brest, Quimper, Saint Brieuc,...) took place in december 2015. During these experimentations, more than 1000 children distributed in 18 schools have participated.

The Intuiscript project has been featured in several french television news bulletin:

- TV France 3 Bretagne (12/10/2015) : https://vimeo.com/142233890
- TV 12h45 de M6 (28/09/2015) : https://vimeo.com/140660028

8 Other Grants and Activities

8.1 National initiatives

8.1.1 IMPACT - Brittany and Pays de la Loire region Grant (ARED)

Participants: Eric Anquetil, ZhaoXin Chen, Christian Viard-Gaudin (IRCCYN NANTES), Harold Mouchère (IRCCYN NANTES).

- Partners: Brittany and Pays de la Loire region
- Contract: INSA
- 36 months (2013-2016)

This project funds (Brittany and Pays de la Loire region/Insa grant) the Ph.D. of ZhaoXin Chen on Recognition and interpretation of structured documents with multi-touch interactions. It aims to extend the composition capability of complex structured documents (architectural plan, diagram, mathematical expressions ...) with multipoint interaction on touch screen (see section 6.3). This project is supervised by Prof. E. Anquetil (INSA of Rennes) and Prof. C. Viard-Gaudin (Univ. of Nantes) in collaboration with assistant Prof. Harold Mouchère(Univ. of Nantes).

8.1.2 DMOS-PI Licensing - University of Nantes

Participants: Jean Camillerapp, Bertrand Coüasnon, Aurélie Lemaitre, Harold Mouchère (IRCCYN NANTES), Geoffrey Roman-Jimenez (IRCCYN NANTES), Christian Viard-Gaudin (IRCCYN NANTES).

- Partners: IRCCYN NANTES
- Contract: INSA
- Since 2016

We started in September 2016 a collaboration with Irccyn - University of Nantes on the layout recognition of registers of the Théâtre-Italien from the 18th century. The University of Nantes is a partner of the ANR CIREFI and will work with the DMOS-PI method, proposed by IntuiDoc, to build a document structure recognition system for these handwritten registers, which will drive handwritten text recognizers, to make a complete information retrieval system.

8.2 International initiatives

8.2.1 Rochester Institute of Technology, USA

Participants: Kwon-Young Choi, Bertrand Coüasnon, Yann Ricquebourg, Richard Zanibbi (RIT).

- Partner: Rochester Institute of Technology (Rochester, NY, USA).
- Since 2016

In collaboration with Richard Zanibbi from the Rochester Institute of Technology (RIT), Rochester, NY, USA, we work on interactive learning by combining deep learning technics, syntactical analysis and user interaction to introduce learning of segmentation. We propose to study the strong combination of deep learning and syntactical methods to build a document structure recognition system able to deal with segmentation problems by learning them. The syntactical part models the structure and brings complex context to the deep learning recognition. This collaboration is done through the co-supervising (Bertrand Coüasnon, Yann Ricquebourg and Richard Zanibbi) of the PhD of Kwon-Young Choi (see section 6.2).

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 “Drawing and Handwriting Processing for User-Centered Systems” IEEE Transactions on Human-Machine Systems (April 2017).
- E. Anquetil is member of the program committee of the Conference of the International Graphonomics Society (IGS 2017).
- E. Anquetil is member of the program committee of the International Conference on Frontiers in Handwriting Recognition (ICFHR 2016).

- E. Anquetil is member of the program committee of the International Workshop on Pattern Recognition (IWPR 2016).
- E. Anquetil and A. Lemaitre are members of the program committee of the "Colloque International Francophone sur l'Écrit et le Document" (CIFED 2016).
- B. Coüasnon is General Chair of the "Colloque International Francophone sur l'Écrit et le Document" (CIFED 2016).
- B. Coüasnon is Guest Editor of the Special Issue "De l'analyse du manuscrit à la recherche d'information dans les réseaux sociaux" of the journal "Document Numérique".
- Bertrand Coüasnon is Associate Editor of Frontiers in Cultural Heritage Digitization.
- B. Coüasnon is member of the program committee of the International Conference on Pattern Recognition (ICPR 2016).
- B. Coüasnon is member of the program committee of the international conference DRR 2016, 23rd Document Recognition and Retrieval Conference, conference IS&T/SPIE of the Electronic Imaging Symposium in San Francisco.
- E. Anquetil, B. Coüasnon and A. Lemaitre are members of the program committee of the International Conference on Document Analysis and Recognition (ICDAR 2017).

9.1.2 Reviewing

- E. Anquetil is a reviewer in 2016 of:
  - IEEE Transactions on Human-Machine Systems ;
  - IJDAR (International Journal on Document Analysis and Recognition).
- B. Coüasnon is a reviewer in 2016 of:
  - IJDAR (International Journal on Document Analysis and Recognition).
- A. Lemaitre is a reviewer in 2016 of:
  - PR (Pattern Recognition);
  - IJDAR (International Journal on Document Analysis and Recognition).

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 and B. Coüasnon take part in the animation structure " SA 5.2-Écrit " of the Pôle 2: ICC (Interaction coopération et communication), Axe 5: " Communication orale, écrite et visuelle " du GDR-PRC CNRS I^3 (Information, Interaction, Intelligence).
- 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.
- E. Anquetil is an elected member of the 27e section of the CNU council of INSA.
- B. Coüasnon is member of the board of Valconum (Centre Européen de Valorisation Numérique).
- B. Coüasnon is an elected member of the laboratory council of the INSA component of IRISA.
- B. Coüasnon is head of the Media and Interactions Department of IRISA since April 2016.

9.1.4 Participation to PhD and HDR defenses

- E. Anquetil was a reviewer for the HDR of Harold Mouchère, Reconnaissance de documents manuscrits structurés. Des équations manuscrites aux documents anciens, Université de Nantes, October 2016.
- B. Coüasnon was a member of the PhD committee of the PhD of Frank Julca-Aguilar, Recognition of Online Handwritten Mathematical Expressions using Contextual Information, Universidade de São Paulo, Université de Nantes, April 2016.
- B. Coüasnon was a member of the PhD committee of the PhD of Sonia Yousfi, Embedded Arabic text detection and recognition in videos, Université de Lyon, July 2015.
- B. Coüasnon was a reviewer for the PhD of Nabil Ghanmi, Segmentation d'images de documents manuscrits composites : application aux documents de chimie, Université de Lorraine, September 2016.

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 is program manager of the Master of Science *"Innovation and Entrepreunership"* of INSA and ESC School of business of Rennes.
- E. Anquetil and B. Coüasnon give lectures at 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 two courses at Master-Research "New technologies applied to History" of the Ecole nationale des Chartes on : "Digital Documents: Textual Documents" and "Automatic Access to Old Documents", Paris, France.

..

10 Bibliography

Major publications by the team in recent years

[1] A. Almaksour, E. Anquetil, "Improving premise structure in evolving Takagi-Sugeno neuro-fuzzy classifiers", *Evolving Systems 2*, 2011, p. 25–33, 10.1007/s12530-011-9027-0, http://dx.doi.org/10.1007/s12530-011-9027-0.

[2] A. Almaksour, E. Anquetil, "ILClass: Error-Driven Antecedent Learning For Evolving Takagi-Sugeno Classification Systems", *Applied Soft Computing*, 0, 2013, http://www.sciencedirect.com/science/article/pii/S1568494613003414.

[3] B. Coüasnon, J. Camillerapp, I. Leplumey, "Access by Content to Handwritten Archive Documents: Generic Document Recognition Method and Platform for Annotations", *International Journal on Document Analysis and Recognition, IJDAR 9*, 2-4, 2007, p. 223–242, http://springerlink.com/content/5843461264501u81/.

[4] B. Coüasnon, "DMOS, a Generic Document Recognition Method: Application to Table Structure Analysis in a General and in a Specific Way", *International Journal on Document Analysis and Recognition, IJDAR 8*, 2-3, June 2006, p. 111–122.

[5] A. Delaye, E. Anquetil, "HBF49 feature set: A first unified baseline for online symbol recognition", *Pattern Recognition 46*, 1, 2013, p. 117 – 130.

[6] A. Ghorbel, A. Lemaitre, E. Anquetil, S. Fleury, E. Jamet, "Interactive interpretation of structured documents: Application to the recognition of handwritten architectural plans", *Pattern Recognition 48*, 8, 2015, p. 2446 – 2458, http://www.sciencedirect.com/science/article/pii/S0031320315000473.

[7] A. Lemaitre, J. Camillerapp, B. Coüasnon, "Multiresolution Cooperation Improves Document Structure Recognition", *International Journal on Document Analysis and Recognition (IJDAR) 11*, 2, November 2008, p. 97–109.

[8] S. Macé, E. Anquetil, "Eager interpretation of on-line hand-drawn structured documents: The DALI methodology", *Pattern Recognition 42*, 12, 2009, p. 3202 – 3214, New Frontiers in Handwriting Recognition.

[9] S. Macé, E. Anquetil, "Exploiting on-the-fly interpretation to design technical documents in a mobile context", *Journal on Multimodal User Interfaces 4*, 2011, p. 129–145.

[10] S. Quiniou, M. Cheriet, E. Anquetil, "Error handling approach using characterization and correction steps for handwritten document analysis", *International Journal on Document Analysis and Recognition (IJDAR) 15*, 2012, p. 125–141.

Books and Monographs

[11] S. Calabretto, B. Coüasnon, L. Goeuriot, S. Barrat (editors), *CORIA 2016 - Conférence en Recherche d'Informations et Applications- 13th French Information Retrieval Conference. CIFED 2016 Colloque International Francophone sur l'Ecrit et le Document, Toulouse, France, March 9-11, 2016, Toulouse, France, March 9-11, 2016*, ARIA-GRCE, 2016.

[12] S. Calabretto, B. Coüasnon, *De l'analyse du manuscrit à la recherche d'information dans les réseaux sociaux*, *Numéro spécial de Document Numérique*, *19*, 2-3, Lavoisier, 2016, http://dn.revuesonline.com/resnum.jsp?editionId=3184.

Doctoral dissertations and "Habilitation" theses

[13] M. Bouillon, *Apprentissage actif en-ligne d'un classifieur évolutif, application à la reconnaissance de commandes gestuelles*, PhD Thesis, Insa de Rennes, March 2016.

[14] C. Carton, *Inférence semi-automatique et interactive de règles avec ou sans vérité terrain pour la reconnaissance de structure de documents*, PhD Thesis, Insa de Rennes, March 2016.

Articles in referred journals and book chapters

[15] M. Bouillon, E. Anquetil, "Online Active Supervision of an Evolving Classifier for Customized-Gesture-Command Learning", *Neurocomputing journal*, Accepted, to appear in 2017.

[16] C. Carton, A. Lemaitre, B. Coüasnon, "Eyes Wide Open: an interactive learning method for the design of rule-based systems", *International Journal on Document Analysis and Recognition (IJDAR)*, To appear in 2017.

[17] D. Simonnet, E. Anquetil, M. Bouillon, "Multi-Criteria Handwriting Quality Analysis with Online Fuzzy Models", *Pattern Recognition*, In minor Revision (2016).

[18] D. Simonnet, E. Anquetil, "Handwriting Quality Analysis of Block Letters and Cursive Words.", *Handwriting Today, Journal of the National Handwriting Association*, 15, Autumn 2016, p. 15–21.

[19] C. Wang, H. Mouchère, A. Lemaitre, C. Viard-Gaudin, "Online Flowchart Understanding by Combining Max-margin Markov Random Field with Grammatical Analysis", *International Journal on Document Analysis and Recognition (IJDAR)*, In minor Revision (2016).

Publications in Conferences and Workshops

[20] E. Anquetil, N. Girard, D. Simonnet, "IntuiScript : suivi et aide à l'apprentissage de l'écriture", *in : International Symposium for Educational Literacy*, Ajaccio, Corse, France, June to appear in 2017.

[21] N. Bonneton-Botté, E. Anquetil, E. Jamet, S. Fleury, "Apprendre à écrire des lettres cursives sur tablette numérique: rôle du feedback", *in : International Symposium for Educational Literacy*, Ajaccio, Corse, France, June to appear in 2017.

[22] S. Y. Boulahia, E. Anquetil, R. Kulpa, F. Multon, "HIF3D: Handwriting-Inspired Features for 3D Skeleton-Based Action Recognition", *in : 23rd International Conference on Pattern Recognition (ICPR 2016)*, IEEE (editor), Cancun, Mexico, December 2016, `https://hal.archives-ouvertes.fr/hal-01376113`.

[23] S. Y. Boulahia, E. Anquetil, R. Kulpa, F. Multon, "3D Multistroke Mapping (3DMM): Transfer of hand-drawn pattern representation for skeleton-based gesture recognition", *in : 12th IEEE Conference on Automatic Face and Gesture Recognition*, Washington, DC, USA, May to appear in 2017.

[24] C. Carton, A. Lemaitre, B. Coüasnon, "Inférence semi-automatique et interactive de règles sans vérité terrain", *in : CORIA 2016 - Conférence en Recherche d'Informations et Applications- 13th French Information Retrieval Conference. CIFED 2016 Colloque International Francophone sur l'Ecrit et le Document, Toulouse, France, March 9-11, 2016, Toulouse, France, March 9-11, 2016.*, p. 397–412, 2016, `http://asso-aria.org/coria/2016/CIFED/15.pdf`.

[25] Z. Chen, E. Anquetil, H. Mouchère, C. Viard-Gaudin, "The MUMTDB dataset for evaluating simultaneous composition of structured documents in a multi-user and multi-touch environment", *in : 15th International Conference on Frontiers in Handwriting Recognition*, Shenzhen, China, October 2016, `https://hal.archives-ouvertes.fr/hal-01398762`.

Internal Reports

[26] N. Ragot, E. Anquetil, "The Prototyping and Focused Discriminating Strategy for Pattern Recognition and one Instantiation: the MELIDIS System", *Research report*, Université François Rabelais Tours, LI (EA 6300), June 2016, https://hal.archives-ouvertes.fr/hal-01332566.