Activity Report 2017

Team INTUIDOC

Intuitive User Interaction for Documents

D6 – Media and Interactions

Contents

1 Team

Permanent Members

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

Jean Camillerapp [Emeritus Professor]

Bertrand Coüasnon [Associate Professor, Insa, HDR]

Nathalie Girard [Associate Professor, Uni. Rennes 1]

Aurélie Lemaitre [Associate Professor, Uni. Rennes 2]

Ivan Leplumey [Associate Professor, Insa]

Yann Ricquebourg [Associate Professor, Insa]

Fixed-term positions

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

Kwon-Young Thomas Choi [Insa PhD student, CD grant]

Omar Krichen [Insa PhD student, from October 1st]

Mickaël Renault [Insa Research Engineer]

Damien Simonnet [Insa Post-Doc]

Simon Corbillé [Insa Research Engineer, from October 1st]

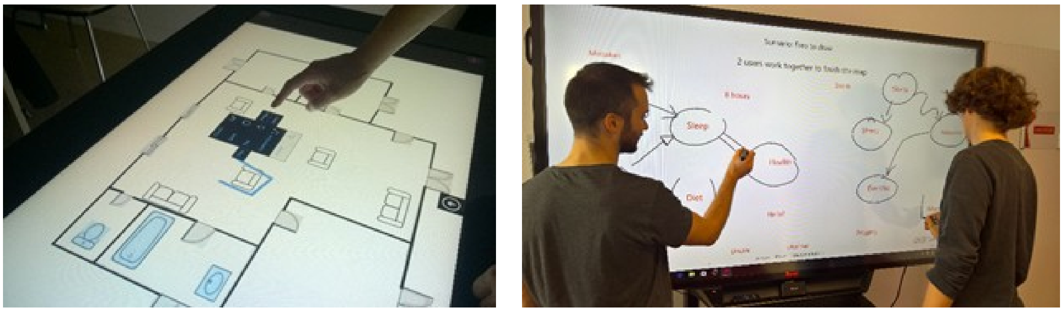
Administrative Assistant

Nadia Saint-Pierre [Inria]

2 Overall Objectives

The research topics of IntuiDoc concern the written communication and the engineering of documents under various aspects: analysis, recognition, composition, interpretation and also graphical/gestural man-document interaction. This research relates to the handwriting and the documents under different forms: manuscript, printed paper form, pen-based and touch-based interaction, graph, images, heterogeneous documents, etc. The roadmap of the IntuiDoc team is on the frontier of several research axes: Pattern recognition, Machine-Learning, Human-Machine Interaction, Uses and Digital Learning. The aim is to explore new scientific challenges of the domain of the Human-Document Interaction with a specific focus on interactive, incremental and evolving learning based on the integration of the user in all the processes of analysis and decision making. Today, four major emerging scientific axes are investigated with strong partnerships with national and international laboratories and companies:

- "On-line" evolving cross-learning of 2D (touch and pen –based) and 3D gestures (Kinect and Leap Motion);
- "On-line" analysis of drawing, sketching and handwriting with pen-based tablet for digital learning (e-education);
- Interactive learning of document structure without ground-truth;
- Document collection analysis for big-data.



(a) (b)

Figure 1: Multi-Touch and Multi-User Interaction

3 Scientific Topics

3.1 On-line evolving cross-learning of 2D and 3D gestures.

3.1.1 2D evolving recognizer for gesture commands

With the increasing use of touch and pen-based sensitive screens, human-computer interactions are evolving. New interaction methods have been designed to take advantage of the new potential of interaction offered by these interfaces. Among them, a new concept has recently appeared: to associate commands to gestures. Those gesture commands enable users to execute various actions simply by drawing symbols. This new man-machine interaction can be used for on-line composition of complex documents such as electrical sketches or floor plan. In order to use such gesture commands, a recognition system is required. For users to easily memorize more than a dozen of gesture commands, it is important to enable gesture set customization. The classifier used to recognize drawn symbols must hence be customizable, able to learn from very few data, and evolving to learn new classes on-the-fly and improve during its use. The objective of this work is to obtain a gesture command system that cooperates as best as possible with the user, learning from its mistakes without soliciting the user too often. Gesture commands lead to a cross-learning situation where the user has to learn and memorize the gestures, and the classifier has to learn and recognize drawn gestures. We study the impact of different strategies to supervise the online training of an evolving recognizer for gesture commands, and how to optimize this cooperation between the user and the recognition system. In particular, we design an inner confidence measure to solicit the user when some data samples don't fit the classifier model, and that it will be very gainful to learn from it.

3.1.2 Multi-touch gesture recognition

Due to the recent prevalence of multi-touch devices, multi-touch gesture recognition has gained a large interest in the last decade. Unlike mono-touch gesture recognition which tracks the movement of a single point of input, multi-touch gesture often tracks many points of contact in parallel as they appear, move and disappear. The recognition for multi-touch gestures is challenging because of the complex chronological relation between the fingers' trajectories. We are going to explore new methods for modelling the shape, relative temporal and motion information in multi-touch gesture by a model of graph and graph embedding approach. In our future work we aim at developing a strategy to detect the pattern of multi-touch gesture at runtime, to be able to address direct manipulation by command gesture.

3.1.3 Multiple users freely-drawn sketch recognition and 3D action gesture recognition

Another scientific challenge is also to address large multi-touch display that allows multiple users to simultaneously interact in the same context and work together. Indeed, many researches and commercial products propose tangible interfaces which support simultaneous participations of multiple users. This is a really new research topic to automatically recognize and interpret in real time the freely-drawn sketch of multiple users.

Finally, in this axe, we investigate the validity of transferring the expertise on hand-drawn symbol representation to recognize 3D action gesture. This new research topic will be conducted in collaboration with MIMETIC project team of Inria. We base this proposition on the observation that patterns produced by a human motion, in particular 2D hand-drawn symbols and 3D actions, share several important properties. They are both governed by kinematic constraints that must be considered while modeling such human motions. We hypothesize that both recognition problems could be addressed in similar ways.

3.2 "On-line" analysis of drawing, sketching and handwriting for e-education

3.2.1 Handwriting analysis for digital learning at school

This axe has been launched with the IntuiScript project founded by the French government as part of innovative national projects (BPI). IntuiScript targets towards offering an advanced digital writing experience at school by using tablets 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:

- it allows children to work in autonomy with an on-line and real time feedback;
- it proposes automatically pedagogical exercises that are adapted to children difficulties based on the automatic analysis of children writing;
- it provides a precise off-line analysis of children writing (i.e. order, direction, shape) to help teachers to understand children writing skills and difficulties.

This project is based on a user-centered design approach that includes several cycles of conception followed by experiments. Therefore, feedback of children and teachers related to these experiments will be used to improve the next version of the education scenario. More than 1,000 primary school students from Brittany take part to an experiment in the project.

Figure 2: Handwriting analysis for digital learning

The scientific problem we tackle here is to quantitatively evaluate a cursive handwriting with respect to a reference model and recommendations of a teacher. In order to be able to teach children how to write, we must be able to analyse their handwriting, to evaluate if the letters, words, sentences are correctly written, and to detail which aspects of the child handwriting do not correspond to the teacher models (corrective feedback). This problem is completely different from the classical task of character recognition, where the challenge is to determine to which class the data samples belongs.

Our objective is to be able to analyse, qualify and evaluate handwriting, with regards to reference models, and for multiple distinct aspects like: shape (for legibility), drawing direction and order (for ductus), speed and fluidity for instance. We use an analysis system based on an evolving fuzzy classifier. It allows to easily define reference models from few data samples to customize "on the fly" the writing exercises to the children. Then, the analysis system can be used to evaluate drawn gestures, regarding a specific feature set, and finally give a confidence score.

3.2.2 Joint laboratory (LabCom: Script&Labs) between IntuiDoc and Script&Go

As we mentioned before, this project is a great success from the academic, pedagogical, industrial and scientific points of view. That is why we have extended this work on "on-line" analysis of drawing, sketching and handwriting by designing new e-education projects based on pen-based tablet devices for primary, middle and upper school. In that way, to ensure the value creation from the research activities, we have built in 2017 the joint laboratory (LabCom: Script&Labs) between IntuiDoc and Script&Go Company on this promising topic.

3.2.3 "On-line" analysis of drawing for new learning strategies based on "generative drawing"

One of the key topics in this axe is how to encourage new learning strategy based on "generative drawing" using pen-based numerical devices. The goal is to improve the learning of students by considering "learning" as a generative activity. In this scope, the potential induced by pen-based tablet is really interesting. The goal is to investigate how we can automatically generate intelligent "corrective" or "predictive" feedbacks to a user during his drawing process: summarizing, mapping, drawing, sketching... We want to explore this new research area in collaboration with researchers in psychology of the LP3C/LOUSTIC Laboratory of Rennes. To support this multi-disciplinary challenge, we have developed the new innovative national project "ACTIF" (BPI – e-FRAN) with the support of the Brittany Region.

3.3 Interactive learning of document structure without ground-truth

3.3.1 Interactive Rule Inference

We work on the interactive learning of document structure, in the context of a thesis that has just ended. This work enables to combine statistical methods with syntactical approaches (grammars). Indeed, statistical methods are not able to convey two-dimension hierarchical structures that are common in document analysis. On the opposite, rule-based syntactical methods often require a fastidious manual step for the specification of the various organizations of the document physical layouts. The objective is to model the logical structure with rules and to learn the physical structure. This learning is based on databases of documents with ground-truth that are really costly to label. The current and future work aims at learning physical properties without ground-truth. The scientific context is to lean on large amount of documents and on generic document system analysis. We want to show that some knowledge can stand out from the repetition of physical structures, thanks to non-supervised learning methods. The challenge is to define strategies to make this learning possible thanks to an interaction with the user, which brings a semantic knowledge to the physical detected elements.

3.3.2 Combination Deep Learning / Syntactical

In collaboration with Richard Zanibbi from the Rochester Institute of Technology (RIT), Rochester, New York, USA, we will continue to work on interactive learning by combining deep learning technics, syntactical analysis and user interaction to introduce learning of segmentation. Deep learning methods like convolutional neural networks or recurrent neural networks have shown very interesting results in recognition by being able to make a common segmentation and recognition, with a good introduction of local context. But they are limited to a local context, which is interesting for the recognition of letters and words in a handwritten text line, but is not enough for a modeling and an understanding of a complex structure like the one we can find in a complex structured document. 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. The objective is to introduce in the architecture of the neural networks the large contextual information and to make the neural networks able to give not only a recognition but also information of localization of the recognized element. Indeed this localization information is important for the syntactical part to continue and explore different solutions in the global recognition of the document. To train the neural networks, we will have to focus also on a semi–automatic generation of datasets and ground truth, made by the grammatical description of the document, in combination with unsupervised clustering and a user interaction to generate ground truth with a minimum of manual work.

3.3.3 Spread Applications

These combinations could open large perspectives by simplifying the grammatical description as much as possible by learning the document structure, including regions of interest (segmentation), region types (classification) and their relationships (parsing/structure). Many applications could be studied on domains where it is important to combine deep learning and strong a priori knowledge. We will also make this combination able to deal with born digital documents (pdf, XML. . .) to address the huge quantities of documents, which need a real understanding for information extraction.

3.4 Document collection analysis for historical big-data

3.4.1 Handwritten Historical Registers

We start 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.

3.4.2 Strategies for Sequential Collection

The DMOS-PI method proposes a framework for the analysis of collections of documents. It enables to share information from the collection between the pages, thanks to an iterative mechanism of analysis. This mechanism also makes it possible to integrate an asynchronous interaction between automatic analysis and human operators. We propose to work on modeling strategies of analysis for the analysis of collections of documents. The strategies could sequence the various iterative treatments of documents pages, the global treatments and the interactions. The interest is to exploit as much knowledge as possible on the collection in order to make the extraction of information in each analyzed pages more reliable, and to make the understanding between the various data at the collection level easier. In this context, the ANR HBDEX project has been selected. It is led by the PSE “Paris School of Economics”

("Ecole d'économie de Paris"), with the LITIS lab in Rouen and the CAMS-EHESS. This project focuses on the extraction of historical big-data for digital humanities, applied to financial data. The objective is to analyse masses of tabular data: daily listing on the Stock Exchange from the 19th and 20th centuries. The analysis will be based on the redundancy between the successive days of listing and the consistency between the global sequences of data. This modeling will enable a fast adaptation to other kinds of historical tabular data that only exist on a paper form (economic, demographic, meteorological), but that is necessary to constitute historical big-data databases. This opens a large possibility of applications on documents found in all statistical institutes.

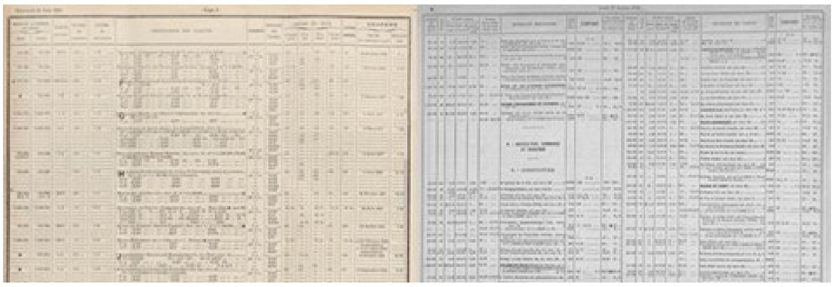


Figure 3: Examples of daily listing on the Paris Stock Exchange

3.4.3 Adaptive Document Layout Analysis

We propose to integrate the interactive document structure learning without ground truth and the collection modeling to generate an adaptive document layout analysis system where a user, with few interactions, could make the recognition system learn new layouts to adapt itself and improve the global recognition quality. We will build this adaptive document layout system on the European project EURHISFIRM (InfraDev). EURHISFIRM designs a world-class research infrastructure (RI) to connect, collect, collate, align, and share detailed, reliable, and standardized long-term financial, governance, and geographical data on European companies. This project is leaded by the PSE "Paris School of Economics" ("Ecole d'économie de Paris"), with seven partners working on quantitative economics and finance, economic and social history, and the LITIS Lab in Rouen working with us on document images analysis. We will work on a system to extract high-quality data from historical serial printed sources, to address three issues: (i) lowering the costs of data extraction from the same source; (ii) lowering the cost of adaptation of the system from one source to the other; (iii) developing effective data validation process. Interactions between the system and experts on the sources lay at the heart of the conception.

4 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).

4.1 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 fifth 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.

4.2 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.

4.3 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

4.4 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*.

4.5 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".

4.6 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] .

4.7 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).

4.8 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 4.7 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.

4.9 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.

4.10 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).

5 New Results

5.1 Segmentation of historical registers

Participants: Aurélie Lemaitre, Jean Camillerapp.

This work focuses on the context of historical registers for a purpose of genealogical researches. This work has been realized in a context of an experimentation of Doptim startup. Their goal is to provide a faster access to historical contents for genealogists. In this work, the*Archives Départementales d'Ille et Vilaine* provided images of parish registers before 1790. Our goal is to build a segmentation of those documents into acts.

The scientific contribution is based on a combination of two sources: some visual clues of text lines combined with rules on the position of text lines inside of a register.

The visual clues on text position are given by an early vision system that detects text lines into a blurred image. The result of this system are pieces of text, or entire text lines.

Then we exploit DMOS method to express some grammatical rules that describes the logical organisation of the document. The difficulties are mainly due to the heterogeneous configurations of the pages: the acts are not always separated by a visible white space, some clues are sometimes present in the margin.

The figure 4 presents some examples of obtained results.

This constitues a preliminary work that demonstrates the complexity of the problem, and shows the interest to recrute a pos-doc position to keep working on this topic. This first work should led to the establishment of a collaboration contract with Doptim society in 2018.

5.2 Text-line localisation in ancient documents

Participants: Bertrand Coüasnon, Aurélie Lemaitre, Jean Camillerapp.

Baseline detection is an open research topic in document analysis and is a preprocessing step for e.g. Handwritten Text Recognition. In this context, we took part to ICDAR 2017 Competition on Baseline Detection (cBAD). The aim of this competition is to evaluate the performance of methods for detecting baselines in archival document images. There were two tracks of participation. The first track deals with the basic baseline detection of handwritten texts in paragraph form. In total 750 pages of handwritten archival documents (no tables or marginalia) with manually annotated baselines and text regions (paragraphs) are prepared. The second track consists of more challenging data including

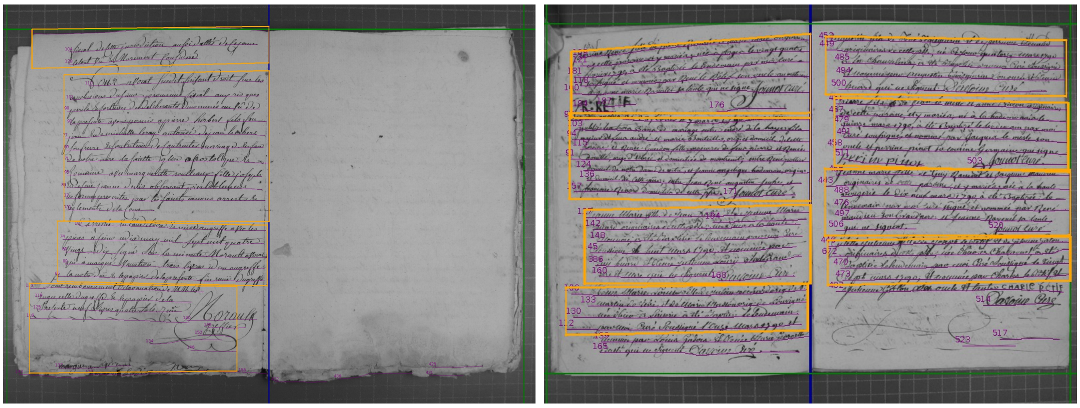


Figure 4: Examples of registers: in purple the clues of text position, in orange the segmentation into acts.

tables, marginalia, and noisy document images. Textlines can be skewed up to 180°. About 1200 pages of archival documents (handwritten and printed documents) have been manually annotated. For both tracks, the images are provided from 9 different archives and document collections.

5.2.1 Baseline Detection, on simple documents

Our baseline detection method is based on a blurred image combined with a description of textlines in the context of the document structure for simple textual documents. This method uses the same blurred image as in the previous method we proposed in [LCC14], but now focused on the lower edges of textlines detected in the blurred image combined with connected components. Hypotheses of textlines produced by this first step are then combined according to a description of textlines defined in EPF, using the DMOS-PI method [LCC10]. This combination build textlines by assembling textlines hypotheses, following rules on their contextual alignments.

5.2.2 Baseline Detection, on complex documents

Our method for complex documents is an extension of the method for simple documents. It uses the same hypotheses of textlines using the blurred image. The description of the structure is also done in EPF, using the DMOS-PI method [LCC10]. The recognition process starts by the detection of table structure, with the description in EPF of simplified table structures limited to columns, based on vertical and horizontal line segments, detected with a Kalman filtering method. Those vertical line segments and all the horizontal line segments (rulings) are removed from the image, and then the blurred image is computed to extract the hypotheses of textlines (without hypotheses that could have been produced by rulings). Textlines hypotheses are cut by the detected vertical rulings. The description of the content of the table structures allows to build textlines limited to each column. When no table is detected, the textline description is used to build textlines in the same way it is done on simple documents.

[LCC14] A. Lemaitre, J. Camillerapp, B. CoÃ¼asnon, "Handwritten text segmentation using blurred image", *in: Document Recognition and Retrieval XXI*, 2014. to appear.

[LCC10] A. Lemaitre, J. Camillerapp, B. Coüasnon, "Interest of perceptive vision for document structure analysis", *in: Human Vision and Electronic Imaging XV*, San Jose, California, USA, January 2010.

5.2.3 Results

The results presented in [DKF+17] show that we obtain a F-value of 88% on track A (simple documents) and of 73% on track B (complex documents).

[1] Aurélie Lemaitre, Jean Camillerapp, Bertrand Coüasnon. Handwritten text segmentation using blurred image. DRR – Document Recognition and Retrieval XXI, Jan 2014, San Francisco, United States. 2014.

[2] Aurélie Lemaitre, Jean Camillerapp, Bertrand Coüasnon. Interest of perceptive vision for document structure analysis Human Vision and Electronic Imaging XV, Jan 2010, San José, United States. 2010

5.3 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 7.2.1). 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 this work, we propose to use the combination of a syntactical method with deep learning architecture to solve the segmentation and classification of broken and touching symbols in music scores 5.3.

(a) Broken symbol (b) Touching symbols

Figure 5: Examples of segmentation problems

Related Work 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.

In another hand, Deep Learning techniques like convolutional neural network are very effective in order to produce accurate object detection models [FKS+17]. But their use of contextual information is limited to a relative local scope, limiting their ability to recognize complex deep document structure.

[DKF+17] M. Diem, F. Kleber, S. Fiel, B. Gatos, T. Grüning, "cBAD: ICDAR2017 Competition on Baseline Detection", *in: 2017 14th IAPR International Conference on Document Analysis and Recognition*, p. 1355–1360, 2017.

[FKS+17] A. Fathi, A. Korattikara, C. Sun, I. Fischer, J. Huang, K. Murphy, M. Zhu, S. Guadarrama, V. Rathod, Y. Song, Z. Wojna, "Speed and accuracy trade-offs for modern convolutional object detectors", Honolulu, Hawaii, 2017.

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.

Datasets 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. Using DMOS and simple music symbol classifier, we produced a dataset of 2025 music symbols of four different classes and we also generated a lot of reject data that are thumbnails containing no relevant symbols. The dataset was entirely manually checked and ground-truth errors should now be really rare. We are also in the process of selecting around 100 different complex orchestral printed scores published between 1750 and 1950. Our project is to experiment new semi-supervised or unsupervised methods with very little manual annotations in order to produce a reliable ground for this new dataset.

Results In [21], we propose a convolutional neural network single symbol detector by training explicitly an attention mechanism to localize precisely an accidental. We improved the results of this simple detector by augmenting artificially our dataset by randomly sampling the position of the cropped image around the accidental. We also improved the architecture of the detector by integrating some a-priori knowledge of the position of another symbol into the architecture. Using this architecture, we obtained a mAP of 89.7% with an IoU threshold of 0.75. We also experimented state-of-the-art deep learning object detectors like Faster R-CNN, R-FCN and SSD producing respectively 98.3%, 98.7% and 97.8% mAP with an IoU threshold of 0.75. In a collaborative work with Alexander Pacha from Institute for Software Technology and Interactive Systems in Vienna Austria, we also applied these detectors to the newly produced MUSCIMA++ dataset that contains symbol level annotation of 140 pages of handwritten music scores. Preliminary results shows that the best performing model is producing a mean average precision of 80% with a vocabulary of 71 very diverse classes of music symbols. These results will be presented in [PEC+ar].

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

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

This research work is part of a thesis conducted in collaboration with MIMETIC project team of Inria. The main goal of this thesis is to investigate the validity of transferring the expertise on hand-drawn symbol representation to recognize 3D gestures. We base this proposition on the observation that patterns produced by a human motion, in particular 2D hand-drawn symbols and 3D gestures, share several important properties. In this work we particularly consider the modelling and recognition of 3D dynamic hand gestures.

Over the past few years, advances in commercial 3D sensors have substantially promoted the research of dynamic hand gesture recognition. On a other side, whole body gestures recognition has also attracted increasing attention since the emergence of Kinect like sensors. One may notice that both research topics deal with human-made motions and are likely to face similar challenges.

In particular, our aim is to evaluate the applicability of an action recognition feature-set to model dynamic hand gestures using skeleton data. Furthermore, existing datasets are often composed of pre-segmented gestures that are performed with a single hand only. We collected therefore a more challenging dataset, which contains unsegmented streams of 13 hand gesture classes, performed with

[PEC+ar] A. Pacha, H. Eidenberger, K.-Y. Choi, B. Coüasnon, Y. Ricquebourg, R. Zanibbi, "Handwritten Music Object Detection: Open Issues and Baseline Results", *in: Proceedings of the 13th IAPR International Workshop on Document Analysis Systems*, Vienna, Austria, 2018. To appear.

either a single hand or two hands. The overall process opf the proposed appraoch is illustrated in Figure 6.

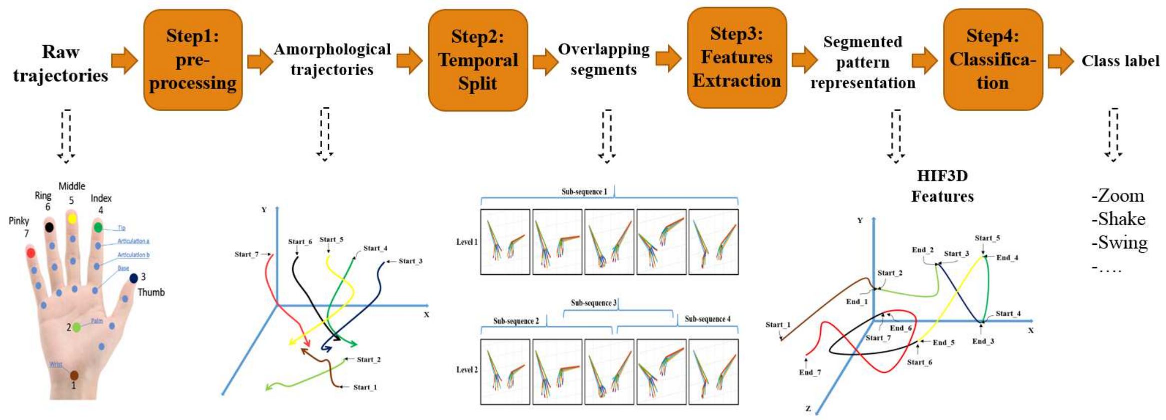


Figure 6: Illustration of the major steps constituting our dynamic hand gesture recognition approach

5.4.1 Proposed dynamic hand gesture recognition approach

Instead of representing a hand gesture with low-level features such as raw positions or angles, we propose to extract the higher level HIF3D features on the hands trajectories. In fact, similarly to the whole body, hands are a set of hierarchical joints. Despite of the trajectories amplitude produced either by the whole body or with hands only, the resulting pattern could be only considered as a 3D motion produced by a human. We therefore propose to draw a parallel between fingers and whole body joints trajectories so as to exploit the valuable improvement in terms of whole body modelling. This is particularly interesting as the employed HIF3D features are themselves the result of a transfer operated from 2D hand-drawn symbols to model 3D whole body gestures. By doing so, we want to go a step further toward the fusion of ways human-made trajectories are modelled.

We therfore consider as input data the raw 3D positions of each finger tip along with the palm and wrist positions (Figure 6-Step1). The positions of each such joint are provided in a reference centred on the recording device (for instance the Leap Motion). When performing a given hand gesture, the successive positions of each joint constitute a 3D trajectory. These trajectories are then assembled to form a single 3D pattern S consisting in a single 3D trajectory. The obtained 3D pattern is then passed to the HIF3D features extractor [BAKM16]. The resulting features report high level information about both local and global shape of the 3D pattern (Figure 6-Step3). HIF3D descriptors are in fact resulting from a improved way of transfering handwriting feature-set, an initial transfer study was presented in [18].

Furthermore, as HIF3D features allow only to express the hand shape variation, i.e. the spatial information, we need to additionally capture the temporal information. In fact, some gestures, such as reversed gestures which have the same spatial characteristics, could only be distinguished based on their temporal sequencing information. We therefore extract HIF3D features according to the so called Temporal Pyramid (TP). This is a commonly used technique in hand gestures recognition and consists in extracting features on overlapping sub-sequences resulting from a temporal split of the global gesture

[BAKM16] S. Y. Boulahia, E. Anquetil, R. Kulpa, F. Multon, "HIF3D: Handwriting-Inspired Features for 3D skeleton-based action recognition", *in: Pattern Recognition (ICPR), 2016 23rd International Conference on*, IEEE, p. 985–990, 2016.

sequence. The temporal split could be performed at different levels and we propose to conduct it at only two levels as illustrated in Figure 6-Step2. As far as the classification is concerned (Figure 6-Step4), we employ the support vector machine (SVM) classifier. This supervised learning classifier is widely used in computer vision classification problems.

5.4.2 Experimental results

We present lastly some results obtained on an existing dataset, namely DHG dataset. Different protocols could be conducted on this dataset. We retain two protocols that consider only 3D skeleton data. According to the first protocol, gestures are grouped into 14 classes such that the number of fingers is not taken into account. In the second protocol, gestures are grouped into 28 classes such that two gestures of the same motion but are performed respectively with one finger and whole hand, belong to two different categories. The second protocol is obviously more challenging as the inter-class variability is lower than in the first protocol.

To allow a fair comparison with previous methods, we use 1960 instances for training the recognition model and 840 instances for evaluation. Table 1 reports the results of our approach along with those of previous methods in the cases of 14 and 28 gestures .

Method	**14 gestures (%)**	**28 gestures (%)**
HoWR [De Smedt et al.,2016]	35.61	-
SoCJ [De Smedt et al.,2016]	63.29	-
HoHD [De Smedt et al.,2016]	67.64	-
- [Oreifej et al.,2013]	78.53	74.03
- [Devanne et al.,2015]	79.61	62.00
SoCJ + HoHD [De Smedt et al.,2016]	82.29	-
- [Guerry et al.,2017]	82.90	71.90
SoCJ + HoHD + HoWR [De Smedt et al.,2016]	83.07	80.00
- [Ohn-Bar et al.,2013]	83.85	76.53
- [De Smedt et al.,2016]	88.24	**81.90**
Our	**90.48**	80.48

Table 1: Comparison between our approach and previous approaches considering 14 and 28 gestures on DHG dataset.

Several conclusions could be drawn based on the achieved results. First, we have shown that recognizing 3D actions and 3D dynamic hand gestures share similar properties as they have been addressed similarly. Results also attest to the merits of HIF3Ds in modelling hand gestures and also the interest of considering high level features instead of using raw data or low level features. Last, it is important to note that our approach uses a subset of 7 joints out of 22 and an interesting perspective would be to study the impact of considering more joints as done by previous methods. More details are provided in the associated conference paper [19].

Future work will probably focus on the segmentation challenges highlighted during our experiments by applying a multiscale approach in which sliding windows of a couple of different lengths would be used.

5.5 IntuiScript project: Handwriting Quality Analysis

Participants: Damien Simonnet, Eric Anquetil, Mickael Renault, Nathalie Girard.

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. The partners are Script&Go

- Learn&Go companies, educational experts (ESPE), Brittany region and LOUSTIC laboratory. It targets the introduction of innovative services and digital contents in the development of fundamental skills at school (see section 6.5). 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, cursive letters, block digits, and preparation to the cursive letter) and on the development of new modules such as stylus pressure learning. Results have been published in the journal *Pattern Recognition* [14], in the international conference *18th International Graphonomics Society Conference (IGS2017)* [22] and in the *International Symposium for Educational Literacy (SILE/ISEL)* [16, 17]. Moreover, the project and the current work on feedback production have been presented during the *workshop EduIHM* ("IHM pour l'Éducation") [25]. Finally, a research paper entitled *Diagnostic of Children Cursive Handwritten Words for e-Education* has been submitted to *Patter Recognition Letters Journal - Special Issue on "Graphonomics for e-citizens: e-health, e-society, e-education"*.

5.5.1 Writing analysis

Different works have shown the importance of feedback in real time [KD96,Shu08], therefore in *IntuiScript* a special attention is paid to this aspect. **In the notebook, each module gives an on-line and real time feedback with a colour-scale indicator** (see Figure 7(a) to Figure 7(d)) that is easy to understand for children. Other visual feedback are accessible to children, as an example the letter model is automatically played when they start an exercise, they can replay this dynamic model as well as their letter when they want. Despite the user-friendliness of the colour-scale feedback, its computation is based on fundamental writing analysis characteristics.

In the literature, handwriting quality is related to legibility and kinematic [Din15]. The former corresponds to letter *shape* and its associated readiness. The latter concentrates on the writing process (*e.g., order*, *direction*, *fluidity*) that must be efficient, as writing is a fundamental skill that is necessary for learning and using knowledge. During the early learning stage of the writing, *i.e.,* for 5 years old children (target of *IntuiScript*), the focus is mainly made on the legibility of the writings (*shape*). Then, several kinematic aspects are integrated especially the *order*, and the *direction* aspects. The French Ministry of Education defines documents that describe a writing convention which is a recommendation for teacher. Based on these recommendation, in the *IntuiScript* notebook, the handwriting analysis is based on three criteria: *shape*, *order* and *direction*. *Shape* combined HBF49-shape features and fuzzy histogram of orientation [DA13],[14]. *Order* and *direction* criteria are correct if the order and direction of elementary strokes are correct. *Order* is determined by identifying the order of median strokes (*i.e.,* the middle of elementary stroke, see thicker lines in Figure 7(h)) in a gesture based on a median stroke classifier. *Direction* uses the fuzzy histogram of direction features to characterise local changes of directions [14]. Finally, the *multi-criteria* classifier combines the scores c_s, c_o and c_d of the *shape*, *order* and *direction* classifiers, as follows:

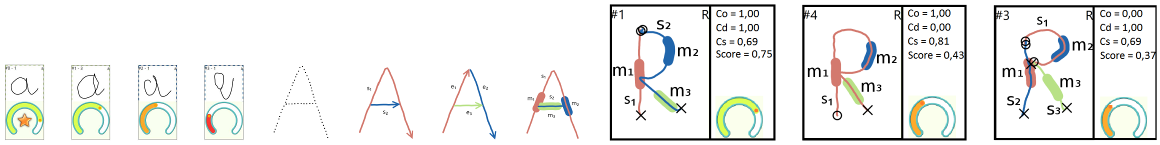
$$\min(\alpha_s c_s, 1.0) \cdot \mathcal{P}(c_o, c_d) \textit{ where } \mathcal{P}(c_o, c_d) = \begin{cases} 1 & \text{if} \quad c_o + c_d = 2 \\ \alpha_1 & \text{if} \quad c_o + c_d = 1 \\ \alpha_0 & \text{if} \quad c_o + c_d = 0 \end{cases} \tag{1}$$

[KD96] A. N. Kluger, A. DeNisi, "The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory", *Psychological Bulletin 119*, 1996, p. 254–284.

[Shu08] V. J. Shute, "Focus on formative feedback", *Review of Educational Research 78*, 2008, p. 153–189.

[Din15] L. H. Dinehart, "Handwriting in early childhood education: Current research and future implications", *Journal of Early Childhood Literacy 15*, 2015, p. 97–118.

[DA13] A. Delaye, E. Anquetil, "HBF49 feature set: A first unified baseline for online symbol recognition", *Pattern Recognition 46*, 1, January 2013, p. 117–130, `https://hal.archives-ouvertes.fr/hal-00933509`.



(a) (b) (c) (d) (e) (f) (g) (h) (i) (j) (k)

Figure 7: The on-line feedback given with a colour-scale indicator drawn under each gesture: very good (green with yellow star) (a), good (green) (b), average (orange) (c) and incorrect (red) (d). An original gesture from a digital device (e), and its decomposition in *strokes* (delimited by pen up) (f) and *elementary strokes* (S_n) (g) with the drawing direction and order. *Median strokes* are represented by thicker lines in (h). And the scores and sub-scores computed from different Block letters in (i), (j), (k).

where $\alpha_1 = 0.49$, $\alpha_0 = 0.35$ and $\alpha_s = 1.1$. This output score belongs to $[0, 1]$ and is then linked to the colour-scale indicator (see Figure 7(i) to Figure 7(k)).

5.5.2 Experiments and first observations

The project is based on a user-centred design approach that involves several cycles of design followed by in-class test campaigns[1]. For the project, three levels of in-class experiments are scheduled. The first level aims for validating and consolidating in a progressive manner each of the designed modules (one by one). The in-class experiments last 1 to 2 hours in the presence of children and teachers. This level of experiments is followed by a consolidation step allowing the improvement of the notebook in order to conduct the second level of tests in good conditions. The second level of tests, in short-term immersion (1 to 2 weeks), aims for experiments with the *Authoring Mode* and the Personalised Pedagogical Progression, *i.e.,* teachers configure their exercises by their own and, conducted their classes as they currently process during handwriting learning session. After the consolidation step of the second level, the last level of tests is dedicated to long-term immersion (3 months), that will allow an analysis of the notebook usefulness for the handwriting learning (quality, fluency, etc.).

In the context of short-term tests, the use of digital devices is organised around digital workshops of twenty minutes with six tablets. These digital workshops come in addition to the traditional workshop on papers to understand the impact of digital technologies on the handwriting learning.

From the past tests, we have extracted **anonymous information about all writings of around** 465 **children**. Currently **our database contains more than** 27.000 **entries describing different features of children writings**. Each entry corresponds to a digit, a capital or a cursive letter written by a children. Among the different features, the digital tablet allows the observation of the time required to write a specific symbol, the number of strokes used to write it, and also the pressure applied by the children on the stylus all along each drawing. Moreover, the score progression during the writing sequence of a child is recorded and each score can be easily linked to specific criteria in the gesture done.

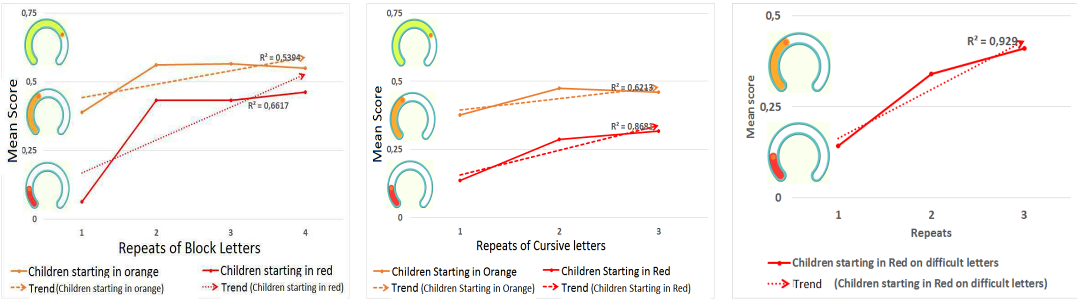
A file with all entries -the letters, anonymous ID, repeat number, age, gender, laterality, duration, number of strokes, average pressure, variance of pressure, scores- is available online[2]. Among all these records, we extracted a subpart according to the score tuning used (see Table 2). The 111 children in average difficulty for the first repeat of a Block letter produced 299 input, corresponding to mistakes on 21 letters, 481 input were produced by the 117 children in average difficulty on cursive letter. 471 block letters with incorrect feedback were produced by the 149 children in high difficulty, representing

[1]`https://vimeo.com/142233890` ; `https://vimeo.com/140660028`

[2]`https://www-intuidoc.irisa.fr/en/children-handwritings-database/`

Table 2: Description of the data extracted.

Module	Complete Database		Analysed Database			
	#Symbol	#Children	#Symbol	#Children	#Children Starting at least 1 time in average difficulty	#Children Starting at least 1 time in high difficulty
Total	27.073	465				
Block Letter	16.692	409	5675	189	111	149
Cursive Letter	10.381	237	6117	145	117	105



(a) (b) (c)

Figure 8: Letter writing progression during an exercise: Block letters (a), Cursive letters (b), difficult cursive letters (c)

19 different letters. And, 502 cursive letters were produced by the 105 children in high difficulty on 21 different letters. Figure 8 presents the observed progression of children during the writing repetition of several block letters Figure 8(a) and cursive letters Figure 8(b) (initial phase of the cursive writing module). Only Children with difficulties on the first repeat are considered (see Table 2). For block letters, the first repeat corresponds to the letter they initially write in the whole word, then the repeats 2 to 4, are the new trainings they have, to improve their writing. On both figures, we can observe that children improve their scores no matter when the repeats are done (see the linear trend -dashed lines- for each curve). We can also observed that children starting with the lower score reach the average score: increase mean of 0.18 of the score - increase rate 9% for cursive letters, and for block letter: increase mean of 0.39 of the score - increase rate 13%. This progression is also observed when we focus only on the difficult letters ([Dum13] see Figure 8; only the children starting in red are depicted since for the others we do not have enough data - reliability index R^2 close to 0). But, when they start in Orange, they reach the good score only for Block letter: increase mean of 0.16 of the score - increase rate 5%. For cursive letter, they increase their score from 0.08 - increase rate 4%.

5.5.3 Conclusion and future works

These results come from short session (only four/three repeats). In future works we will explore the long-term impact of the approach. Moreover, we will explore several solutions to show children the localisation of their mistakes on their drawings. The following test campaign will be dedicated to this

[Dum13] D. Dumont, *Le système d'écriture des minuscules latines manuscrites en usage dans les écoles françaises : intérêt de la prise en compte de ce systéme pour l'enseignement de l'écriture manuscrite*, PdD Thesis, Université Paris-Descartes, 2013.

task: test several types of feedback and analyse the impact on children handwriting improvement.

5.6 CAMIA/KASSIS: a new digital learning environment (workbook) for active learning with pen-based tablet

Participants: Mickaël Renault, Eric Anquetil.

Keywords: Digital Learning, Handwritten annotation, Pen-based Tablet..

CAMIA/KASSIS is a new digital learning environment (workbook) for active learning for face-to-face lectures. It was designed and developed during the last 2 years in collaboration with lecturers. The concept is based on one of the learning keys which is to make the learner active during his learning. The handwritten note-taking is one of the first elements to make the learner active through a continuous effort of synthesis and emphasis which allows a better appropriation of the delivered discourse. The CAMIA/KASSIS solution is based on the networking (autonomous network or existing network) of a set of pen-based hybrid tablet. It is an interactive digital solution based on several innovative concepts that are integrated within the same digital learning ecosystem:

- it offers to students and lecturers a direct handwritten note-taking on the slides to facilitate the composition of formulas, diagrams, sketches, drawings, annotations (going beyond limited keyboard input);
- it offers handwritten advanced graphical quizzes composition for teachers: questions can be expressed on-the fly, by sketching or drawing directly on the tablet display (going beyond rigid and planned interaction based quizzes or push button);
- it offers real time aggregation, analysis and restitution of the responses collected for immediate feedback for students and teacher.

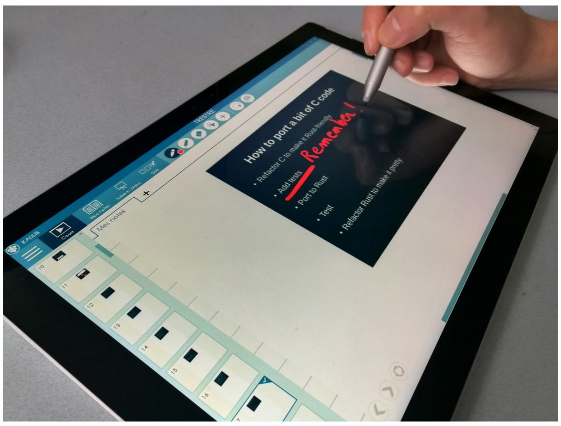


Figure 9: KASSIS IHM

From this application, two research projects have been developed:

- The national project "ACTIF" (BPI – e-FRAN) with the support of the Brittany Region,
- and Labex (CominLabs) eFil (e-Feedback for interactive Lecture).

The former project goal is to go beyond 'classical' active learning methods to improve learning of middle-school students using and evaluating a new digital learning environment based on CAMIA/KASSIS digital learning ecosystem. The partners are LP3C and LOUSTIC laboratories, Learn&Go companies, educational experts (ESPE) and Brittany region.

The latter will focus on higher education in collaboration with the LP3C (University of Rennes 2) and the LS2N (University of Nantes) laboratories. In the eFIL Project, CAMIA/KASSIS digital learning ecosystem will be adapted and tested in classroom settings during lectures. Traces will be collected and analyzed both to explore learning dynamics in classroom, and to provide a dashboard allowing instructors to monitor their activities during lectures as well as in-between sessions.

CAMIA has been transferred to the society *Learn&Go*, which is commercialized under the name *KASSIS*.

5.7 Real-time interpretation of geometric shapes for digital learning

Participants: Omar Krichen, Eric Anquetil, Nathalie Girard, Mickaël Renault, Simon Corbillé.

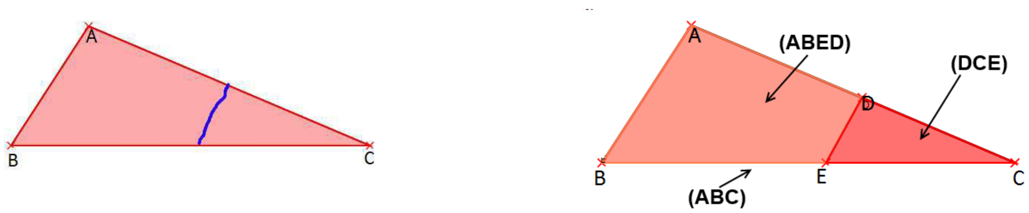
Keywords: Online recognition, Hand-draw stroke analysis, Digital learning, gestures and stylus interaction.

5.7.1 Real-time interpretation of geometric shapes for digital learning

This work is in the context of the *e-FRAN ACTIF* project (*c.f* section 6.2), which aims to use pen-based tablets in an educational context to foster active and collaborative learning in French middle-schools. The partners are LP3C and LOUSTIC laboratories, Learn&Go company, educational experts (ESPE) and Brittany region.

The goal is to allow students to draw geometric shapes on the fly, given a teacher's instruction. The pattern recognition and analysis system has to recognize on the fly the user's drawings in order to produce real time visual, corrective and guidance feed-backs.

We base our work on the visual grammar CD-CMG [8] (Context Driven Multi-set Grammar), to model the domain knowledge and interpret the hand-drawn sketches on the fly. This generic formalism has been applied on various forms of documents, such as electrical sketches or architectural plans. We adapted this grammar to the Geometry domain to cover the concepts taught in middle-school. Even though the formalism is expressive enough to model the domain knowledge, the multiple interactions between geometric objects, *e.g* creating sub-figures from existing ones (*c.f* figure 10), generates combinatorics problems in the analysis process. Thus, the formalism could not cope with the exigence of real-time interaction.



(a) Drawn stroke (in blue) (b) 2 triangles, 1 quadrilateral, 8 segments

Figure 10: Sub-figures creation

Our early works focus on extending this formalism, by the addition of new operators that affect the analysis process and reduce the search space size. This allowed the design of a system with acceptable performance in relation to the constraint of real-time user interaction. Table 3 presents the impact of our optimization on a complex figure analysis in terms of number of reduced production rules (iterations, interpretations), time analysis and triggered rules. A research paper describing these early works, entitled "Real-time intepretation of geometric shapes for digital learning" will be presented to the next International Conference on Pattern Recognition and Artificial Intelligence (ICPRAI 2018). Our future works consist in the design of an author mode, where the teacher can create customized exercises. The goal will be to generate the procedure, or drawing plan, from the teacher's production. This will allow to give personalized corrective and guidance feedbacks to the students. The collaboration with LOUSTIC and LP3C laboratories and the users tests conducted in pilot middle-schools will monitor the system's evolution towards an efficient tutoring tool.

Approach	Iterations	Interpretations	Time	Triggered
CD-CMG	42	6	30 s	2472
Extended formalism	7	1	**1.5 s**	167

Table 3: Impact of formalism extension on system performance

A first prototype have already been designed. This is an e-learning system for geometry learning in middle-school. The aim is to offer a personalized pedagogical experience where the child is autonomous. The idea here is to perform a geometry exercise on digital medium by simulating the traditional pen and paper figure sketching. The user can manipulate digital tools like the ruler, the compass or the protractor with theirs fingers while drawing his figures with the stylus. Thus, the learning process is similar to the one used on traditional method. Figure 11 presents the prototype interface. The student is guided by real-time feedbacks. The coupling of free drawing and feedback interventions will foster the child's active learning and improve his performance. Even though it is early ages, the prototype has been validated by pedagogical experts from the Academy of Rennes. Plus, this application was presented at the inauguration of the Living Lab Interactik in Lorient on 30 november. First test of Geometry will be start in pilot middle-schools in 2018. The result of these experimentations will allow to study the impact of this type of application on the learning process and validate our approach.

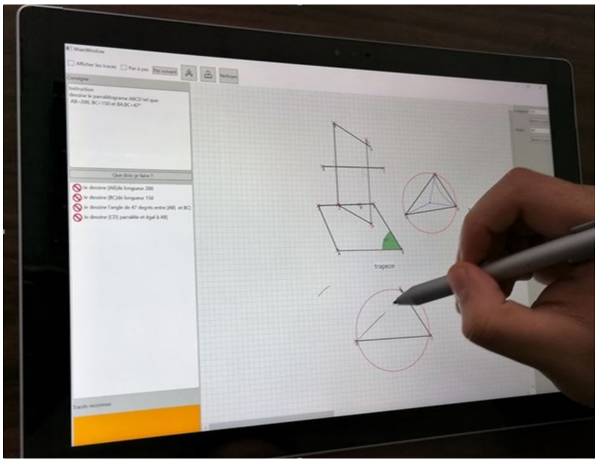


Figure 11: Geometry IHM

6 Contracts and Grants with Industry

6.1 Script&Labs/ANR joint Laboratory (LabCom): On-line handwriting and drawing recognition and interpretation for active Learning in e-education

Participant: Eric Anquetil, Nathalie Girard, Damien Simonnet, Mickael Renault.

- 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…

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

Participant: Eric Anquetil, Nathalie Girard, Yann Ricquebourg, Simon Corbillé.

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

The project "ACTIF" is one of the 22 selected national project from the "e-fran / innovative national project " call. This is a 4 years project (2017-2020) for a total amount of 1 143 856€, 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 is to investigate 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.

6.3 eFil: e-Feedback for Interactive Lecture - Labex CominLabs project

Participant: Eric Anquetil, Nathalie Girard.

- Partner: *LP3C (University of Rennes 2)*, *LS2N (University of Nantes)*
- 24 months (2017-2019).
- Contract: INSA

The e-Fil Cominlabs project is based on CAMIA/KASSIS which is a new digital learning environment (workbook) for active learning with pen-based tablet. The goal of the eFIL Project is to optimize and to evaluate this digital learning environment in classroom settings during lectures (Optimizing with user-centered design). Traces will be collected and analyzed both to explore learning dynamics in classroom, and to provide a dashboard allowing instructors to monitor their activities during lectures as well as in-between sessions. The goal is to explore and analyze traces to monitor learning activities in classroom with Educational Data Mining and Learning Analytics.

6.4 Industrial software licensing with Learn&GO company

Participant: Eric Anquetil.

- Partners: *Learn&GO company*
- Since 2017
- Contract: INSA

The IntuiDoc team has close links with the Learn&GO company for transferring its research results for e-education. This partnership is now based on several licensing agreements. They cover various technologies of the Intuidoc team such as handwriting recognition and analysis.

This partnership is also supported by several collaborative projects: in particular the ANR LabCom "Script&Labs" and the *IntuiScript* BPI Project (http://intuiscript.com/).

6.5 IntuiScript: National Innovative BPI Project

Participants: Eric Anquetil, Mickael Renault, Damien Simonnet, Nathalie Girard.

- Partners: *SCRIPT&GO company*, *Microsoft*, *Academy of Brittany*, *Region of Brittany*, *LOUSTIC laboratory of Rennes*
- 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 three months with a large number of French primary school students in Brittany. 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

7 Other Grants and Activities

7.1 National initiatives

7.1.1 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.

7.2 International initiatives

7.2.1 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 5.3).

8 Dissemination

8.1 Leadership within scientific community

8.1.1 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 Workshop on Graphics Recognition and Graphical Document Analysis (GREC 2017).
- B. Coüasnon is Associate Editor of Frontiers in Cultural Heritage Digitization.
- B. Coüasnon is member of the program committee of the 4th International Workshop on Historical Document Imaging and Processing (HIP 2017).
- B. Coüasnon is member of the program committee of the Doctoral Consortium of the International Conference on Document Analysis and Recognition 2017 (DC ICDAR 2017).
- E. Anquetil, B. Coüasnon and A. Lemaitre are members of the program committee of the International Conference on Document Analysis and Recognition (ICDAR 2017).

8.1.2 Reviewing

- E. Anquetil is a reviewer in 2017 of:
  - Pattern Recognition Letters.
- A. Lemaitre is a reviewer in 2017 of:
  - IJDAR (International Journal on Document Analysis and Recognition).
- N. Girard is a reviewer in 2017 of:
  - ImaVis (Image and Vision Computing);
- B. Coüasnon is a reviewer in 2017 of:
  - Project funding "Partenariats Hubert Curien" (PHC).

8.1.3 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 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 scientific head of the Media and Interactions Department of IRISA.
- B. Coüasnon is member of the scientific board of Irisa.

8.1.4 Participation to PhD and HDR defenses

- E. Anquetil was member of the HDR committee of the HDR of Nicolas Ragot, Contributions à la reconnaissance de formes et applications à l'analyse de l'écrit et des documents, Université de Tours, juin 2017.
- E. Anquetil was the president of the PhD committee of Marc Dupont, Glove-based gesture recognition for real-time outdoors robot control, Université de Bretagne-Sud, Janvier 2017.
- E. Anquetil was a reviewer of the PhD of Marion Morel, Modélisation de séries temporelles multi-dimensionnelles. Application à l'évaluation générique et automatique du geste sportif, Université Paris 6, Novembre 2017.
- N. Girard was a member of the PhD committee of the PhD of Ngoc Bich Dao, Réduction de dimension de sac de mots visuels grâce à l'Analyse Formelle de Concept, Université de La Rochelle, juin 2017.

8.2 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 N. Girard give lectures at Master-Research *d'informatique* of University of Rennes 1.
- E. Anquetil is in charge of the module "Analysis, Interpretation and Recognition of 2D (touch) and 3D Gestures for New Man-Machine Interactions" (AIR) of the Master-Research *d'informatique* of University of Rennes 1.

- E. Anquetil is in charge of the module "Motion Analysis and Gesture Recognition (2D / 3D)" (AMRG) of the Computer Science Dept. of INSA Rennes.
- 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.

8.3 Patent and Deposit of digital creations (APP)

8.3.1 Deposit of Digital creations (APP)

- E.Anquetil and M. Renault deposited a new digital creations (APP) on the "CAMIA" technology: IDDN.FR.001.150033.000.S.P.2017.000.10000 (see section 5.6)

9 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.

Articles in referred journals and book chapters

[11] M. Bouillon, E. Anquetil, "Online active supervision of an evolving classifier for customized-gesture-command learning", *Neurocomputing 262*, November 2017, p. 77 – 89, https://hal.inria.fr/hal-01575805.

[12] 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 20*, 2, 2017, p. 91–103, https://hal.inria.fr/hal-01493442.

[13] G. Pirlo, R. Plamondon, E. Anquetil, "Guest Editorial Special Issue on Drawing and Handwriting Processing for User-Centered Systems", *IEEE TRANSACTIONS ON HUMAN-MACHINE SYSTEMS 47*, 2, April 2017, p. 4, https://hal.inria.fr/hal-01591261.

[14] D. Simonnet, E. Anquetil, M. Bouillon, "Multi-Criteria Handwriting Quality Analysis with Online Fuzzy Models", *Pattern Recognition*, 2017, https://hal.archives-ouvertes.fr/hal-01515397.

[15] 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 20*, 2, June 2017, p. 123–136, https://hal.inria.fr/hal-01503366.

Publications in Conferences and Workshops

[16] E. Anquetil, N. Girard, D. Simonnet, "IntuiScript : suivi et aide à l'apprentissage de l'écriture", *in: International Symposium for Educational Literacy (SILE/ISEL)*, Ajaccio, Corse, France, June 2017, https://hal.inria.fr/hal-01551684.

[17] 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 (SILE/ISEL)*, Ajaccio, Corse, France, June 2017, https://hal.inria.fr/hal-01625130.

[18] 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 International Conference on Automatic Face and Gesture Recognition (FG 2017)*, p. 462 – 467, Washington, DC., United States, May 2017, https://hal.archives-ouvertes.fr/hal-01555452.

[19] S. Y. BOULAHIA, E. Anquetil, F. Multon, R. Kulpa, "Dynamic hand gesture recognition based on 3D pattern assembled trajectories", *in: IPTA 2017 - 7th IEEE International Conference on Image Processing Theory, Tools and Applications*, p. 1–6, Montreal, Canada, November 2017, https://hal.archives-ouvertes.fr/hal-01666377.

[20] Z. Chen, H. Mouchère, E. Anquetil, C. Viard-Gaudin, "Early Recognition of Handwritten Gestures based on Multi-classifier Reject Option", *in: 14th IAPR International Conference on Document Analysis and Recognition (ICDAR2017)*, Kyoto, Japan, November 2017, https://hal.archives-ouvertes.fr/hal-01653154.

[21] K.-Y. Choi, B. Coüasnon, Y. Ricquebourg, R. Zanibbi, "Bootstrapping Samples of Accidentals in Dense Piano Scores for CNN-Based Detection", *in: 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR)*, IEEE, Kyoto, Japan, November 2017, https://hal.archives-ouvertes.fr/hal-01712213.

[22] N. Girard, D. Simonnet, E. Anquetil, "IntuiScript a new digital notebook for learning writing in elementary schools: 1st observations", *in: 18th International Graphonomics Society Conference (IGS2017), Proceedings of IGS 2017*, p. 201–204, Gaeta, Italy, June 2017, https://hal.inria.fr/hal-01548200.

[23] M. Mehri, P. Héroux, R. Mullot, J.-P. Moreux, B. Coüasnon, B. Barrett, "HBA 1.0: A Pixel-based Annotated Dataset for Historical Book Analysis", *in: International Workshop on Historical Document Imaging and Processing (HIP)*, Kyoto, Japan, November 2017, https://hal.archives-ouvertes.fr/hal-01637826.

Miscellaneous

[24] B. Coüasnon, A. Lemaitre, "DMOS, It's your turn !", 1st International Workshop on Open Services and Tools for Document Analysis (ICDAR-OST), November 2017, https://hal.inria.fr/hal-01659131.

[25] N. Girard, D. Simonnet, M. Renault, E. Anquetil, "IntuiScript : le cahier numérique pour le suivi et l'aide à l'apprentissage de l'écriture à l'école", GT "IHM POUR L'EDUCATION", August 2017, https://hal.inria.fr/hal-01579798.

[26] A. Lemaitre, B. Coüasnon, "DMOS : A vous de jouer !", Symposium International Francophone sur l'Ecrit et le Document (SIFED), June 2017, https://hal.inria.fr/hal-01534885.