INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Project-Team Qgar

Querying Graphics Through Analysis and Recognition

Nancy - Grand Est

THEME COG

Activity Report

2007

Table of contents

1. Team

Head of project-team

Salvatore Tabbone [assistant professor, Université de Nancy 2, head of project-team since February, HdR]

Administrative assistants

Sandra Bezon [TR INRIA, from November]
Zohra Marzak [TR INRIA, from January to October]
Françoise Laurent [SASU INPL for budget support at INPL/École des Mines de Nancy]

Permanent researchers

Gérald Masini [research scientist, temporarily head of the Computing Resources Department since February 2007]
Karl Tombre [research director, on temporary leave from professor position at INPL/École des Mines de Nancy, HdR]

Faculty members

Suzanne Collin [assistant professor at UHP/ESIAL]
Philippe Dosch [assistant professor at Université de Nancy 2]
Bart Lamiroy [assistant professor at INPL/École des Mines de Nancy]
Benoît Naegel [assistant professor at UHP/IUT SAINT-DIÉ, since September]
Laurent Wendling [assistant professor at UHP/ESIAL, HdR]

Post-doctoral fellow

Oriol Ramos Terrades [ATER at Université de Nancy 2, since September]

Technical staff

Vitor Vasconcelos Araújo Silva [engineer under ODL contract since January]
Laurent Fritz [engineer employed by INPL/École des Mines de Nancy on contract FRESH, until March]

Ph. D. students

Sabine Barrat [CIFRE contract with Netlor Concept, defense planned in 2009]
Oanh Nguyen [joint supervision with IFI (Hanoi, Vietnam), defense planned in 2009]
Jan Rendek [CIFRE contract with France Télécom R&D, defense planned in June 2008]
Salim Jouili [employed by INRIA on Navidomass project, defense planned in November 2010]
Jean-Pierre Salmon [employed by INPL on European project FRESH, defense planned in March 2008]

External collaborators

Matias Mateu Graside [INRIA internship, University of Urugay, visiting Qgar from May to September]
Thai Hoang Van [Master student, MICA Hanoi, visiting Qgar one month in September]
Wan Zhang [PhD student at City University Hong Kong, visiting Qgar from September to December]
Djemel Ziou [professor at Université de Sherbrooke (Canada), visiting Qgar from June to July]

2. Overall Objectives

2.1. Overall Objectives

The Qgar project-team works on the conversion of weakly structured information—an image of a paper document, or a PDF file, for example—into "enriched" information, structured in such a way that it can be directly handled within information systems. Our research belongs to the document analysis field, and more precisely to the graphics recognition community. However, as the semantics (or domain knowledge) of the processed information cannot be fully taken into account, this community is well aware of the fact that recognition alone will not lead to completely automated back-conversion.

We study the use of graphics recognition methods to index and organize weakly structured graphical information, contained in *graphics-rich documents*, such as technical documentation. In this context, we experiment the capacity of pattern recognition methods to compute useful features for indexing and information retrieval. It appears that text-based retrieval (based on annotations or textual references, for instance) must be complemented with the handling of purely graphical information, such as symbols or drawing parts.

3. Scientific Foundations

3.1. Introduction

Our scientific foundations belong in the domain of image analysis and pattern recognition. For many years, the main contributions of our project-team were in the area of algorithms and methods for image analysis and segmentation, with a specific thrust on images of graphics-intensive documents. In the last years, while keeping a regular activity in this domain, we have moved our main effort towards pattern recognition methods, especially for symbol recognition and spotting. But, of course, recognition tasks also require the prior extraction of features, using image processing and segmentation methods.

3.2. Feature Extraction and Segmentation

Keywords: *Binarization*, *image processing*, *text-graphics segmentation*, *vectorization*.

As conversion from pixels to features raises a great deal of problems, our project-team had to design several algorithms and methods for binarization, vectorization or text-graphics segmentation. However, designing new methods, or variants of older ones, is not enough. We also must be able to characterize and to evaluate the performances of the methods we use, to study their robustness and reliability, and to develop stable implementations for them (hence, the focus on software, cf. § 5.1).

Vectorization is the conversion of a binary image into a series of graphical primitives, mainly line segments and arcs of circle, generically called “vectors”, which are a good representation of the original graphics. Existing methods generally suffer from two major drawbacks, over-segmentation and poor geometric precision, especially at the junctions between vectors. We have worked for many years on this matter and proposed several techniques to overcome these limitations. In particular, we defined a new method of detection of arcs and segments extracted from curves. The strong variations of the curvature radius are assumed to be the break points between two primitives.

Performance evaluation is a major concern in document analysis, and more generally in image processing, pattern recognition and computer vision. A way of approaching the problem considers a method to be evaluated as a completely separate module, which is fed with synthetic or real data. Evaluation is then carried out by comparing the results supplied by the module with some ground-truth. The performances of a segmentation method may also be evaluated according to the observed qualities of recognition steps, using the features provided by the method. This is sometimes called “goal-oriented performance evaluation”.

We are actively involved in the organization of performance evaluation campaigns for symbol recognition, at national and international levels. More particularly, our project-team is leader of the Épeires project (cf. § 8.1.1), affiliated to the Techno-Vision program. This project aims at providing a complete environment of performance evaluation, for our own needs—as organizers—but also for the needs of any team working on symbol recognition or using recognition methods.

Performance evaluation of graphics recognition methods is related to several open scientific questions, including intricate problems such as defining simple and non-biased metrics and matching procedures between the ground-truth and the output of recognition methods, when the answer is more complex than a simple “recognized” or “not recognized” label—a good example is the evaluation of a vectorization method. Another potential problem is the generation of large sets of training or benchmarking data, using image degradation models. Such models have to be realistic and as interchangeable as possible with real data. In this perspective we have launched a new joint project with the Computer Center Vision at *Universitat Autonoma de Barcelona* on the characterization and evaluation of shape descriptors (cf. § 6.4).

3.3. Symbol Recognition and Spotting

Symbol recognition is the localization and identification of symbols in documents [30], to get natural features to be used in indexing and retrieval applications. Whereas it has ancient and solid foundations, and has proved to be mature for character recognition, for instance, symbol recognition still remains an open question when dealing with complex symbols having large variations. Our attention is focused on the weaknesses of the existing recognition methods, which make them difficult to use:

- In querying and browsing applications, it is often impossible to work with a database of reference symbols *a priori* known. It is more often the case that the user delineates an arbitrary symbol in a drawing and queries for similar symbols in the set of available documents. We therefore have to design methods able to recognize, or at least to spot "on the fly", without prior learning or precompilation of models.
- We are interested in coping with cases when a prior segmentation of the image is difficult, or even impossible, i.e., when it is necessary to design segmentation-free recognition methods or methods which simultaneously perform segmentation and recognition.
- Efficient methods for recognizing a symbol among 10 or 20 different models are currently available. They work even when the symbol to be recognized is distorted by noise or other touching graphics. However, they do not scale well to the recognition of a number of symbols an order of magnitude larger. Thus, there are both computational complexity issues and open questions about the discrimination power of the methods chosen for recognition.

Signatures are often used for indexing and retrieval purposes, but most work has concentrated on text-based or image-based signatures. Nevertheless, we think that there is also room for graphics-specific signatures to achieve an efficient localization and recognition of symbols, and we currently work in two directions:

- Quick and robust symbol localization through image-based signatures: we propose to combine a feature descriptor method with a structural representation of symbols. We define a robust structural representation based on key points, which allows a quick localization of candidate symbols within documents. Each candidate is then recognized using a combination of shape descriptors. To improve the recognition steps, the processing is directly performed on the grey-level image of the document, without any previous segmentation.
- Structural shape signatures: direct work on the raw image data is not always necessary as, in many cases, vector data can be obtained from available CAD files or similar electronic representations, or can be captured through raster-to-vector conversion. It is therefore interesting to use signatures directly computed on such data.

When dealing with a large number of symbols, both signatures and structural recognition methods may not be powerful enough to discriminate. Combining outputs of classifiers or descriptors is one of the strategies used to improve recognition rates.

4. Application Domains

4.1. Application Domains

Keywords: *Documentation*, *document analysis*, *graphics*, *indexation*, *navigation*.

Our main application domain is the processing and analysis of documents—i.e., information produced by humans for communicating with other humans—which convey a huge amount of information in very "poor" formats: paper documents, or low-level, poorly structured digital formats such as Postscript, PDF or DXF.

We are more specifically interested in graphics-rich documents, typically technical documents containing text mixed with lots of graphics. The usual text-based indexing and retrieval methods are still of interest, but we also need additional ways of accessing the information conveyed by the documents: recurring symbols, connections between textual descriptions and drawing parts, etc. Within this general application area, we work on two major kinds of document analysis applications:

- Specific documentation referring to a well-known framework of technical knowledge: a good knowledge about the type of information which has to be extracted from the documents is usually available, in terms of models of the symbols to be recognized. This is the case of electrical wiring diagrams in aircrafts, like those of the European FRESH project in which we participate (cf. § 7.2).
- Open documentation: few or even no strong assumptions are made on the kind of information to handle. This is typically the case of applications for browsing large sets of heterogeneous documents, when the user provides "on the fly" information about the symbols or structures he is looking for. We are currently working with France Télécom R&D on this topic (cf. § 7.1).

The growing inclusion of cameras in mobile devices provides significant potential for new computing applications. We recently started a collaboration with Netlor Concept (cf. § 7.3), to develop a prototype system to identify city buildings from still images: a central computer system receives images captured by users' mobile devices and identifies them from a large image database including a wide range of viewpoints of each building.

5. Software

5.1. Qgar Software

Participants: Gérald Masini, Vitor Vasconcelos Araújo Silva.

Overview. Since several years, the QGAR project-team has devoted much effort to the construction of a software environment, to be able to reuse whole or part of software implemented during previous work, as well as collected experience. This environment includes three main parts:

- *QgarLib*, a library of C++ classes implementing basic graphics analysis and recognition methods,
- *QgarApps*, an applicative layer, including high-level applications (binarization, edge detection, text-graphics separation, thick-thin separation, vectorization, etc.),
- *QgarGUI*, a graphical interface to design and run applications, providing data manipulation and display capabilities.

The Qgar system is registered with the French agency for software protection (APP) and may be freely downloaded from its web site (http://www.qgar.org). The whole system is written in C++ and includes about 100,000 lines of code, including unit test procedures. A particular attention has been paid to the support of "standard" formats (PBM+, DXF, SVG), high-quality documentation, configuration facilities (using CMake), and support of Unix/Linux and Windows operating systems.

Application management is plugin-based. Each executable binary file is paired with a XML description file which is parsed when the user interface is launched: the corresponding application is then dynamically integrated into the menus of the interface, and dialog boxes to access the documentation and run the application are dynamically generated. In this way, any application may be easily coupled with a remote system based on a similar approach. Conversely, as the integration (or removal) of an application does not imply any modification of the user interface, the installation of remote applications, provided by partners for testing for example, is easy. This is particularly useful when comparing different methods performing the same task, in the context of performance evaluation, a topic which is part of our current research work, as previously mentioned. For these reasons, the system has been and is used within the context of several cooperation projects, like the FRESH European project (cf. § 7.2).

New results. The system reached a stable and matured architecture at the end of 2004, and most of our effort was then devoted to the improvement of the quality of its components, especially library QgarLib and applications QgarApps (input/output, connected components, text-graphic separation, shape context descriptor, etc.), simultaneously with a continuing work of documentation and unit test writing. The last step of this consolidation effort consisted in providing a version which compiles on Linux as well as Windows platforms, using the same scripts in both cases. This new version (3.0) will be issued at the beginning of January 2008:

- The whole source code repository was moved to InriaGforge, using Subversion (instead of CVS) to manage the complete set of files of the software: documentation, configuration and installation scripts, etc.
- The file structure was reorganized and a separate "development library" was created to include algorithms and tools being currently developed by members of the Qgar project-team. This library is distributed without any guarantee of correctness and efficiency, but once a module will be consolidated, it will be transferred to the QgarLib standard library. The source files of the documentation have now the same hierarchical structure as the source files of the software (API, libraries, installation tools, etc.).
- All compilation scripts of the previous version were rewritten, using CMake (instead of autoconf/automake) so that the same set of scripts can be used to compile and install the software on platforms Linux (with compiler gcc) and Windows (with compilers MinGW and Visual Studio).
- For Linux, packages are available to download and automatically install the software. For Windows, an installer executable file (an alpha version may be downloaded from the Qgar site of InriaGforge) was elaborated with the Nullsoft Scriptable Install System (NSIS). A script controls the installation process and provides means to insert the presentation of Qgar licenses, to create shortcuts, and so on.
- The Qgar graphical interface is coded using the graphical toolkit of Qt3, a software by TrollTech. Unfortunately, the new version, Qt4, is not backward compatible with Qt3. Moving from Qt3 to Qt4 is a hard work as the organization of the Qt classes and communication between objects have been completely changed. Presently, only a simplified version of the Qgar graphical interface is available for Linux and Windows platforms (but Microsoft Visual Studio cannot be used for compilation).

Perspectives. As explained previously, moving from Qt3 to Qt4 is undoubtedly the more difficult and time consuming work, to be continued in 2008. Indeed, although Trolltech has fully reengineered Qt, it does not provide any efficient nor powerful tool to refactor existing code. In compensation, the new version of Qt offers its own SAX implementation, which should fruitfully replace the current Qgar home-made implementation. An intensive use of XML opens an interesting perspective on the support of the SVG vectorial format, which is used by a lot of researchers in the field, including those of the Qgar project-team. Serialization and persistence of XML objects as well as the adoption of XML as a default data format should substantially increase interest for the Qgar library as it could be jointly used with all external tools based on the related formats. Simultaneously, the experimental code of the Qgar development library will be progressively refactored to be included in the standard QgarLib library.

5.2. The Épeires Environment

Participant: Philippe Dosch.

In the context of the Épeires performance evaluation project (cf. § 8.1.1), we have improved the Épeires framework, a complete information system, capable of managing all required data related to performance evaluation of symbol recognition methods. It includes the management of data themselves, but also of their classification, of the automatic degradation processes, of the participants profiles, and of the available tests and result storage. This environment has been developed using the *Ruby on Rails* framework (http://www.rubyonrails.org/), and is available at the web site of the project, hosted at http://www.epeires.org.

The web site is also the location where all resources related to the project are freely available for the scientific community. In this way, users are able to manage their methods, to generate testing evaluations for specific purposes or to download existing ones, and to send and analyze their results. The same functionalities are provided to the organizers (Épeires Consortium).

We also developed a collaborative ground-truth managing software, called Picvert, coded in Java and hosted at http://gforge.inria.fr/. It is used to create, review and validate the labelling of test images with respect to reference symbols (ground-truth). Its architecture is based on a client/server model connected to the information system.

The web site is now fully operational, and we continue to improve it, by adding new functionalities and by refining the existing content. It has been used to organize the third International Symbol Recognition Contest, which has been held during GREC'2007 in Curitiba, Brazil. As we still extend both web site functionalities and data, we plan to use it to organize several other campaigns during 2008.

6. New Results

6.1. Combination of Classifiers and Shape Descriptors

Participants: Jan Rendek, Jean-Pierre Salmon, Oriol Ramos Terrades, Salvatore Tabbone, Laurent Wendling.

Combining outputs of classifiers or descriptors is one of the strategies used to improve classification rates in common classification. We tackled the problem of combining classifiers within a non-Bayesian framework, considering both two-class and multi-class classifiers. A classifier is assimilated to a random variable and finding the best classifier combination is expressed as solving an optimization problem. This approach has been applied to the combination of a set of ridgelets descriptors [2].

We are also interested in combining images descriptors and text modes. We have chosen to use probabilistic graphical models, especially Bayesian networks. In fact, these models allow for combining different types of information inside a same network, and for managing missing data. We have proposed an original adaptation of such networks to the problem of visual descriptors combination for image classification [11], [17].

We also proposed an approach to automatically extract a pertinent subset of soft output classifiers and to aggregate them to provide a global decision using the Choquet integral [18]. This approach relies on two key points. Firstly, a learning method based on a 2-class model provides a new set of decisions rules to recognize one class from another one. All the associated capacities are aggregated again at a high level to recognize symbols. Secondly, a selection method discards weak or redundant decision rules, keeping only the most relevant subset. An experimental study, based on real world data, is then described. It analyzes the improvements achieved by these points, first when used independently, then when combined together.

6.2. Scaling of Symbol Recognition Methods

Participants: Sabine Barrat, Bart Lamiroy, Jean-Pierre Salmon, Salvatore Tabbone, Karl Tombre, Laurent Wendling.

The FRESH project (cf. § 7.2) gives a new dimension to the recognition/classification problem as it mainly aims at distinguishing between roughly 500 different symbols either very similar, only differing by slight details, or completely dissimilar from a visual point of view, or even composed of other known and significant symbols, possibly not connected. Furthermore, the applicative framework of the project requires the analysis of a very large number of very large images. The main difficulty is the preservation of the scalability with respect to a substantial increase of the amount of data: structural relationships between (parts of) symbols should be available to perform semantical interpretation as well as electric simulations from wiring diagrams, robustness should be tuned so that image noise does not perturb recognition but fine details are captured as well, fine details should not hinder scalability, etc.

Global signal-based shape descriptors present many drawbacks in this context. They do not handle well connected symbols and composition, and they are generally not very suitable to the capture of small detail changes, though this tends to make them fault tolerant to image distortions. They nevertheless scale very well and allow easy insertion of new symbols in the database. Structural shape descriptors are very segmentation-dependent and do not scale easily when they are based on graph representations. They also lack robust expressive power when hierarchically organized. However, they capture details very well and provide interesting hooks for semantic interpretation.

Our contribution consists in designing a composite descriptor, robust and compact, with a low computational complexity. It is composed of smaller elementary descriptors, yielding integer or boolean values, which are thoroughly evaluated with respect to robustness and reliability. They express information like the number of small occlusions, the occurrence of text, the number of connected components, the presence of symmetrical extensions, full circles [12] or rectangles, etc.

As the main goal is scalability, we have developped a second stage operating disambiguation on the remaining candidates. We improved a pattern recognition scheme based on genetic algorithms, by integrating circular arc features and by adding the ability to take feature topology into account [21].

Storing and querying high-dimensional data are important problems in designing an information retrieval system [23]. When handling a large database, a system needs an efficient index mechanism to retrieve data by their contents. In this perspective, Salim Jouili started a PhD thesis about methods of information retrieval in masses of documents represented in a structural form. When used for graphic documents, such methods are well appropriate to pattern recognition because they offer a rich framework to describe symbols of unspecified shapes and structural relations between them. This research takes place in the scientific environment of project Navidomass (Navigation In Document Masses of the call to project ANR MDCA) which relates to indexing of large bases of ancient cultural heritage documents.

6.3. On-the-fly Recognition and Symbol Spotting

Participants: Philippe Dosch, Bart Lamiroy, Oanh Nguyen, Salvatore Tabbone, Wan Zhang.

The SYMBOLREC associated team with CVC of Barcelona (cf. § 8.2.1) also gave rise to the opportunity of continuing previously started work on on-the-fly recognition [13]. This work falls within the domains of syntactical and sketch-based symbol recognition, and is related to grammatical inference. Once topological properties of symbols are expressed using an adjacency grammar, a set of local geometric invariants is produced by the grammar rules which are triggered by the recognition process, in order to further disambiguate topologically similar configurations. The combination of both steps results in a very robust and efficient recognition method well-suited to user-drawn sketches. We show that the same approach can be easily adapted to the generation of adjacency grammars, or can be used as an interface for navigation and querying based on hand-drawn examples.

In the PhD thesis of Oanh Nguyen we focused on approaches to search and browse symbols in all kinds of document without prior knowledge on their contents. Our objective is to build a feature vocabulary adapted to graphic documents.

6.4. Evaluation and Benchmarking

Participants: Philippe Dosch, Gérald Masini, Jan Rendek, Salvatore Tabbone, Oriol Ramos Terrades, Karl Tombre, Wan Zhang.

Since the end of 2004, our project-team is leader of the Épeires project affiliated to the Techno-Vision campaign. The objective was the construction of a complete environment for performance evaluation of symbol recognition and localization methods. This topic has gained increasing interest in the last years, as demonstrated by the creation of two international contests on symbol recognition methods [27] [9]. In addition, we have worked with the other Épeires partners on the definition of the related metrics and protocols required by evaluation campaigns. This also includes the development of an open source collaborative ground-truth management software for dataset labelling purpose (cf. § 5.2).

In collaboration with CVC at University of Barcelona as part of the SYMBOLREC associated team (cf. § 8.2.1), we also started a new project about characterization of shape descriptors, in order to define a "genetic map" of selected descriptors, that is to say a list of properties some families among them share. These properties must not be dataset-dependent. We hope that such a map will lead to the definition of usage profiles, so that a user facing a practical pattern recognition problem can get help in choosing the most appropriate family of descriptors.

Such a protocol must be independent of datasets to be of general use. Evaluation models are easy to implement when they are dataset-driven, as they just require the computation of some recognition rates on a given dataset. However, they do not give much information about the behavior of a descriptor on a different dataset. Our aim is the design of dataset-independent models, which is a more ambitious and a more complex task.

We are interested both in pixel-based descriptors, computed on all pixels of the shape or on a subset of these pixels (contours, or regions, for example), and in structural descriptors, computed from the components of the shape and from the relationships between them. Each descriptor must be evaluated in terms of computational complexity, of robustness to geometric transformations and image degradations, of genericity, and of separability and compactibility. The last two criteria give a measure of the discrimination power of a descriptor when the number of classes and the variability of the shapes grow.

Evaluation is also the matter of our collaboration with City University Hong Kong. During her visit to LORIA, Wan Zhang integrated a descriptor developed by her team [31] in the shape descriptor characterization project, in order to evaluate its discrimination power and its efficiency using different databases. The Hong Kong group is developing an online symbol recognition system, which improves the efficiency and effectiveness of the user's input by presenting him/her with possible candidates dynamically selected from the symbol database. Wan Zhang worked on applying this idea to help annotating the ground-truth in the Épeires environment.

7. Contracts and Grants with Industry

7.1. France Télécom R&D

Participants: Bart Lamiroy, Jan Rendek, Karl Tombre, Laurent Wendling.

We have a long-lasting partnership with France Télécom R&D (FT R&D) on various issues within document image analysis. From 2001 to 2003, we were members of the RNTL DocMining project led by FT R&D. Since the end of 2004, we have a new partnership on the topic of on-the-fly symbol recognition (cf. § 6.3). More specifically, FT R&D paid the salary of Jan Rendek, a PhD student under a CIFRE contract.

7.2. The Fresh Project and Algo'tech

Participants: Bart Lamiroy, Laurent Fritz, Jean-Pierre Salmon, Karl Tombre, Laurent Wendling.

We contribute to the recapture of electrical wiring diagrams, with a particular focus on recognition methods for complex symbols. More precisely, we are developing a hierarchical symbol recognition method to progressively "zoom in" on a few models out of several hundred possible ones, by computing a set of simple geometric descriptors and relationships between them [12]. We have also improved a pattern recognition scheme based on a genetic algorithm [21] by adding the ability to take arcs of circle and connected components into account. The first prototype of the software was presented at the Bourget International Aerospace Fair.

The partners of the project are Algo'tech Informatique (France), ESTIA (France), Euro Inter (France), ECM Drawings (France), CEIT (Spain), Rector (Poland), Tekever (Portugal), and Zenon (Greece).

7.3. Netlor Concept

Participants: Sabine Barrat, Salvatore Tabbone.

Netlor Concept is a French company located in Nancy, which develops web services to optimize company management and to dynamize information exchange. It pays the salary of Sabine Barrat, a PhD student under a CIFRE contract, who works on objects recognition (cf. § 4.1).

8. Other Grants and Activities

8.1. National Actions

8.1.1. Techno-Vision ÉPEIRES

Participants: Philippe Dosch, Gérald Masini, Karl Tombre.

Techno-Vision is a French national program to fund projects related to performance evaluation of vision algorithms in computer science. We are leaders of the Épeires project[1] on performance evaluation of symbol spotting and recognition (2005–2007) and, as such, in charge of the scientific animation, of the creation of the information system related to the project, and of the creation of the testing data which have been used during an evaluation campaign opened to all interested participants at the end of 2007, during GREC'2007 in Curitiba, Brazil.

The funded partners are the universities of La Rochelle (L3I laboratory), Rouen (PSI laboratory) and Tours (LI laboratory), and the Algo'Tech company. The non-funded partners are the City University of Hong Kong, the DAG team of the Computer Vision Center of Barcelona, and the ONE laboratory of France Télécom R&D.

8.1.2. Navidomass

Participants: Salim Jouili, Salvatore Tabbone.

The NAVIDOMASS project was accepted by the scientific committee of ANR in Fall 2006 and started at the beginning of 2007. It is a research proposal to the ANR call on Data Masses (ARA MDCA). The general purpose is the construction of a framework to digitally preserve and provide universal access to heritage document collections in libraries, museums and public archives. The main focus is put on the navigation in large databases of archives containing text, images, illustrations and schemas, through the extraction of knowledge from images of documents.

Our partners are the IMADOC group at IRISA in Rennes, Centre d'Études Supérieures de la Renaissance in Tours (CESR, UMR 6576), and the universities of La Rochelle, Tours, Rouen and Paris 5.

8.2. International Cooperation

8.2.1. CVC Barcelona

We intensified our long-lasting scientific cooperation with the Computer Vision Center at *Universitat Autònoma de Barcelona*, including joint PhD supervisions, student, regular researcher and post-doc exchanges, collaboration in the Techno-Vision Épeires project, INRIA associated team SYMBOLREC[2], funding through PAI Picasso for joint work, joint European proposal, and joint organization of international symbol recognition contests.

We also continue our joint project on the characterization and evaluation of shape descriptors (cf. § 6.4), and organized two joint workshops. The first, held in Nancy on June 21–22, examined the progress of the work on shape descriptors evaluation, while the second, held in Barcelona in December 13–14, examined the progress of this work and discussed the different ways of continuation for 2008.

[1] http://www.epeires.org/
[2] http://dag.cvc.uab.es/symbolrec/

The work on the evaluation includes:

- Definition and discussion of several measures to characterize the performance of shape descriptors. Several criteria for assessing separability and compactability have been proposed and discussed in order to define the set of rating indices to be used.
- Implementation of several standard descriptors to be included in the final evaluation.
- Definition and generation of several datasets to be used to characterize descriptors according to constraints of genericity, to robustness to distortions, and to domain independence. Standard databases of graphic symbols, logos, compact shapes and characters have been used, and a protocol to generate different kinds of distortions in these databases has been defined.
- Evaluation of a limited set of shape descriptors, to test the general framework (evaluation criteria, definition of datasets, types of descriptors included in the framework, etc.).

In addition to the main activity on the evaluation of standard shape descriptors we are investigating new descriptors for symbol recognition. Joint research activities are studying shape signature models for symbol recognition:

- *O. Ramos Terrades, S. Tabbone, and E. Valveny*. At the image level, the co-supervised thesis of Oriol Ramos Terrades [28], defended in October 2006, proposed a classifier combination scheme of multiresolution descriptors based on the ridgelets transform (cf. § 6.1). We continued investigating a better application of the combination scheme to the descriptor obtained with the ridgelets transform.
- *M. Rusiñol, J. Lladós, and P. Dosch*. Querying symbols by shape in large databases has a great interest when working with technical CAD documents. The interest of indexing methods is the fact that they are able to tackle with recognition and segmentation at the same time. They provide regions of interest, in which the queried symbol has a high probability to be found. We are working on vectorial signatures, that is to say descriptors to index symbols represented by vectorial data (cf. § 6.3). A symbol is described by polylines encoded as attributed strings. A lookup table is constructed to "index" symbol queries in terms of polylines. We proposed a new signature model, using accumulated length and accumulated turning angle as shape descriptor to guarantee invariance to segment cardinality. It has been applied to symbol detection in real images in camera-based frameworks and classical symbol spotting in document images,
- *J. Mas, B. Lamiroy, G. Sánchez, and J. Lladós*. They are working on syntactical approaches for symbol representation, in particular using adjacency grammars for sketch recognition (hand-drawn symbols [13]). The joint work focused on infering grammars representing symbol classes from hand-drawn examples. The inference mechanism is based on the most representative rulesets of geometrical and relational constraints among primitives. Starting from this syntactical model, an incremental on-line parsing algorithm has been proposed to recognize sketched diagrams.

8.2.2. City University of Hong Kong

We continued our research collaboration with Wenyin Liu's group at the City University of Hong Kong (CityU). Our scientific objective is a common work on symbol recognition for information retrieval in technical documentation, through the construction of a reference database and the study of performance evaluation of graphics recognition methods. The project is partially funded by PAI Procore and by *Région Lorraine*, and also includes a collaboration within the Techno-Vision Épeires project (§ 8.1.1). In May, Philippe Dosch and Karl Tombre visited CityU and had exchanges with Wenyin Liu and Wan Zhang about symbol recognition and construction of a reference database. The exchanges have been also related to ongoing work done by Wan Zhang on vectorial signatures of graphical symbols.

From September to December, Wan Zhang, a PhD student at CityU, stayed in Nancy and took part in our work on performance characterization and evaluation (cf. § 6.4). As in 2006, she also worked on the design of a graph of spatial relations between basic geometric shapes, to make the best possible use of the respective contributions of LORIA and CityU: decomposition of a symbol into basic geometric shapes for the former, hierarchical representation of relations between basic geometric shapes for the latter.

In November, Bart Lamiroy visited CityU for one week, in order to collaborate with Wenyin Liu on arc segmentation based on [12] and to participate in the GREC'2007 Arc Segmentation Contest.

8.2.3. SEPIA

Participants: Oanh Nguyen, Oriol Ramos Terrades, Salvatore Tabbone.

SEPIA, *Système d'Étude du Patrimoine des Inscriptions Anciennes au Vietnam*, is a CNRS international project with *Mica* Lab in Hanoi (Vietnam). SEPIA deals with the problem of indexing and searching for information (typically ideograms) in images of ancient Vietnamise steles. As it started in September, we cannot present significant results yet.

8.2.4. IDEA

Participants: Benoît Naegel, Oriol Ramos Terrades, Salvatore Tabbone.

IDEA (Images of natural Disasters from robot Exploration in urban Area) is a regional programme *Ict*-Asia. Its aim is the development of an expertise about image processing and computer vision for urban natural disasters, when using a camera mounted on a patrolling robot as a mobile sensor on disaster sites. More precisely, it proposes a new research direction to make it easy to identify images and to take decision about rescue management after a natural disaster. Our partners in this project are team *Modélisation et Simulation Informatique de systèmes complexes* at *Institut de la Francophonie pour l'Informatique* of Hanoi (Vietnam), Institute of Information Technology of Hanoi, *Universiti Kuala Lumpur Malaysia*, and *Laboratoire Informatique, Image, Interaction* (L3I) at *Université de La Rochelle*.

8.2.5. Miscellaneous

Karl Tombre is working with University of Yaoundé 1 and *École Polytechnique de Yaoundé* (Cameroon), helping in the supervision of two PhD students on the topics of geographical information systems and map analysis. One of these PhDs was defended in November.

9. Dissemination

9.1. Animation of Scientific Community

9.1.1. Journals

Karl Tombre is editor in chief of the International Journal on Document Analysis and Recognition (IJDAR), member of the advisory board of Electronic Letters on Computer Vision and Image Analysis (ELCVIA), and member of the editorial board of Machine Graphics & Vision and of *Revue Africaine de la Recherche en Informatique et Mathématiques Appliquées* (ARIMA).

9.1.2. Associations

Karl Tombre is the elected president of the International Association for Pattern Recognition (IAPR, see http://www.iapr.org/) for a two-years term from 2006 to 2008.

Karl Tombre is the president of the French Association for Pattern Recognition and Image Processing (AFRIF, *Association Française pour la Reconnaissance et l'Interprétation des Formes*).

9.1.3. Other Responsibilities

- Bart Lamiroy is elected to the scientific council of INPL.

  He is a member of *Comité de suivi de l'espace transfert*, which follows and evaluates spin-offs and start-ups resulting from research works carried out in LORIA.

  He is an active member of the Education Commission of the Information Technology Service Management Forum (itSMF) France.
- Gérald Masini is the president of the Commission of Computing Resources Users (COMIN), and also head (by special delegation of the director) of the IT Service at LORIA.
- On February 1st, Karl Tombre was appointed director of the INRIA Nancy - Grand Est Research Center. He is also the director of LORIA. His main duty is therefore to manage a large research center of about 450 persons.

  Karl Tombre was member of various evaluation committees for French research institutions: LIRIS (Rouen), LISyC & E3I2 (Brest), École des Mines de Douai.
- Laurent Wendling is a member of the council of LORIA and is elected to the studies council (CEVU) of Université Henri Poincaré Nancy 1.

9.2. Collaborations within INRIA

A cooperation with Isabelle Debled-Rennesson (project-team ADAGIO) is ongoing. We have proposed a new way to extract arcs of circle and line segments from curvature profiles by defining adapted filters in the splitting process [29].

We regularly work with the IMADOC group at IRISA, a partner of our new NAVIDOMASS project (cf. § 8.1.2) on heritage documents.

We also have recurring exchanges with the TexMex group at IRISA, and serious intentions to collectively develop research on combined text and graphics mining.

9.3. Teaching

Most members of the QGAR project-team are university faculty members and, as such, have a statutory teaching service in their respective universities, cumulated, for several of them, with major organizational and administrative responsibilities. They have teaching positions at various institutions:

- Suzanne Collin, at Université Henri Poincaré Nancy 1/ESIAL (engineering school, master of engineering level).
- Philippe Dosch, at Université de Nancy 2, at bachelor level.

  He is the director of studies for the bachelor degree "Administration of open source systems, networks and applications".
- Bart Lamiroy, at *Institut National Polytechnique de Lorraine/École des Mines de Nancy* (engineering school, master of engineering level).

  He heads the Department of Computer Science, and is the technical coordinator of the IPISO specialized degree. He is also vice president of the recruitment committee in computer science (*Commission de Spécialistes, 27^{e} section*) at Institut National Polytechnique de Lorraine.
- Salvatore Tabbone, at Université de Nancy 2, at bachelor and master level.

  He is vice-president of the recruitment committee in computer science (*Commission de Spécialistes, 27^{e} section*) at Université de Nancy 2.
- Laurent Wendling, at Université Henri Poincaré Nancy 1/ESIAL (engineering school, master of engineering level).

9.4. Conference and Workshop Committees

- Philippe Dosch is member of the program committee of CIFED'2008 (10e Colloque International Francophone sur l'Écrit et le Document, Rouen, France).
- Bart Lamiroy is member of the program committee of CIFED'2008 (10e Colloque International Francophone sur l'Écrit et le Document, Rouen, France).
- Salvatore Tabbone is general chairman of CIFED'2008 (10e Colloque International Francophone sur l'Écrit et le Document, Rouen, France). He was/is member of the program committees of ICDAR'2009 (10th International Conference on Document Analysis and Recognition, Barcelona, Spain), CIFED'2008 (10e Colloque International Francophone sur l'Écrit et le Document, Rouen, France), CARI'2008 (9th African Conference on Research in Computer Science and Applied Mathematics, Rabat, Maroc), ACM-SAC'2008 (23rd ACM Symposium on Applied Computing, Fortaleza, Brazil), ICCR'2008 (2nd International Conference on Cognition and Recognition, India), ACM-SAC'2007 (22th ACM Symposium on Applied Computing, Seoul, Korea), ORASIS'2007 (11e Congrès Francophone des Jeunes Chercheurs en Vision par Ordinateur, Obernai, France), and TAIMA'2007 (Traitement et Analyse de l'Information : Méthodes et Applications, Hammamet, Tunisie).
- Karl Tombre was/is member of the program committee of ICDAR'2009 (10th International Conference on Document Analysis and Recognition, Barcelona, Spain), DAS'2008 (8th International Workshop on Document Analysis Systems, Nara, Japan), RFIA'2008 (16e Congrès Francophone AFRIF-AFIA de Reconnaissance des Formes et Intelligence Artificielle, Amiens, France), CIFED'2008 (10e Colloque International Francophone sur l'Écrit et le Document, Rouen, France), CIARP'2008 (13th Iberoamerican Congress on Pattern Recognition, Havana, Cuba), ICDAR'2007 (9th International Conference on Document Analysis and Recognition, Curitiba, Brazil), GREC'2007 (7th International Workshop on Graphics Recognition, Curitiba, Brazil), MVA'2007 (10th IAPR International Conference on Machine Vision Applications, Tokyo, Japan), ICIAP'2007 (14th International Conference on Image Analysis and Processing, Modena, Italy), and IbPRIA'2007 (3rd Iberian Conference on Pattern Recognition and Image Analysis, Girona, Spain).
- Laurent Wendling was/is member of the program committee of IASTED VIIP'2007 (6th International Conference on Visualization, Imaging, and Image Processing, Palma de Mallorca, Spain), IASTED VIIP'2008 (7th International Conference on Visualization, Imaging, and Image Processing, Palma de Mallorca, Spain), IASTED ICVGIP'2007 (5th Indian Conference on Computer Vision, Graphics and Image Processing, Madurai, India), CIFED'2008 (10e Colloque International Francophone sur l'Écrit et le Document, Rouen, France), and ORASIS'2007 (11e Congrès Francophone des Jeunes Chercheurs en Vision par Ordinateur, Obernai, France).

10. Bibliography

Major publications by the team in recent years

[1] X. HILAIRE, K. TOMBRE. *Robust and Accurate Vectorization of Line Drawings*, in "IEEE Transactions on Pattern Analysis and Machine Intelligence", vol. 28, no 6, 2006, p. 890–904, http://hal.inria.fr/inria-00000394/en/.

[2] O. RAMOS TERRADES, E. VALVENY, S. TABBONE. *On the combination of ridgelets descriptors for symbol recognition*, in "Proceedings of the 7th International Workshop on Graphics Recognition (Curitiba, Brazil)", 2007, p. 18–19, http://hal.inria.fr/inria-00179493/en/.

[3] S. TABBONE, L. ALONSO, D. ZIOU. *Behavior of the Laplacian of Gaussian extrema*, in "Journal of Mathematical Imaging and Vision", vol. 23, no 1, 2005, p. 107–128, http://hal.inria.fr/inria-00001244/en/.

[4] S. TABBONE, L. WENDLING, J.-P. SALMON. *A new shape descriptor defined on the Radon transform*, in "Computer Vision and Image Understanding", vol. 102, n^{o} 1, April 2006, p. 42–51, http://hal.inria.fr/inria-00000643/en/.

[5] S. TABBONE, L. WENDLING, K. TOMBRE. *Matching of graphical symbols in line-drawing images using angular signature information*, in "International Journal on Document Analysis and Recognition", vol. 6, n^{o} 2, 2003, p. 115–125, http://hal.inria.fr/inria-00099540/en/.

[6] L. WENDLING, S. TABBONE. *A new way to detect arrows in line drawings*, in "IEEE Transactions on Pattern Analysis and Machine Intelligence", vol. 26, n^{o} 7, 2004, p. 935–941, http://hal.inria.fr/inria-00100177/en/.

Year Publications

Articles in refereed journals and book chapters

[7] S. BARRAT, S. TABBONE. *A progressive learning method for symbol recognition*, in "Journal of Universal Computer Science, special issue", 2008 (to be published).

[8] J. RENDEK, L. WENDLING. *On determining suitable subsets of decision rules using Choquet integral*, in "International Journal of Pattern Recognition and Artificial Intelligence", 2008 (to be published).

[9] E. VALVENY, P. DOSCH, A. WINSTANLEY, Z. YU, Y. SU, Y. LUO, W. LIU, D. ELLIMAN, M. DELALANDRE, E. TRUPIN, S. ADAM, J.-M. OGIER. *A general framework for the evaluation of symbol recognition methods*, in "International Journal On Document Analysis and Recognition", vol. 9, 2007, p. 59–74, http://hal.inria.fr/inria-00176899/en/.

Publications in Conferences and Workshops

[10] S. BARRAT, S. TABBONE. *A progressive learning method for symbols recognition*, in "Proceedings of the 22nd ACM Symposium on Applied Computing Document Engineering Track (Seoul, Korea)", 2007, p. 627–631, http://hal.inria.fr/inria-00118309/fr/.

[11] S. BARRAT, S. TABBONE, P. NOURRISSIER. *A Bayesian classifier for symbol recognition*, in "Proceedings of the 7th International Workshop on Graphics Recognition (Curitiba, Brazil)", 2007, p. 24–25, http://hal.inria.fr/inria-00179764/fr/.

[12] B. LAMIROY, O. GAUCHER, L. FRITZ. *Robust circle detection*, in "Proceedings of the 9th International Conference on Document Analysis and Recognition (Curitiba, Brazil)", 2007, p. 526–530, http://hal.inria.fr/inria-00184490/fr/.

[13] J. MAS ROMEU, G. SANCHEZ, J. LLADÓS, B. LAMIROY. *An incremental on-line parsing algorithm for recognizing sketching diagrams*, in "Proceedings of the 9th International Conference on Document Analysis and Recognition (Curitiba, Brazil)", 2007, p. 452–456, http://hal.inria.fr/inria-00186621/fr/.

[14] C. MAZAUD, J. RENDEK, V. BOMBARDIER, L. WENDLING. *A feature selection method based on Choquet integral and typicality analysis*, in "Proceedings of the 16th IEEE International Conference on Fuzzy Systems (London, United Kingdom)", 2007.

[15] C. MAZAUD, J. RENDEK, V. BOMBARDIER, L. WENDLING. *Une méthode de sélection de caractéristiques fondées sur l'intégrale de Choquet et l'analyse des typicalités*, in "Actes du 21e Colloque Gretsi sur le Traitement du Signal et des Images (Troyes, France)", 2007.

[16] O. RAMOS TERRADES, S. TABBONE, E. VALVENY. *A review of shape descriptors for document analysis*, in "Proceedings of the 9th International Conference on Document Analysis and Recognition (Curitiba, Brazil)", 2007, http://hal.inria.fr/inria-00179453/en/.

[17] O. RAMOS TERRADES, S. TABBONE, E. VALVENY. *Optimal linear combination for two-class classifiers*, in "Proceedings of the 6th International Conference on Advances in Pattern Recognition (Kolkata, India)", 2007, http://hal.inria.fr/inria-00179484/en/.

[18] J. RENDEK, L. WENDLING. *Symbol Recognition using a 2-class hierarchical model of Choquet integrals*, in "Proceedings of the 9th International Conference on Document Analysis and Recognition (Curitiba, Brazil)", 2007, p. 634-638.

[19] M. RUSIÑOL, P. DOSCH, J. LLADÓS. *Boundary shape recognition using accumulated length and angle information*, in "Proceedings of the 3rd Iberian Conference on Pattern Recognition and Image Analysis (Girona, Spain)", Lecture Notes in Computer Science 4478, 2007, p. 210–217, http://hal.inria.fr/inria-00142213/fr/.

[20] M. RUSIÑOL, J. LLADÓS, P. DOSCH. *Camera-based graphical symbol detection*, in "Proceedings of the 9th International Conference on Document Analysis and Recognition (Curitiba, Brazil)", vol. 2, 2007, p. 884–888, http://hal.inria.fr/inria-00176893/fr/.

[21] J.-P. SALMON, L. WENDLING. *ARG based on arc and line segments to improve the symbol recognition by genetic algorithm*, in "Proceedings of the 7th International Workshop on Graphics Recognition (Curitiba, Brazil)", 2007, http://hal.inria.fr/inria-00184880/fr/.

[22] S. TABBONE, J. LLADÓS. *A propos de la reconnaissance de documents graphiques : synthèse et perspectives*, in "Actes du congrès de Traitement et Analyse de l'Information : Méthodes et Applications (Hammamet, Tunisie)", 2007, http://hal.inria.fr/inria-00179440/en/.

[23] S. TABBONE, D. ZUWALA. *An indexing method for graphical documents*, in "Proceedings of the 9th International Conference on Document Analysis and Recognition (Curitiba, Brazil)", 2007, http://hal.inria.fr/inria-00204098/fr/.

[24] E. VALVENY, S. TABBONE, O. RAMOS TERRADES, E. PHILLIPPOT. *Performance characterization of shape descriptors for symbol representation*, in "Proceedings of the 7th International Workshop on Graphics Recognition (Curitiba, Brazil)", 2007, p. 82–83, http://hal.inria.fr/inria-00179490/en/.

[25] W. ZHANG, P. DOSCH, S. TABBONE. *Performance evaluation on several recognition algorithms*, in "Proceedings of the 7th International Workshop on Graphics Recognition (Curitiba, Brazil)", 2007, p. 86–89, http://hal.inria.fr/inria-00193106/en/.

Miscellaneous

[26] K. TOMBRE. *Graphics recognition for querying and browsing technical documentation*, January 2007, Invited talk at Symposium on Machine Learning in Image and Document, held for the 10th Anniversary

of LIAMA (Sino-French Laboratory for Computer Science, Automation and Applied Mathematics), Institute of Automation, Chinese Academy of Sciences, Beijing (China).

References in notes

[27] P. DOSCH, E. VALVENY. *Report on the Second Symbol Recognition Contest*, in "Graphics recognition—Ten years review and future perspectives (revised selected papers from GREC'05)", W. LIU, J. LLADÓS (editors), Lecture Notes in Computer Science 3926, Springer Verlag, 2006, p. 381–397, http://hal.inria.fr/inria-00091343/en/.

[28] O. RAMOS TERRADES. *Linear combination of multiresolution descriptors: application to graphics recognition*, Thèse de doctorat en cotutelle, Universitat Autònoma de Barcelona and Université de Nancy 2, October 2006.

[29] J.-P. SALMON, I. DEBLED-RENNESSON, L. WENDLING. *A new method to detect arcs and segments from curvature profiles*, in "Proceedings of the 18th International Conference on Pattern Recognition (Hong Kong, China)", vol. 3, August 2006, p. 387–390, http://hal.inria.fr/inria-00104414/en/.

[30] K. TOMBRE, S. TABBONE, P. DOSCH. *Musings on symbol recognition*, in "Graphics recognition—Ten years review and future perspectives (revised selected papers from GREC'05)", W. LIU, J. LLADÓS (editors), Lecture Notes in Computer Science 3926, Springer Verlag, 2006, p. 23–34, http://hal.inria.fr/inria-00000230/en/.

[31] W. ZHANG, L. WENYIN, K. ZHANG. *Symbol recognition with kernel density matching*, in "IEEE Transactions on Pattern Analysis and Machine Intelligence", vol. 28, n° 12, December 2006, p. 2020–2024.