



Activity Report 2011

Project-Team IMEDIA

Image and multimedia indexing, browsing and retrieval

RESEARCH CENTER
Paris - Rocquencourt

THEME
**Vision, Perception and Multimedia
Understanding**

Table of contents

| | |
|--|-----------|
| 1. Members | 1 |
| 2. Overall Objectives | 2 |
| 2.1. Introduction | 2 |
| 2.2. Highlights | 2 |
| 3. Scientific Foundations | 3 |
| 3.1. Introduction | 3 |
| 3.2. Modelling, construction and structuring of the feature space | 3 |
| 3.3. Pattern recognition and statistical learning | 4 |
| 3.4. Interactive search and personalisation | 4 |
| 4. Application Domains | 5 |
| 5. Software | 5 |
| 6. New Results | 6 |
| 6.1. Feature space modelling | 6 |
| 6.1.1. A novel shape boundary based description for leaf identification | 6 |
| 6.1.2. Visual-based plant species identification from crowdsourced data | 7 |
| 6.1.3. Spatial relations between salient points on a leaf | 7 |
| 6.1.4. 3D mesh segmentation by example | 7 |
| 6.2. Feature space structuring | 9 |
| 6.2.1. Random Maximum Margin Hashing | 9 |
| 6.2.2. Scalable information retrieval in distributed architectures | 9 |
| 6.2.3. Visual similarity sensitive hashing methods for semantic image search in very large collections of images | 10 |
| 6.3. Pattern recognition and statistical learning | 10 |
| 6.4. Interactive search and personalisation | 11 |
| 6.4.1. Database denoising and multi visual queries | 11 |
| 6.4.2. Object-based Visual Query Suggestion | 12 |
| 6.4.3. Interpretable Visual Models for Human Perception-based Object Retrieval | 12 |
| 6.4.4. Relevance feedback on local features : Application to plants annotations and identification | 12 |
| 6.5. Software | 13 |
| 7. Partnerships and Cooperations | 13 |
| 7.1. National Initiatives | 13 |
| 7.1.1. PI@ntNet project [2009-2013] | 13 |
| 7.1.2. Other collaborations with INRA | 14 |
| 7.1.3. ANR project SCAR-FACE [2008-2011] | 14 |
| 7.2. European Initiatives | 14 |
| 7.2.1.1. CHORUS+ | 14 |
| 7.2.1.2. GLOCAL | 14 |
| 7.2.1.3. I-SEARCH | 15 |
| 7.3. International Initiatives | 15 |
| 8. Dissemination | 15 |
| 8.1. Animation of the scientific community | 15 |
| 8.1.1. Nozha Boujemaa | 15 |
| 8.1.2. Alexis Joly | 16 |
| 8.1.3. Anne Verroust-Blondet | 16 |
| 8.2. Teaching | 16 |
| 9. Bibliography | 17 |

Project-Team IMEDIA

Keywords: Image Processing, Multimedia, Computer Vision, Information Indexing And Retrieval

1. Members

Research Scientists

Anne Verroust-Blondet [Interim team leader since November 1st 2010, Junior Researcher, HdR]
Nozha Boujemaa [Senior Researcher, HdR]
Alexis Joly [until June 30th 2011, Junior Researcher]

External Collaborators

Michel Crucianu [Professor at CNAM, HdR]
Itheri Yahiaoui [Associate Professor at Reims University]

Technical Staff

Jean-Paul Chièze [Senior Technical Staff INRIA (half-time) until June 1st 2011]
Vera Bakić [Engineer INRIA since February 1st 2011]
Vincent Ladeveze [Technical Staff INRIA until September 15th 2011]
Souheil Selmi [Junior Technical Staff INRIA]

PhD Students

Esmâ Elghoul [INRIA grant, Télécom ParisTech since October 15th 2010]
Amel Hamzaoui [INRIA grant, Paris-Sud University since October 1st 2008]
Pierre Letessier [CIFRE with INA since October 1st 2009]
Sofïène Mouine [INRIA grant, Télécom ParisTech since October 15th 2010]
Saloua Ouertani-Litayem [INRIA grant, Télécom ParisTech since October 1st 2009]
Wajih Ouertani [INRA grant, Paris-Sud University since October 1st 2008]
Ahmed Rebai [INRIA grant, Paris-Sud University until March 31st 2011]
Asma Rejeb Sfar [INRIA grant, Télécom ParisTech since October 15th 2010]
Mohamed Riadh Trad [INRIA grant, Télécom ParisTech since October 1st 2009]
Olfa Mzoughi [internship INRIA since March 1st 2011]

Post-Doctoral Fellows

Raffi Enciclaud [Expert engineer INRIA until March 31st 2011]
Hervé Goëau [Expert engineer INRIA]
Laurent Joyeux [Expert engineer INRIA]
Sébastien Poullot [Expert engineer INRIA since February 1st 2011]

Visiting Scientist

Donald Geman [Professor at John Hopkins University, HdR]

Administrative Assistant

Laurence Bourcier [Secretary (TRS) INRIA (shared with Salsa, Smis and Micmac project-team)]

2. Overall Objectives

2.1. Introduction

One of the consequences of the increasing ease of use and significant cost reduction of computer systems is the production and exchange of more and more digital and multimedia documents. These documents are fundamentally heterogeneous in structure and content as they usually contain text, images, graphics, video and sounds and may contain 3D objects. Information retrieval can no longer rely on text-based queries alone; it will have to be multi-modal and to integrate all the aspects of the multimedia content. In particular, the visual content has a major role and represents a central vector for the transmission of information. The description of that content by means of image analysis techniques is less subjective than the usual keyword-based annotations, whenever they exist. Moreover, being independent from the query language, the description of visual content is becoming paramount for the efficient exploration of a multimedia stream. In the IMEDIA group we focus on the intelligent and efficient access by visual content. With this goal in mind, we develop methods that address key issues such as content-based indexing, interactive search and image database navigation, in the context of multimedia content (text, image, video, 3D models). Content-based image retrieval systems provide help for the automatic search and assist human decisions. The user remains the *maître d'oeuvre*, the only one able to take the final decision. The numerous research activities in this field during the last decade have proved that retrieval based on the visual content was feasible. Nevertheless, current practice shows that a usability gap remains between the designers of these techniques/methods and their potential users. One of the main goals of our research group is to reduce the gap between the real usages and the functionalities resulting from our research on visual content-based information retrieval. Thus, we apply ourselves to conceive methods and techniques that can address realistic scenarios, which often lead to exciting methodological challenges. Among the "usage" objectives, an important one is the ability, for the user, to express his specific visual interest for a part of a picture. It allows him to better target his intention and to formulate it more accurately. Another goal in the same spirit is to express subjective preferences and to provide the system with the ability to learn those preferences. When dealing with any of these issues, we keep in mind the importance of the scalability of such interactive systems in terms of indexing and response times. Of course, what value these times should have and how critical they are depend heavily on the domain (specific or generic) and on the cost of the errors. Our research work is then at the intersection of several scientific domains. The main ones are image analysis, pattern recognition, statistical learning, human-machine interaction and database systems. It is structured into the following main themes:

1. 2D or 3D indexing: this part mainly concerns modelling the visual aspect of images and of 3D shapes, by means of image analysis techniques. It leads to the design of signatures that can then be obtained automatically.
2. feature space structuring: to increase in efficiency the search by content in very large collections of images.
3. Interactive search and personalization: to let the system take into account the preferences of the user, who usually expresses subjective or high-level semantic queries.

More generally, the research work and the academic and industrial collaborations of the IMEDIA team aim to answer the complex problem of the intelligent and efficient access to multimedia content.

2.2. Highlights

We organized the Plant Identification task at ImageCLEF 2011 (cf. <http://imageclef.org/2011/plants> and [19]). We submitted two runs [20], which won respectively the Scans images and the Scan-like images tasks.

3. Scientific Foundations

3.1. Introduction

We group the existing problems in the domain of content-based image indexing and retrieval in the following themes: image indexing and efficient search in image collections, pattern recognition and personalisation. In the following we give a short introduction to each of these themes.

3.2. Modelling, construction and structuring of the feature space

Participants: Nozha Boujemaa, Jean-Paul Chièze, Esma Elghoul, Raffi Enficiaud, Hervé Goëau, Amel Hamzaoui, Alexis Joly, Pierre Letessier, Sofiène Mouine, Olfa Mzoughi, Saloua Ouertani-Litayem, Ahmed Rebai, Mohamed Riadh Trad, Anne Verroust-Blondet, Itheri Yahiaoui.

The goal of the IMEDIA team is to provide the user with the ability to do content-based search into image databases in a way that is both intelligent and intuitive to the users. When formulated in concrete terms, this problem gives birth to several mathematical and algorithmic challenges.

To represent the content of an image, we are looking for a representation that is both compact (less data and more semantics), relevant (with respect to the visual content and the users) and fast to compute and compare. The choice of the feature space consists in selecting the significant *features*, the *descriptors* for those features and eventually the encoding of those descriptors as image *signatures*.

We deal both with generic databases, in which images are heterogeneous (for instance, search of Internet images), and with specific databases, dedicated to a specific application field. The specific databases are usually provided with a ground-truth and have an homogeneous content (faces, medical images, fingerprints, etc.)

Note that for specific databases one can develop dedicated and optimal features for the application considered (face recognition, etc.). On the contrary, generic databases require generic features (colour, textures, shapes, etc.).

We must not only distinguish generic and specific signatures, but also local and global ones. They correspond respectively to queries concerning parts of pictures or entire pictures. In this case, we can again distinguish approximate and precise queries. In the latter case one has to be provided with various descriptions of parts of images, as well as with means to specify them as regions of interest. In particular, we have to define both global and local similarity measures.

When the computation of signatures is over, the image database is finally encoded as a set of points in a high-dimensional space: the feature space.

A second step in the construction of the index can be valuable when dealing with very high-dimensional feature spaces. It consists in pre-structuring the set of signatures and storing it efficiently, in order to reduce access time for future queries (tradeoff between the access time and the cost of storage). In this second step, we have to address problems that have been dealt with for some time in the database community, but arise here in a new context: image databases. Today's scalability issues already put brake on growth of multi-media search engines. The searchable space created by the massive amounts of existing multimedia files greatly exceeds the area searched by today's major engines. Consistent breakthroughs are therefore urgent if we don't want to be lost in data space in ten years. We believe that reducing algorithm complexity remains the main key. Whatever the efficiency of the implementation or the use of powerful hardware and distributed architectures, the ability of an algorithm to scale-up is strongly related to its time and space complexities. Nowadays, efficient multimedia search engines rely on various high level tasks such as content-based search, navigation, knowledge discovery, personalization, collaborative filtering or social tagging. They involve complex algorithms such as similarity search, clustering or machine learning, on heterogeneous data, and with heterogeneous metrics. Some of them still have quadratic and even cubic complexities so that their use in the large scale is not affordable if no fundamental research is performed to reduce their complexities. In this way, efficient and generic high-dimensional similarity search structures are essential for building scalable content-based search systems.

Efficient search requires a specific structuring of the feature space (multidimensional indexing, where indexing is understood as data structure) for accelerating the access to collections that are too large for the central memory. The applications we have in mind are related to biodiversity (as in Pl@ntNet), to the detection of illegal copies of images and video (with INA) and to video surveillance and monitoring (with AVT).

3.3. Pattern recognition and statistical learning

Participants: Nozha Boujemaa, Michel Crucianu, Donald Geman, Alexis Joly, Wajih Ouertani, Ahmed Rebai, Asma Rejeb Sfar.

Statistical learning and classification methods are of central interest for content-based image retrieval. We consider here both supervised and unsupervised methods. Depending on our knowledge of the contents of a database, we may or may not be provided with a set of *labelled training examples*. For the detection of *known* objects, methods based on hierarchies of classifiers have been investigated. In this context, face detection was a main topic, as it can automatically provide a high-level semantic information about video streams. For a collection of pictures whose content is unknown, e.g. in a navigation scenario, we are investigating techniques that adaptively identify homogeneous clusters of images, which represent a challenging problem due to feature space configuration.

Object detection is the most straightforward solution to the challenge of content-based image indexing. Classical approaches (artificial neural networks, support vector machines, etc.) are based on induction, they construct generalisation rules from training examples. The generalisation error of these techniques can be controlled, given the complexity of the models considered and the size of the training set.

Our research on object detection addresses the design of invariant kernels and algorithmically efficient solutions as well as boosting method for similarity learning. We have developed several algorithms for face detection based on a hierarchical combination of simple two-class classifiers. Such architectures concentrate the computation on ambiguous parts of the scene and achieve error rates as good as those of far more expensive techniques.

Unsupervised clustering techniques automatically define categories and are for us a matter of visual knowledge discovery. We need them in order to:

- Solve the "page zero" problem by generating a visual summary of a database that takes into account all the available signatures together.
- Perform image segmentation by clustering local image descriptors.
- Structure and sort out the signature space for either global or local signatures, allowing a hierarchical search that is necessarily more efficient as it only requires to "scan" the representatives of the resulting clusters.

Given the complexity of the feature spaces we are considering, this is a very difficult task. Noise and class overlap challenge the estimation of the parameters for each cluster. The main aspects that define the clustering process and inevitably influence the quality of the result are the clustering criterion, the similarity measure and the data model.

We investigate a family of clustering methods based on the competitive agglomeration that allows us to cope with our primary requirements: estimate the unknown number of classes, handle noisy data and deal with classes (by using fuzzy memberships that delay the decision as much as possible).

3.4. Interactive search and personalisation

Participants: Nozha Boujemaa, Michel Crucianu, Donald Geman, Amel Hamzaoui, Wajih Ouertani, Sébastien Poullot, Anne Verroust-Blondet, Jean-Paul Chièze.

We are studying here the approaches that allow for a reduction of the "semantic gap". There are several ways to deal with the semantic gap. One prior work is to optimise the fidelity of physical-content descriptors (image signatures) to visual content appearance of the images. The objective of this preliminary step is to bridge what we call the numerical gap. To minimise the numerical gap, we have to develop efficient images signatures. The weakness of visual retrieval results, due to the numerical gap, is often confusingly attributed to the semantic gap. We think that providing richer user-system interaction allows user expression on his preferences and focus on his semantic visual-content target.

Rich user expression comes in a variety of forms:

- allow the user to notify his satisfaction (or not) on the system retrieval results—method commonly called relevance feedback. In this case, the user reaction expresses more generally a subjective preference and therefore can compensate for the semantic gap between visual appearance and the user intention,
- provide precise visual query formulation that allows the user to select precisely its region of interest and pull off the image parts that are not representative of his visual target,
- provide interactive visualisation tools to help the user when querying and browsing the database,
- provide a mechanism to search for the user mental image when no starting image example is available. Several approaches are investigated. As an example, we can mention the logical composition from visual thesaurus. Besides, learning methods related to information theory are also developed for efficient relevance feedback model in several context study including mental image retrieval.

4. Application Domains

4.1. Application Domains

- **Security applications** Examples: Identify faces or digital fingerprints (biometry). Biometry is an interesting specific application for both a theoretical and an application (recognition, supervision, ...) point of view. Two PhDs were defended on themes related to biometry. Our team also worked with a database of images of stolen objects and a database of images after a search (for fighting pedophilia).
- **Audio-visual applications** Examples: Look for a specific shot in a movie, documentary or TV news, present a video summary. Help archivists to annotate the contents. Detect copies of a given material in a TV stream or on the web. Our team has a collaboration with INA (French TV archives), IRT (German broadcasters) and press agencies AFP and Belga in the context of an European project. Text annotation is still very important in such applications, so that cross-media access is crucial.
- **Scientific applications** Examples: environmental images databases: fauna and flora; satellite images databases: ground typology; medical images databases: find images of a pathological character for educational or investigation purposes. We have an ongoing project on multimedia access to biodiversity collections for species identifications.
- **Culture, art and design** IMEDIA has been contacted by the French ministry of culture and by museums for their image archives.
Finding a specific texture for the textile industry, illustrating an advertisement by an appropriate picture. IMEDIA is working with a picture library that provides images for advertising agencies. IMEDIA has been involved in TRENDS European project dedicated to provide designers (CRF Fiat, Stile Bertone) with advanced content selection and visualisation tools.

5. Software

5.1. IKONA/MAESTRO Software

Participants: Vera Bakić, Nozha Boujemaa, Jean-Paul Chièze, Raffi Enficiaud, Alexis Joly, Laurent Joyeux, Olfa Mzoughi, Souheil Selmi, Itheri Yahiaoui.

IKONA is a generalist software dedicated to content-based visual information indexing and retrieval. It has been designed and implemented in our team during the last years [21]. Its main functionalities are the extraction, the management and the indexing of many state-of-the-art global and local visual features. It offers a wide range of interactive search and navigation methods including query-by-example, query-by-window, matching, relevance feedback, search results clustering or automatic annotation. It can manage several types of input data including images, videos and 3D models.

Based on a client/server architecture, it is easily deployable in any multimedia search engine or service. The communication between the two components is achieved through a proprietary network protocol. It is a set of commands the server understands and a set of answers it returns to the client. The communication protocol is extensible, i.e. it is easy to add new functionalities without disturbing the overall architecture. can be replaced by any new or existing protocol dealing with multimedia information retrieval.

The main processes are on the server side. They can be separated in two main categories:

- off-line processes: data analysis, features extraction and structuration
- on-line processes: answer the client requests

Several clients can communicate with the server. A good starting point for exploring the possibilities offered by IKONA is our web demo, available at http://www-roc.inria.fr/cgi-bin/imedia/circario.cgi/bio_diversity?select_db=1. This CGI client is connected to a running server with several generalist and specific image databases, including more than 23,000 images. It features query by example searches, switch database functionality and relevance feedback for image category searches. The second client is a desktop application. It offers more functionalities. More screen-shots describing the visual searching capabilities of IKONA are available at <http://www-rocq.inria.fr/imedia/cbir-demo.html>.

IKONA is a pre-industrial prototype, with exploitation as a final objective. Currently, there does not exist a licensed competitor with the same range of functionalities. It exists several commercial softwares or systems exploiting technologies similar to some functionalities of IKONA but usually not the most advanced ones. We can for example cite the SDK developed by LTU company, the service proposed by AdVestigo company, etc. Many prototypes and demonstrators, industrial or academic, share some functionalities of IKONA but here again not the most advanced (e.g. Google Image Similarity Search Beta, IBM Muffin, etc.).

The main originality of IKONA is its **genericity** (in terms of visual features, metrics, input data, storage format, etc.), its **adaptivity** (to new visual features, new indexing structures or new search algorithms), its innovative **interactive** search functionalities (Local and Global Relevance Feedback, Local Search with Query Expansion, Search results clustering, etc.) and its **scalability** thanks to a generic indexing structure module than can support the integration of any new advances.

Current Users of IKONA include European and National Projects Participants through its integration in prototype multimedia systems, commercial companies through user trials (EXALEAD, INA, BELGA, AFP), General or Specific Public through Web demos (PI@ntNet leaf identification demo).

IKONA software provides a high degree of visibility to IMEDIA scientific works through demos in commercial, scientific and general public events (notably in most INRIA national showrooms). It is also the mainstay of several Multimedia Systems developed at the European level, in conjunction with many Leader European Companies and Research Centers.

6. New Results

6.1. Feature space modelling

6.1.1. A novel shape boundary based description for leaf identification

Participants: Itheri Yahiaoui, Olfa Mzoughi, Nozha Boujema.

The problem of automatic leaf identification is particularly difficult for two main reasons: (i) the first is the enormous number of leaf species and (ii) the second, which is relevant for some special species but more complex, is the high inter-species and the low intra-species similarity.

Our research has focused on analysing leaf morphology in order to determine a numeric key description for leaf species robust to all the above mentioned constraints. The approach that we propose is a shape boundary description that combines two complementary information: (i) the first one outlines local variations of the leaf margin. This is performed using the Directional Fragment Histogram (DFH), introduced in [25], which encodes the relative frequency distribution of groups of contour points with uniform orientation, (ii) the second property emphasizes the spatial distribution of contour points (in terms of distances). This is done by comparing the shape to standard geometric ones (such as circle, rectangle, ellipse, convex hull, etc.).

This descriptor was evaluated within the framework of ImageCLEF 2011 plant task where a crowd-sourced database, called PI@ntLeaves [13], was used and a high number of image retrieval techniques was tested (a total of 8 groups from all around the world that have submitted 20 runs [19]). Our descriptor brought the best rate for scan-like pictures and was close to the best rate for scan pictures. Besides to the accuracy, this descriptor requires very low computational time, which accomplishes a basic condition for real world application.

6.1.2. Visual-based plant species identification from crowdsourced data

Participants: Hervé Goëau, Alexis Joly, Souheil Selmi, Laurent Joyeux.

Inspired by citizen sciences, the main goal of this work is to speed up the collection and integration of raw botanical observation data, while providing to potential users an easy and efficient access to this botanical knowledge. We therefore designed and developed an original crowdsourcing web application dedicated to the access of botanical knowledge through automated identification of plant species by visual content.

Technically, the first side of the application deals with content-based identification of plant leaves. Whereas state-of-the-art methods addressing this objective are mostly based on leaf segmentation and boundary shape features, we developed a new approach based on local features and large-scale matching. This approach obtained the best results within one sub-task of ImageCLEF 2011 plant identification benchmark [19]. The second side of the application deals with interactive tagging and allow any user to validate or correct the automatic determinations returned by the system.

Overall, this collaborative system allows to enrich automatically and continuously the visual botanical knowledge and therefore to increase progressively the accuracy of the automated identification. A demo of the developed application was presented at ACM Multimedia conference [13]. This work was done in collaboration with INRIA team ZENITH and with the botanists of the AMAP UMR team (CIRAD). It is also closely related to a citizen science project around plant's identification that we developed with the support of the TelaBotanica social network inside the PI@ntNet project.

6.1.3. Spatial relations between salient points on a leaf

Participants: Sofiène Mouine, Itheri Yahiaoui, Anne Verroust-Blondet.

In the scope of the PI@ntNet project, our recent work has consisted in finding spatial relationships between salient points on a leaf. As a first step, classic detectors were used to find significant points in the leaf area and then the Shape context descriptor, originally applied on contour points, was introduced to measure a spatial relation between interest points. We have tested different configurations by varying the set of voting points. First results confirm that including spatial relations enriches the local description of each point. We are currently improving a veins and landmark extraction approach [12] in order to include also veins points in the voting set.

6.1.4. 3D mesh segmentation by example

Participants: Esma Elghoul, Anne Verroust-Blondet.

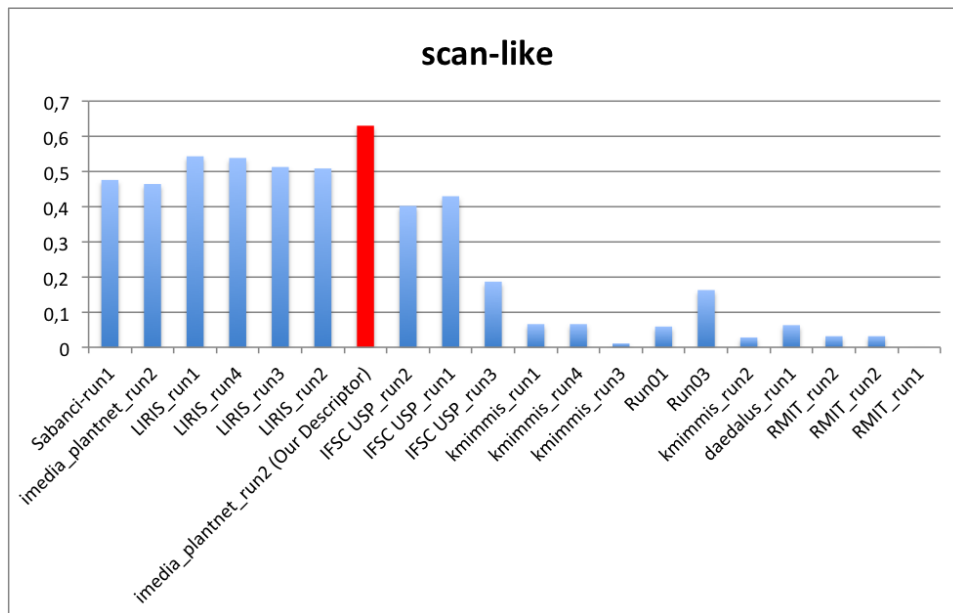
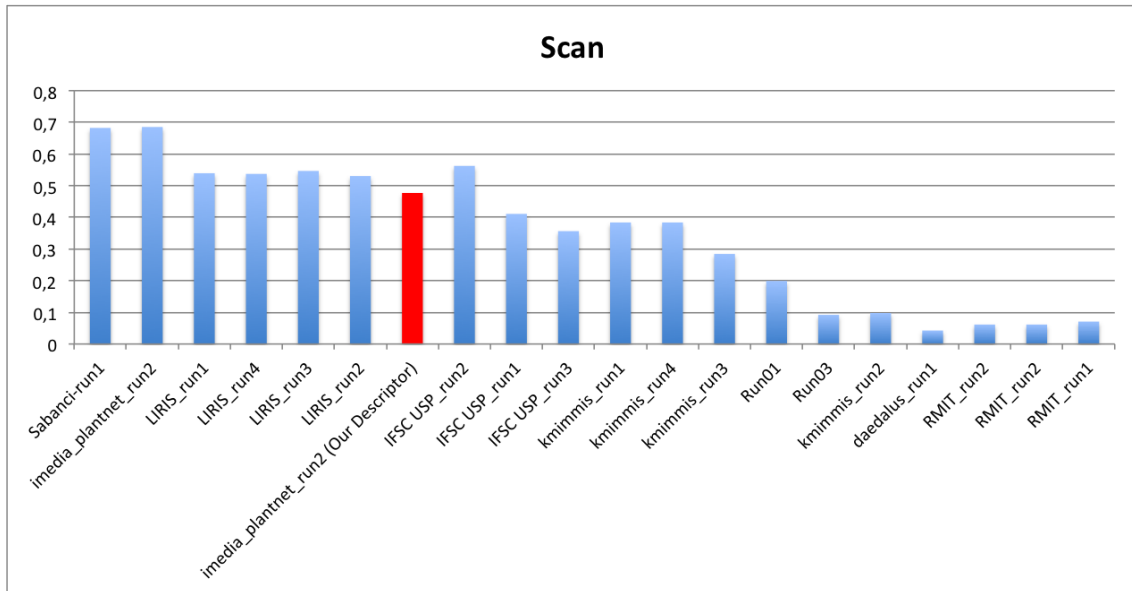


Figure 1. Normalized classification scores for scan images (up) and scan-like images (down). Our approach bin is in red, other ImageCLEF methods in blue.

In recent years, there has been an increasing interest for automatic 3D segmentation. Indeed, segmentation of 3D objects is an important step in many applications such as part indexing of 3D objects, pattern recognition, compression, morphing, texture mapping and simplification. It refers to the process of partitioning 3D shapes into multiple parts, based on semantic criteria and/or geometric criteria.

Our work consists in introducing an approach to segment a 3D object class referring to a given segmented object from this class (we called it segmentation by example). The considered segmentation method is not automatic: we want to use interactive tools that proved advantageous to segment a 3D shape into relevant parts.

As a first task, we reviewed the state of the art in 3D segmentation techniques recently proposed in the literature. The different techniques were evaluated and classified for the purpose of choosing the more appropriate one for our work. We opted for extending the technique of random walks [24] to build an interactive tool of 3D segmentation.

For the second task, we had to solve a basic problem: that of similarly direct objects belonging to the same class. Indeed, each 3D model is provided in a random orientation in the space. In order to align objects of a same class, we used the alignment approach developed in [22] which computes 3 alignment axes. To properly orient our objects between them, we had to develop an additional process to the last one. It combines a 2D ICP and a 3D ICP approaches to give the best orientation among 48 possibilities and pair each two objects meshes.

So having a user-supplied already segmented model (model (1)) and a second model (model (2)) belonging to the same class (not segmented but similarly oriented), we developed a method to put into correspondence segmented parts of model (1) with faces of model (2). Then we computed a segmentation of model (2) using a derivative approach of the random walks. We applied this technique as well to segment all the objects that belong to the class of model (1). Our approach provides good results for manufactured object classes such as chairs and tables.

6.2. Feature space structuring

6.2.1. Random Maximum Margin Hashing

Participant: Alexis Joly.

Following the success of hashing methods for multidimensional indexing, more and more works are interested in embedding visual feature space in compact hash codes. Such approaches are not an alternative to using index structures but a complementary way to reduce both the memory usage and the distance computation cost. Several data dependent hash functions have notably been proposed to closely fit data distribution and provide better selectivity than usual random projections such as LSH. However, improvements occur only for relatively small hash code sizes up to 64 or 128 bits due to the lack of independence between the produced hash functions. In this work, we introduced a new hash function family that attempts to solve this issue in any kernel space. Rather than boosting the collision probability of close points, this method focus on data scattering. By training purely random splits of the data, regardless the closeness of the training samples, it is indeed possible to generate consistently more independent hash functions. On the other side, the use of large margin classifiers allows to maintain good generalization performances. Experiments did show that our new Random Maximum Margin Hashing scheme (RMMH) outperforms four state-of-the-art hashing methods, notably in kernel spaces. Overall, this new concept of randomly trained classifiers opens the door to many other problems including large-scale learning, visual vocabulary construction or distributed content-based retrieval methods. A paper describing RMMH was published in the proceedings of CVPR 2011 [14].

6.2.2. Scalable information retrieval in distributed architectures

Participants: Mohamed Riadh Trad, Alexis Joly, Nozha Boujemaa.

Organizing media according to the occurrence of real-life events is attracting increasing interest in the multimedia community. However, whereas text based methods are now mature enough to deal with huge datasets, there are still some challenging issues managing multimedia contents. This becomes even more challenging in the context of User Generated Contents. Low-level visual metadata are indeed not simple textual or scalar values, their management requires efficient similarity search in high dimensional spaces.

Similarity search in high dimensional spaces has been the focus of many works in the database community in the recent years. State-of-the-art methods focus mainly on space partitioning techniques and more recently on hash-based probabilistic algorithms.

Although, hash-based approaches proved to be scalable, the computational cost is still too high for some real world applications and K-Nearest Neighbours Graph constructed can be more desirable than the costly online K-NN search. In fact, the basic LSH algorithm partitions the space uniformly and thus it does not exploit the clustering property of the data, which may result in slow query response and wasted space with additional hash tables. These limitations were pointed out with our scalable prototype for large scale event matching [18].

Scaling up LSH-based techniques and applications is then closely related to buckets occupations and objects distribution within the index structure. Recent works achieve better data distribution over the buckets with guarantees on occupation. As one result, we easily bound the similarity join size and evaluate bound algorithms complexity.

Based on these works, we designed and implemented a scalable prototype for distributed similarity search and K-NN graph construction. We have made several experiments querying real world large datasets. The prototype proved to be efficient for both search and K-NN graph construction.

Ongoing experiments process a 1.2 million images dataset. Results will be submitted for publication.

6.2.3. Visual similarity sensitive hashing methods for semantic image search in very large collections of images

Participants: Saloua Ouertani-Litayem, Alexis Joly, Nozha Boujemaa.

With the rapid development of information acquisition technology, we have witnessed an explosive growth in the scale of shared data collections. Then, it is now possible to tackle fundamental problems with very large datasets' context. Especially those addressing challenging tasks in machine learning for developing large scale approaches for multimedia retrieval and mining. Computer Vision is experiencing this paradigm shift, with large annotated image and video datasets becoming available. Indeed, various benchmark datasets for image classification have been released such as image-net and LabelMe. Therefore, a key challenge is taken up through out the Phd aiming to build efficient methods for training and matching efficiently very large collections of images.

We proposed several SVM-based strategies to build new supervised hash function families from large annotated collections of features. We indeed investigated with an approach consisting in benefiting from different embedding approaches in order to build compact codes indexed with efficient similarity search structures. Therefore, we have extended a kernelized hashing method [14] with multi-class SVM to solve a K-class classification problem by choosing the maximum applied to the outputs of K SVMs. We indeed proposed hashing methods based on the multi-class SVM classification strategies: One vs One (OVO) And One vs All (OVA). An important task during this process was to experimentally evaluate the quality lost induced by such representations with respect to the efficiency gains. We then compared multi-class SVM strategies with different underlying kernels.

Inspired by state of the art hashing in kernel space methods we investigated an approach consisting of benefiting from both semantic hashing like techniques and kernel embedding approach in order to build compact category aware codes indexed with efficient similarity. Experiments, are performed on image-net ILSVRC 2010 dataset [23]. Results will be submitted for publication.

6.3. Pattern recognition and statistical learning

6.3.1. Machine identification of biological shapes

Participants: Asma Rejeb Sfar, Donald Geman, Nozha Boujemaa.

Stored images of biological objects are accumulating at a staggering rate due to new sensor technologies, expanding use in medical diagnostics, web-based search engines and growing demands for web-based services in traditional sciences such as botany. These developments have been accompanied by an increasing demand for the automated analysis of these data, such as counting cell types, detecting lesions and other abnormalities in medical images, and identifying botanical shapes.

All these tasks have one feature in common: massive diversity among the shapes. Indeed, such shapes display enormous within-class variation and are generally highly deformable. Also, they often exhibit a hierarchical organization resulting from evolutionary processes.

There is currently no existing methodology in image analysis and computer vision which can be applied to a multi-class shape recognition problem of this complexity. Consequently, there is a need for a new, generic methodology for categorizing hierarchically-structured families of deformable shapes, particularly when both the number of categories and the within-category variation are very large.

We proposed a coarse-to-fine (CTF) approach in both shape representation and image parsing. The representation is hierarchical in both class and pose.

We focused on botanical shapes, specifically categorizing simple leaves according to species. So, we determined a suitable representation for the pose of a simple leaf and designed and tested a two-stage pose detector. Then, we constructed classifiers based on the plant taxonomy.

Results will be submitted for publication.

6.4. Interactive search and personalisation

6.4.1. Database denoising and multi visual queries

Participant: Sébastien Poullot.

One of IMEDIA's task inside the SCARFACE project is to introduce and develop a character retrieval system. For this purpose, we take as entries the tracking of the persons in video sequences computed by Thalès and construct a database of the profiles. A profile is a 3D frame, a bounding box that changes along the time line. Two original works have been proposed for searching in the profile database.

The first one consists in analysing features of each profile with respect to all the profiles in order to extract relevant features from it, and construct more representative databases.

The second one is to be able to search inside the database with a set of queries (pictures of the same person). An a priori work can be done on this set of queries in order to extract the relevant features (and remove the irrelevant ones). On the other side an a posteriori work can be done on late merging depending on the specificities of each sub query.

TRECVID Instance Search 2011

Before starting the developments for SCARFACE, we tested various algorithms in TRECVID 2011 INS (instance search) task. This task is close to the SCARFACE one: from a set of captures of one object, one should find its occurrences in a set of video sequences. This work has been done during the stay of Sébastien Poullot at NII (the Japan National Institute of Informatics) in July and August 2011.

The differences with SCARFACE are:

- a high diversity in the type of the objects (people but also, places, vehicle, animals, etc),
- the location of the object in the database is not given.

Our approach obtains good results (above the median scores of all teams) and works in a very short time (and without indexing system for speeding up the process) [15]. The choice for SCARFACE's method partially depends on these results. We still continue on the INS task in order to achieve better scores (various descriptors and various post and late fusion between sub queries).

Query generative models

Moreover, in order to enhance visual query results, we want to create some visual query generative models. It is directly linked to SCARFACE (a priori processes) and TRECVID works: given a set of images (considered as queries), we extract what gather them and what separate them in order to construct artificial relevant queries. For now we essentially work on some logo databases.

6.4.2. *Object-based Visual Query Suggestion*

Participants: Amel Hamzaoui, Pierre Letessier, Alexis Joly, Nozha Boujemaa.

After our work on the shared neighbours clustering methods in multi-sources case published in [10], we are interested now to the case of a bipartite graph that we apply to object-based visual query suggestion using the visual words mining technique [16]. In fact, state-of-the-art visual search systems allow to retrieve efficiently small rigid objects in very large datasets. They are usually based on the query-by-window paradigm: a user selects any image region containing an object of interest and the system returns a ranked list of images that are likely to contain other instances of the query object. User's perception of these tools is however affected by the fact that many submitted queries actually return nothing or only junk results (complex non-rigid objects, higher-level visual concepts, etc..). We address the problem of suggesting only the object's queries that actually contain relevant matches in the dataset. This requires to first discover accurate object's clusters in the dataset (as an off-line process); and then to select the most relevant objects according to user's intent (as an online process). We therefore introduce a new object's instances clustering framework based on two main contributions: efficient object's seeds discovery with adaptive weighted sampling and bipartite shared-neighbours clustering. Experiments show that this new method outperforms state-of-the-art object mining and retrieval results on OxfordBuilding dataset. We finally describe two object-based visual query suggestion scenarios using the proposed framework and show examples of suggested object queries.

6.4.3. *Interpretable Visual Models for Human Perception-based Object Retrieval*

Participants: Ahmed Rebai, Alexis Joly, Nozha Boujemaa.

Understanding the results returned by automatic visual concept detectors is often a tricky task making users uncomfortable with these technologies. In this work we attempt to build humanly interpretable visual models, allowing the user to visually understand the underlying semantic. We therefore proposed a supervised multiple instance learning algorithm that selects as few as possible discriminant local features for a given object category. The method finds its roots in the lasso theory where a L1-regularization term is introduced in order to constraint the loss function, and subsequently produce sparser solutions. Efficient resolution of the lasso path is achieved through a boosting-like procedure inspired by BLasso algorithm. Quantitatively, the method achieved similar performance as current state-of-the-art, and qualitatively, it allows users to construct their own model from the original set of patches learned, thus allowing for more compound semantic queries. This work is part of the PhD of Ahmed Rebai [8] and it was published in ICMR 2011 proceedings [17]. This work was then extended to using geometrically checked feature sets rather than using single local features to describe the content of visual patches. We did show that this allows drastically reducing the number of the selected visual words while improving their interpretability. A publication was submitted to pattern recognition journal [11].

6.4.4. *Relevance feedback on local features : Application to plants annotations and identification*

Participants: Wajih Ouertani, Michel Crucianu, Nozha Boujemaa.

As biological image databases are increasing rapidly, automated species identification based on digital data is of great interest for accelerating biodiversity assessment, researches and monitoring. In this context, our work falls within an investigation of computer vision techniques or more precisely: object recognition and content based image retrieval techniques to help botanist identifying and organizing his digital images' collections. Under believe that perception, recognition and decision are parts of human skills, this work focus on an interactive mechanism which tries to extract useful information from the user and gives him help to deal with large data amount. We adopted an explicit relevance feedback (RF) schema and we worked on extending it to deal with local intention through local features (LF) description. This mechanism helps discovering

and dynamically defining new concept and interesting plant parts and feed identification ways interactively. Moreover since it relies to the content rather than labels one direct application is to fill the initially sparse annotation space with right annotations and in a reasonable time and with the introduce of one or many expertises. We recently explored and tested images local features matching involving high order features and non-rigid adaptation tentative to structure database with a patterns' discovery stage. Using those type of methods we expect to introduce a high level appearance information that tends to go beyond classical bag of features and histogram based distances at least from semantic gap and interpretation point of view. We argue our exploration way with the fact that initial search space can be exceedingly rich. By pre-structuring it we can hope to obtain a smaller search space together with more reliable inference. Also learning parts interactively with localized local features may require a lot of interaction since it requires a considerable number of examples. We experienced the design of combined machine learning and prior mining of matches which we are actually improving.

6.5. Software

6.5.1. IKONA/MAESTRO software

Participants: Vera Bakić, Jean-Paul Chièze, Raffi Enficiaud, Alexis Joly, Laurent Joyeux, Pierre Letessier, Olfa Mzoughi, Wajih Ouertani, Ahmed Rebai, Souheil Selmi.

This year, IKONA has been extended in the context of PI@ntNet, Glocal, I-SEARCH and R2I projects. For the PI@ntNet project, along the continuing improvements in the MAESTRO software, a number of new features were added. Namely the support for the automatic image segmentation and subsequent use of segmented regions; descriptors with the various shape's geometric parameters; use of multiple orientations for Harris points; run-time additions to the external database and immediate availability of the new images for the search; descriptors to facilitate external data usage; colour SIFT and Affine Covariance descriptors; integration of the thesis work of Ahmed Rebai for objects retrieval; and tools for statistical tests.

In addition, a number of new web services were developed and deployed: the dynamic indexation system of the on-line pad ("carnet-en-ligne" of Tela Botanica)images; the search with multiple views; the update of PI@ntNet internal demonstration allowing to present features such as visual similarity search, textual search, filtering (pre- and post-filtering), and different methods of research; the implementation of the organ prediction web service and other web services of botanical information statistics; the administration of the indexation system and the experimentation of new research methods (GPS spatial and temporal search).

For the Glocal project, an interface was developed for the demonstrations of a search engine in large scale events database (the queries are event images and the result is a list of the closest events in terms of time alignment and image content), and new web services were developed and updated according to the data exchange format and the middle-ware of the project - among others: fraud detection, import media from the web, associate media with existing event in the repository, and event matching web services. The queries are composed of either a medium link (an external image) or an event link (set of external images).

For the I-SEARCH project, an integration was performed to provide global and local 2D image low level descriptors. For videos, an automated extraction of visual words tool was integrated to show to users image patches which are the most meaningful.

For the R2I project, a detailed technical documentation of the procedure of maestro's installation, web services and tomcat server were provided to Exalead partner.

7. Partnerships and Cooperations

7.1. National Initiatives

7.1.1. PI@ntNet project [2009-2013]

It is a joint project with AMAP (CIRAD, INRA, IRD, Montpellier) and Tela Botanica, an international botanical network with 8,500 members and an active collaborative web platform (10,000 visits /day). The project has its financial support from Agropolis International Foundation (<http://www.agropolis.fr/>) and is titled “Plant Computational Identification and Collaborative Information System”.

7.1.2. *Other collaborations with INRA*

The PhD thesis of Wajih Ouertani, financed by INRA, in the context of a strategic collaboration between INRIA and INRA, addresses interactive species identification through advanced relevance feedback mechanisms based on local image information.

7.1.3. *ANR project SCAR-FACE [2008-2011]*

SCAR-FACE (*Semantic Characterization And Retrieval of FACES*) objective is to develop new interactive technologies for recognizing people in public places provided with videosurveillance networks.

Other partners: Univ Caen - INRIA LEAR, EADS, SPIKENET, IREENAT

IMEDIA activities within the project started with the arrival of Sébastien Poullot (see section 6.4.1).

7.2. European Initiatives

7.2.1. *FP7 Projects*

7.2.1.1. *CHORUS+*

Title: CHORUS+ Network of Audio-Visual Media Search

Type: CAPACITIES (ICT)

Defi: Networked Medias & 3D Internet

Instrument: Coordination and Support Action (CSA)

Duration: January 2010 - December 2012

Coordinator: JCP-Consult (France)

Others partners: UNITN (Italy), HES-so (Switzerland), Thomson R&D (France), JCPC (France), CERTH (Greece), TU Wien (Austria), ENG (Italy), IPTS (Belgium)

See also: <http://www.ist-chorus.org/>

Abstract: CHORUS + has been funded in the continuity of the former CHORUS initiative thanks to its success. Beyond CHORUS coordination objectives, CHORUS+ includes new key issues such as extended cooperation and coordination to Asian countries and US, support to integration and implementation, support to coordinated research evaluations or support to results dissemination of EU projects. Nozha Boujemaa and Alexis Joly are part of the management board of the project.

7.2.1.2. *GLOCAL*

Title: Glocal (*Event-Based Retrieval of Networked Media*)

Type: COOPERATION (ICT)

Defi: Networked Medias & 3D Internet

Instrument: Integrated Project (IP)

Duration: December 2009 - November 2012

Coordinator: Univ. Degli Studi di Trento (Italy)

Others partners: UNITN (Italy), ISOCO (Spain), ALINARI (Italy), CERTH (Greece), Yahoo Iberia SL (Spain), AFP (France), DFKI (Germany), Exalead (France), LUH (Germany), BUT (Czech Republic)

See also: <http://www.glocal-project.eu/>

Abstract: The key idea underlying the project is to use events as the primary means for organizing and indexing media. Within networked communities, common (global) descriptions of the world can be built and continuously enriched by a continuous flow of individual (local) descriptions. With two leading search companies and four content providers, the consortium attempts to realize and evaluate this approach in several application domains, which will involve professional and amateur users dealing with professional and generic contents. IMEDIA is responsible of three research tasks related to visual-based event indexing, retrieval and mining, notably in distributed contexts.

7.2.1.3. I-SEARCH

Title: I-SEARCH (*A unified framework for multimodal content SEARCH*)

Type: COOPERATION (ICT)

Defi: Networked Medias & 3D Internet

Instrument: Specific Targeted Research Project (STREP)

Duration: January 2010 - December 2012

Coordinator: CENTRE FOR RESEARCH AND TECHNOLOGY HELLAS (Greece)

Others partners: CERTH (Greece), JCPC (France), ATTC (Greece), ENG (Italy), Google (Germany), UNIGE (Italy), Exalead (France), FHE (Germany), ANSC (Italy), EGR (Germany)

See also: <http://www.isearch-project.eu/isearch/>

Abstract: The I-SEARCH project aims to provide a novel unified framework for multimodal content indexing, sharing, search and retrieval. The I-SEARCH framework will be able to handle specific types of multimedia and multimodal content (text, 2D image, sketch, video, 3D objects and audio) alongside with real world information, which can be used as queries and retrieve any available relevant content of any of the aforementioned types. IMEDIA is workpackage leader of "RUCOD COMPLIANT Descriptor Extraction".

7.3. International Initiatives

7.3.1. Visits of International Scientists

Don Geman from John Hopkins University.

7.3.1.1. Internship

Olfa MZOUGHFI (from March 2011 until August 2012)

Subject: Analyse et description de la morphologie foliaire: Application à la classification et l'identification d'espèces de plantes

Institution: Université de Tunis El Manar - Faculté des Sciences (Tunisia)

8. Dissemination

8.1. Animation of the scientific community

8.1.1. Nozha Boujemaa

- Re-elected member in the Steering Board of NEM ETP (Networked and Electronic Media European Technology Platform) and acting as INRIA representative.
- Scientific coordinator of PI@ntNet.
- Member of the steering committee of CHORUS+.

- Founding member of the ACM ICMR "ACM International Conference on Multimedia Retrieval" born from the fusion of: ACM MIR (International Conference on Multimedia Information Retrieval) and ACM CIVR (International Conference on Image and Video Retrieval)
- Member of the steering committee of ACM ICMR.
- Responsibilities within INRIA: Director of the INRIA Saclay Ile-de-France research centre.

8.1.2. Alexis Joly

- Member of the steering committee of PI@ntNet: WP2 co-leader.
- Member of the steering committee of CHORUS+ and of GLOCAL.
- Member of the organizing committee of ImageCLEF evaluation forum (<http://www.imageclef.org/2011>) and co-organizer, with Hervé Goëau, of an evaluation task on plant images retrieval within the 2011 evaluation campaign of ImageCLEF.
- Chair of a community networking session on the access to scientific multimedia data within CLEF 2011 international conference, http://avmediasearch.eu/index.php?article_id=246&page=115&action=article&=
- Member of the conference program committee of NEM Summit 2011.

8.1.3. Anne Verroust-Blondet

- Member of the steering committee of PI@ntNet (WP5 co-leader: "Dissemination").
- Member of the steering committee of I-SEARCH (WP4 leader).
- Member of the Humanities and Social Sciences committee for the "Blanc" and "Young researcher" 2011 programmes of the French National Research Agency (ANR).
- Member of the steering committee of the CNRS GDR IG (Informatique Graphique) ;
- Member of the technical programme committee of the Eurographics Workshop on 3D Object Retrieval in 2011 and of the Third International Conference on Advances in Multimedia (MMEDIA 2011),
- Member of the editorial board of the "Revue Electronique Francophone d'Informatique Graphique".

8.2. Teaching

Mohamed Riadh Trad:

Master: Advanced Java, 27 hours/year, M1 level, Paris Dauphine University, France.

Master: JAVA, 27 hours/year, M1 & M2 levels, Paris Dauphine University, France.

Amel Hamzaoui:

Licence: Algorithm and programming, 48 hours/year, L1 level, Orsay University Paris 11, France

Licence: Web programming, 24 hours/year, L1 level, Orsay University Paris 11, France

Sofiène Mouine

Licence: C2I, 36h, level (L2), University Paris 1 (Panthéon-Sorbonne), France.

PhD & HdR:

PhD : Ahmed Rebai, Interactive Object Retrieval using Interpretable Visual Models, Université Paris Sud - Paris XI, May 2011, Nozha Boujemaa and Alexis Joly

PhD in progress : Esma Elghoul, Segmentation d'Objets 3D - Indexation Partielle, October 2010, Anne Verroust-Blondet

PhD in progress : Amel Hamzaoui, Shared-Neighbours methods for visual content structuring and mining, October 2008, Nozha Boujemaa and Alexis Joly

PhD in progress : Pierre Letessier, Analyse, Structuration et Enrichissement de contenus vidéos à large échelle Exploitation de statistiques à grande échelle, October 2009, Nozha Boujemaa, Olivier Buisson and Alexis Joly

PhD in progress : Sofière Mouine, Geometric models for a local description of images, October 2010, Anne Verroust-Blondet and Itheri Yahiaoui

PhD in progress : Saloua Ouertani-Litayem, Large scale supervised image retrieval through hashing methods, October 2009, Nozha Boujemaa and Alexis Joly

PhD in progress : Wajih Ouertani, Relevance feedback on local features : Application to plants annotations and identification, October 2008, Nozha Boujemaa and Michel Crucianu

PhD in progress : Asma Rejeb Sfar, Machine Identification of Biological Shapes, October 2010, Nozha Boujemaa and Donald Geman

PhD in progress : Mohamed Riadh Trad, Scalable information retrieval in distributed architectures, October 2009, Nozha Boujemaa and Alexis Joly

9. Bibliography

Major publications by the team in recent years

- [1] J. AMORES, N. SEBE, P. RADEVA. *Context-Based Object-Class Recognition and Retrieval by Generalized Correlograms*, in "IEEE Transactions on Pattern Analysis and Machine Intelligence", 2007, vol. 29, n° 10, p. 1818 – 1833.
- [2] M. CHAOUCH, A. VERRROUST-BLONDET. *Alignment of 3D models*, in "Graphical Models", March 2009, vol. 71, n° 2, p. 63–76.
- [3] Y. FANG, D. GEMAN, N. BOUJEMAA. *An Interactive System for Mental Face Retrieval*, in "7th ACM SIGMM International Workshop on Multimedia Information Retrieval", Singapore in conjunction with ACM Multimedia 2005, November 10–11 2005.
- [4] J. FAUQUEUR, N. BOUJEMAA. *Mental image search by boolean composition of region categories*, in "Multimedia Tools and Applications", September 2006, p. 95-117.
- [5] M. FERECATU, N. BOUJEMAA, M. CRUCIANU. *Semantic interactive image retrieval combining visual and conceptual content description*, in "ACM Multimedia Systems", 2007.
- [6] N. GRIRA, M. CRUCIANU, N. BOUJEMAA. *Active semi-supervised fuzzy clustering*, in "Pattern Recognition", 2008, vol. 41, n° 5, p. 1834-1844.

Publications of the year

Doctoral Dissertations and Habilitation Theses

- [7] A. REBAI. *Interactive Object Retrieval using Interpretable Visual Models*, Université Paris Sud - Paris XI, May 2011, <http://hal.inria.fr/tel-00608467/en>.
- [8] A. REBAI. *Recherche Interactive d'Objets à l'Aide de Modèles Visuels Interprétables*, Université Paris Sud - Paris XI, May 2011, <http://hal.inria.fr/tel-00596916/en>.

Articles in International Peer-Reviewed Journal

- [9] P. DARAS, A. AXENOPOULOS, V. DARLAGIANNIS, D. TZOVARAS, X. LE BOURDON, L. JOYEUX, A. VERRONST-BLONDET, V. CROCE, T. STEINER, A. MASSARI, A. CAMURRI, S. MORIN, A.-D. MEZAOUR, L. SUTTON, S. SPILLER. *Introducing a Unified Framework for Content Object Description*, in "International Journal of Multimedia Intelligence and Security", 2011, to appear.
- [10] A. HAMZAOU, A. JOLY, N. BOUJEMAA. *Multi-source shared nearest neighbours for multi-modal image clustering*, in "Multimedia Tools Appl.", 2011, vol. 51, n^o 2, p. 479-503.
- [11] A. REBAI, A. JOLY, N. BOUJEMAA. *BLasso for Object Categorization and Retrieval: Towards Interpretable Visual Models*, in "Pattern Recognition", 2011, to appear.

International Conferences with Proceedings

- [12] R. ENFICIAUD, S. MOUINE. *Landmark extraction from leaves with palmate venation*, in "ICPRAM", 2012, to appear.
- [13] H. GOËAU, A. JOLY, S. SELMI, P. BONNET, E. MOUYSET, L. JOYEUX. *Visual-based plant species identification from crowdsourced data*, in "ACM Multimedia 2011", Scottsdale, United States, November 2011, <http://hal.inria.fr/hal-00642236/en>.
- [14] A. JOLY, O. BUISSON. *Random Maximum Margin Hashing*, in "CVPR'11 - IEEE Computer Vision and Pattern Recognition", Colorado springs, United States, June 2011, p. 873-880, <http://hal.inria.fr/hal-00642178/en/>.
- [15] D.-D. LE, C.-Z. ZHU, S. POULLOT, S. SATOH. *National Institute of Informatics, Japan at TRECVID 2011*, in "TRECVID 2011", Gaithersburg, Maryland, USA, December 2011.
- [16] P. LETESSIER, O. BUISSON, A. JOLY. *Consistent visual words mining with adaptive sampling*, in "ICMR", Trento, Italy, 2011, <http://hal.inria.fr/hal-00642202/en/>.
- [17] A. REBAI, A. JOLY, N. BOUJEMAA. *Interpretable Visual Models for Human Perception-Based Object Retrieval*, in "ICMR 2011", Trento, Italy, April 2011, <http://hal.inria.fr/hal-00642232/en>.
- [18] M. R. TRAD, A. JOLY, N. BOUJEMAA. *Large scale visual-based event matching*, in "ICMR", Trento, Italy, 2011, <http://hal.inria.fr/hal-00642210/en/>.

Conferences without Proceedings

- [19] H. GOËAU, P. BONNET, A. JOLY, N. BOUJEMAA, D. BARTHÉLÉMY, J.-F. MOLINO, P. BIRNBAUM, E. MOUYSSET, M. PICARD. *The ImageCLEF 2011 plant images classification task*, in "ImageCLEF", Amsterdam, Netherlands, September 2011, <http://hal.inria.fr/hal-00642197/en>.
- [20] H. GOËAU, A. JOLY, I. YAHIAOUI, P. BONNET, E. MOUYSSET. *Participation of INRIA & Pl@ntNet to ImageCLEF 2011 plant images classification task*, in "ImageCLEF 2011", Amsterdam, Netherlands, September 2011, <http://hal.inria.fr/hal-00642239/en>.

References in notes

- [21] N. BOUJEMAA, J. FAUQUEUR, M. FERECATU, F. FLEURET, V. GOUET-BRUNET, B. LE SAUX, H. SAHBI. *Ikona: Interactive specific and generic image retrieval*, in "International workshop on Multimedia Content-Based Indexing and Retrieval (MMCBIR'2001)", 2001.
- [22] M. CHAOUCH, A. VERRONST-BLONDET. *Alignment of 3D models*, in "Graphical Models", 2009, vol. 71, p. 63-76.
- [23] J. DENG, W. DONG, R. SOCHER, L.-J. LI, K. LI, L. FEI-FEI. *ImageNet: A large-scale hierarchical image database*, in "IEEE Conference on Computer Vision and Pattern Recognition (2009)", 2009, p. 248–255.
- [24] Y.-K. LAI, S.-M. HU, R. MARTIN, P. L. ROSIN. *Rapid and effective segmentation of 3D models using random walks*, in "Computer Aided Geometric Design", August 2009, vol. 26, n^o 6, p. 665-679.
- [25] I. YAHIAOUI, N. HERVÉ, N. BOUJEMAA. *Shape-based image retrieval in botanical collections*, in "Lecture Notes in Computer Science including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics", 2006, vol. 4261 LNCS, p. 357-364.