



IN PARTNERSHIP WITH:
CNRS

**Institut national des sciences
appliquées de Rennes**

Université Rennes 1

Activity Report 2014

Project-Team LINKMEDIA

Content-based multimedia linking and collection structuring

IN COLLABORATION WITH: Institut de recherche en informatique et systèmes aléatoires (IRISA)

RESEARCH CENTER
Rennes - Bretagne-Atlantique

THEME
**Vision, perception and multimedia
interpretation**

Table of contents

1. Members	1
2. Overall Objectives	2
2.1. Context	2
2.2. Scientific objectives	2
3. Research Program	3
3.1. Scientific background	3
3.2. Workplan	3
3.2.1. Unsupervised motif discovery	3
3.2.2. Describing and structuring	4
3.2.3. Linking	4
4. Application Domains	4
4.1. Asset management in the entertainment business	4
4.2. Multimedia Internet	5
4.3. Multiscreen TV	5
4.4. E-learning	5
5. New Software and Platforms	5
5.1. Software	5
5.1.1.1. Peyote	5
5.1.1.2. Yael	5
5.2. The AllGO web services	6
5.3. Experimental platform	6
6. New Results	6
6.1. Highlights of the Year	6
6.2. Unsupervised motif discovery	6
6.2.1. Clustering by diverting supervised machine learning	6
6.2.2. Spoken term discovery applied to audio thumbnailing	6
6.2.3. Unsupervised video structure mining with grammatical inference	7
6.2.4. Efficient indexing for content retrieval	7
6.3. Multimedia content description and structuring	7
6.3.1. Linguistic knowledge extraction	7
6.3.1.1. Identifying events in texts	7
6.3.1.2. Morpho-semantic analysis of terms	8
6.3.1.3. Distributional semantics	8
6.3.2. OCR and speech content-based description	8
6.3.2.1. Use of stress information for robust speech recognition	8
6.3.2.2. Boosting bonsai trees for handwritten/printed text discrimination	8
6.3.2.3. Speaker role detection from spoken document	9
6.3.3. Image and video description and classification	9
6.3.3.1. Fine-grain image classification	9
6.3.3.2. Aggregation of local descriptors	9
6.3.3.3. Action localization in videos	9
6.3.4. Text description for information retrieval	10
6.4. Linking, navigation and analytics	10
6.4.1. NLP-driven hyperlink construction in broadcast videos	10
6.4.2. Analytics in collections of art critics	10
6.4.3. Data models for navigation	11
6.4.4. Exploiting k-nn graphs for image retrieval	11
6.5. Content-based information retrieval and security	11
6.5.1. Analysis of privacy preserving data aggregation for recommendation systems	11

6.5.2.	Content based image retrieval with privacy	12
6.5.3.	Privacy protection in face recognition	12
6.5.4.	Tardos code in practice	12
7.	Bilateral Contracts and Grants with Industry	13
8.	Partnerships and Cooperations	13
8.1.	Regional Initiatives	13
8.1.1.	CominLabs Project CominWeb	13
8.1.2.	CominLabs Project Linking Media in Acceptable Hypergraphs (LIMAH)	13
8.2.	National Initiatives	13
8.2.1.	ANR Project FIRE-ID	13
8.2.2.	ANR Project Secular	14
8.3.	European Initiatives	14
8.3.1.1.	VIAMASS	14
8.3.1.2.	Forensic Image Identifier and Analyzer	14
8.4.	International Initiatives	15
8.4.1.	Inria Associate Teams	15
8.4.2.	Inria International Partners	15
8.4.3.	Other International Programs	15
8.5.	International Research Visitors	15
8.5.1.	Visits of International Scientists	15
8.5.2.	Internships	16
8.5.3.	Visits to International Teams	16
9.	Dissemination	16
9.1.	Promoting Scientific Activities	16
9.1.1.	Scientific events organisation	16
9.1.1.1.	General chair, Scientific chair	16
9.1.1.2.	Member of the organizing committee	16
9.1.2.	Scientific events selection	17
9.1.2.1.	Responsibility in conference program committee	17
9.1.2.2.	Membership of conference program committee	17
9.1.2.3.	Reviewer	17
9.1.3.	Journal	17
9.1.3.1.	Member of the editorial board	17
9.1.3.2.	Reviewer	18
9.2.	Teaching - Supervision - Juries	18
9.2.1.	Teaching	18
9.2.2.	Supervision	19
9.2.3.	Juries	19
9.3.	Community animation	20
9.4.	Popularization	20
10.	Bibliography	20

Project-Team LINKMEDIA

Keywords: Multimedia, Computer Vision, Natural Language, Data Mining, Machine Learning

This report covers the activities of Texmex from Jan. 2014 to June 2014 and of Linkmedia from June 2014.

Creation of the Project-Team: 2014 July 01.

1. Members

Research Scientists

Laurent Amsaleg [CNRS, Researcher, HdR]
Vincent Claveau [CNRS, Researcher]
Teddy Furon [Inria, Researcher]
Guillaume Gravier [CNRS, Senior Researcher, Team leader, HdR]
Patrick Gros [Inria, Senior Researcher, HdR]
Hervé Jégou [Inria, Researcher, HdR]

Faculty Members

Philippe-Henri Gosselin [ENSEA, Professor, Inria chair until Aug. 2014, HdR]
Ewa Kijak [Univ. Rennes I, Associate Professor]
Simon Malinowski [Univ. Rennes I, Associate Professor]
Christian Raymond [INSA Rennes, Associate Professor]
Pascale Sébillot [INSA Rennes, Professor, HdR]
François Poulet [Univ. Rennes I, Associate Professor, until June 2014, HdR]

Engineers

Sébastien Champion [Inria]
Jonathan Delhumeau [Inria]
Mathieu Desoubieux [Inria]
Sébastien Le Maguer [Inria, from Feb. 2014 to Sep. 2014]
Arthur Masson [Inria, until Aug. 2014]
Charly Maupetit [Inria, from Oct. 2014]

PhD Students

Raghavendran Balu [Inria]
Rémi Bois [CNRS, from Oct. 2014]
Mohamed-Haykel Boukadida [Orange Labs]
Ahmet Iscen [Inria, from Feb. 2014]
Himalaya Jain [Inria, from Dec. 2014]
Mihir Jain [Inria, until Apr. 2014]
Abir Ncibi [Inria]
Bingqing Qu [INA]
Anca-Roxana Şimon [Univ. Rennes I]
Vedran Vukotić [INSA Rennes, from May 2014]

Post-Doctoral Fellows

Andrei Bursuc [Inria, from Mar. 2014]
Gylfi Gudmundsson [Inria]
Ronan Sicre [Inria, from Oct. 2014]
Giorgos Toliás [Inria]
Li Weng [Inria, from May 2014]

Administrative Assistant

Aurélie Patier [Univ. Rennes I]

2. Overall Objectives

2.1. Context

Linked media appears today as a major challenge, with numerous potential applications in all areas of multimedia. The strong increase of ubiquitous access to the Internet and the resulting convergence of media on the network open countless opportunities for linked media and reinforce the key role of such a challenge. New applications centered on the notion of linked media are emerging today, such as second screen applications and recommendation services. However, because of the lack of adequate technology, linking related content is mostly deferred to human operators in current applications or to user behavior analysis, e.g., via collaborative filtering, thus indirectly considering the content. This fact severely limits the opportunities offered by a web of media, in terms of creativity, scalability, representativeness and completeness, thus negatively impacting the spread of linked media and the development of innovative services in the Internet of media.

Most of the research effort in automatic multimedia content analysis has been devoted so far to describing and indexing content on which core tasks around information retrieval and recommendation are built to develop multimedia applications. This general philosophy mostly reposes on a vision where documents are considered as isolated entities, i.e., as a basic unit which is indexed or analyzed regardless of other content items and of context. Considering documents in isolation has enabled key progress in content-based analysis and retrieval on a large scale: e.g., design of generic descriptors, efficient techniques for content-based analysis, fast retrieval methodology. But ignoring the links, implicit or explicit, between content items also appears as a rather strong assumption with straightforward consequences on algorithms and applications, both in terms of performance and in terms of possibilities.

2.2. Scientific objectives

LinkMedia investigates a number of key issues related to multimedia collections structured with explicit links: Can we discover what characterizes a collection and makes its coherence? Are there repeating motifs that create natural links and which deserve characterization and semantic interpretation? How to explicitly create links from pairwise distances? What structure should a linked collection have? How do we explain the semantic of a link? How explicit links can be used to improve information retrieval? To improve user experience? In this general framework, the global objective of LinkMedia is to develop the scientific, methodological and technological foundations facilitating or automating the creation, the description and the exploitation of multimedia collections structured with explicit links. In particular, we target a number of key contributions in the following areas:

- developing efficient methods dedicated to multimedia indexing and unsupervised motif discovery: efficiently comparing content items on a large scale and finding repeating motifs in an unsupervised manner are two key ingredients of multimedia linking based on a low-level representation of the content;
- improving techniques for structuring and semantic description: better description of multimedia content at a semantic—i.e., human interpretable—level, making explicit the implicit structure when it exists, is still required to make the most of multimedia data and to facilitate the creation of links to a precise target at a semantic level;
- designing and experimenting approaches to multimedia content linking and collection structuring: exploiting low-level and semantic content-based proximity to create explicit links within a collection requires specific methodology departing from pairwise comparison and must be confronted with real data;
- studying new paradigms for the exploitation of linked multimedia content as well as new usages: explicit links within media content collections change how such data is processed by machines and ultimately consumed by humans in ways that have yet to be invented and studied.

3. Research Program

3.1. Scientific background

LINKMEDIA is a multidisciplinary research team, with multimedia data as the main object of study. We are guided by the data and their specificity—semantically interpretable, heterogeneous and multimodal, available in large amounts, unstructured and disconnected—, as well as by the related problems and applications.

With multimedia data at the center, orienting our choices of methods and algorithms and serving as a basis for experimental validation, the team is directly contributing to the following scientific fields:

- multimedia: content-based analysis; multimodal processing and fusion; multimedia applications;
- computer vision: compact description of images; object and event detection;
- natural language processing: topic segmentation; information extraction;
- information retrieval: high-dimensional indexing; approximate k-nn search; efficient set comparison;

LINKMEDIA also takes advantage of advances in the following fields, adapting recent developments to the multimedia area:

- signal processing – image processing; compression;
- machine learning – deep architectures; structured learning; adversarial learning;
- security – data encryption; differential privacy;
- data mining – time series mining and alignment; pattern discovery; knowledge extraction;

3.2. Workplan

Research activities in LINKMEDIA are organized along three major lines of research which build upon the scientific domains already mentioned.

3.2.1. *Unsupervised motif discovery*

As an alternative to supervised learning techniques, unsupervised approaches have emerged recently with the goal of discovering directly patterns and events of interest from the data, in a totally unsupervised manner. In the absence of prior knowledge on what we are interested in, meaningfulness can be judged based on one of three main criteria: unexpectedness, saliency and recurrence. This last case posits that repeating patterns, known as motifs, are potentially meaningful, leading to recent work on the unsupervised discovery of motifs in multimedia data [77], [75], [76].

LINKMEDIA seeks to *develop unsupervised motif discovery approaches which are both accurate and scalable*. In particular, we consider the discovery of repeating objects in image collections and the discovery of repeated sequences in video and audio streams. Research activities are organized along the following lines:

- developing the scientific basis for scalable motif discovery: sparse histogram representations; efficient co-occurrence counting; geometry and time aware indexing schemes;
- designing and evaluating accurate and scalable motif discovery algorithms applied to a variety of multimedia content: exploiting efficient geometry or time aware matching functions; fast approximate DTW; symbolic representations of multimedia data, in conjunction with existing symbolic data mining approaches;
- developing methodology for the interpretation, exploitation and evaluation of motif discovery algorithms in various use-cases: image classification; video stream monitoring; transcript-free NLP for spoken document;

3.2.2. Describing and structuring

Content-based analysis has received a lot of attention from the early days of multimedia, with an extensive use of supervised machine learning for all modalities [78], [72]. Progress in large scale entity and event recognition in multimedia content has made available general purpose approaches able to learn from very large data sets and performing fairly decently in a large number of cases. Current solutions are however limited to simple, homogeneous, information and can hardly handle structured information such as hierarchical descriptions, tree-structured or nested concepts.

LINKMEDIA aims at *expanding techniques for multimedia content modeling, event detection and structure analysis*. The main transverse research lines that LINKMEDIA will develop are as follows:

- context-aware content description targeting (homogeneous) collections of multimedia data: latent variable discovery; deep feature learning; motif discovery;
- secure description to enable privacy and security aware multimedia content processing: everaging encryption and diversity; exploring adversarial machine learning in a multimedia context; privacy-oriented image processing;
- multilevel modeling with a focus on probabilistic modeling of structured multimodal data: multiple kernels; structured machine learning; conditionnal random fields;

3.2.3. Linking

Creating explicit links between media content items has been considered on different occasions, with the goal of seeking and discovering information by browsing, as opposed to information retrieval via ranked lists of relevant documents. Content-based link creation has been initially addressed in the hypertext community for well-structured texts [71] and was recently extended to multimedia content [79], [74], [73]. The problem of organizing collections with links remains mainly unsolved for large heterogeneous collections of unstructured documents, with many issues deserving attention: linking at a fine semantic grain; selecting relevant links; characterizing links; evaluating links; etc.

LINKMEDIA targets pioneering research on media linking by **developing scientific ground, methodology and technology for content-based media linking** directed to applications exploiting rich linked content such as navigation or recommendation. Contributions are concentrated along the following lines:

- algorithmic of linked media for content-based link authoring in multimedia collections: time-aware graph construction; multimodal hypergraphs; large scale k-nn graphs;
- link interpretation and characterization to provide links semantics for interpretability: text alignment; entity linking; intention vs. extension;
- linked media usage and evaluation: information retrieval; summarization; data models for navigation; link prediction;

4. Application Domains

4.1. Asset management in the entertainment business

Regardless of the ingestion and storage issues, media asset management—archiving, describing and retrieving multimedia content—has turned into a key factor and a huge business for content and service providers. Most content providers, with television channels at the forefront, rely on multimedia asset management systems to annotate, describe, archive and search for content. So do archivists such as the Institut National de l'Audiovisuel, the Nederlands Instituut voor Beeld en Geluid or the British Broadcast Corporation, as well as media monitoring companies, such as Yacast in France. Protecting copyrighted content is another aspect of media asset management.

4.2. Multimedia Internet

One of the most visible application domains of linked multimedia content is that of multimedia portals on the Internet. Search engines now offer many features for image and video search. Video sharing sites also feature search engines as well as recommendation capabilities. All news sites provide multimedia content with links between related items. News sites also implement content aggregation, enriching proprietary content with user-generated content and reactions from social networks. Most public search engines and Internet service providers offer news aggregation portals.

4.3. Multiscreen TV

The convergence between television and the Internet has accelerated significantly over the past few years, with the democratization of TV on-demand and replay services and the emergence of social TV services and multiscreen applications. These evolutions and the ever growing number of innovative applications incurred offer a unique playground for multimedia technologies. Recommendation plays a major role in connected TV. Enriching multimedia content, with explicit links targeting either multimedia material or knowledge databases, appears as a key feature in this context, at the core of rich TV and second screen applications.

4.4. E-learning

On-line courses are rapidly gaining interest with the recent movement for massive open on-line courses (MOOCs). Such courses usually aggregate multimedia material, such as a video of the course with handouts and potentially text books, exercises and other related resources. This setting is very similar to that of the media aggregation sites though in a different domain. Automatically analyzing and describing video and textual content, synchronizing all material available across modalities, creating and characterizing links between related material or between different courses are all necessary features for on-line courses authoring.

5. New Software and Platforms

5.1. Software

5.1.1. News and updates

5.1.1.1. Peyote

Participants: Sébastien Champion, Hervé Jégou [correspondent].

Peyote is a framework for Video and Image description, indexation and nearest neighbor search. It can be used as-is by a video-search or image-search front-end with the implemented descriptors and search modules. It can also be used via scripting for large-scale experimentation. Finally, it is modular and as such can be used for scientific experimentation on new descriptors or indexation methods. Peyote is used in the AABOT software and was used for the Mediaeval Placing task and the Trecvid Instance Search task.

Peyote is used by Lamark, a start-up that is currently being incubating at Inria Rennes.

Last APP deposit: IDDN.FR.001.420008.001.S.A.2012.000.21000.

5.1.1.2. Yael

Participant: Hervé Jégou [correspondent].

This software is jointly maintained by Matthijs Douze, from Inria Grenoble.

Yael [30] is a C/python/Matlab library providing implementations of computationally demanding functions. In particular, the library provides very optimized functions for k-means clustering and exact nearest neighbor search. It is maintained and continuously improved. This year, we have in particular added a few tutorials implementing two simple image search systems, see http://yael.gforge.inria.fr/tutorial/tuto_imgindexing.html.

The current release (v401) was registered at APP under no IDDN.FR.001.220014.002.S.P.2010.000.10000 on July 2014.

5.2. The AllGO web services

Participants: Sébastien Campion [correspondent], Guillaume Gravier.

Available at <http://allgo.irisa.fr>, the AllGO platform allows for the easy deployment of the technology developed in the team as web services. The engineer hired by SED in October 2013 developed several new features that enable software providers to deploy autonomously their algorithm. Dedicated hardware equipment was also purchased in 2014, composed by a main server with 1.3 TB of storage and 3 nodes for computing task and setup, and will be available in 2015. In addition to a strong involvement in the development of the platform, LINKMEDIA contributed several services.

5.3. Experimental platform

Participant: Sébastien Campion [correspondent].

Our experimental platform, consisting of dedicated equipments to experiment on very large collections of multimedia data, was upgraded in 2014. In order to replace old hardware, we acquired 380 TB of additional disk space. Divided in two categories, the first part (140 TB) must replace in 2015 our current network area storage where datasets are stored. The second part (240 TB) is dedicated to a distributed storage filesystem (CEPH), used to store our experimental results with high IO performances for use with the mesocluster IGRIDA.

6. New Results

6.1. Highlights of the Year

BEST PAPER AWARD :

[21] **Challenging differential privacy: the case of non-interactive mechanisms in European Symposium on Research in Computer Security.** R. BALU, T. FURON, S. GAMBS.

6.2. Unsupervised motif discovery

6.2.1. Clustering by diverting supervised machine learning

Participants: Vincent Claveau, Patrick Gros, Abir Ncibi.

Knowledge discovery aims at bringing out coherent groups of objects and is usually based on clustering which necessitates defining a notion of similarity between objects. In practice, this strong prior is often neither possible nor welcome. We proposed to divert supervised machine learning (ML) techniques in order to calculate, indirectly and without supervision, similarities among objects. Our approach consists in generating artificial labeling problems on the data to reveal regularities between objects through their labeling. In [28], we show how this framework can be implemented and experimented on two information extraction/discovery tasks concerned with named entities. The ML technique diverted to exhibit similarities between with the named entities are the Conditional Random Fields. This same method can also be applied with less common ML techniques: In [59], we show that Inductive Logic Programming can also be used to cluster complex data. Thanks to the ability of ILP to handle data that cannot be expressed under the usual attribute-value representation, we use it to make emerge clusters of TV broadcasts based only on their broadcasting information (date, time, length, etc.).

6.2.2. Spoken term discovery applied to audio thumbnailing

Participants: Sébastien Campion, Guillaume Gravier.

We evaluated a system to create audio thumbnails of spoken content, i.e., short audio summaries representative of the entire content, without resorting to a lexical representation. As an alternative to searching for relevant words and phrases in a transcript, unsupervised motif discovery is used to find short, word-like, repeating fragments at the signal level without acoustic models. The output of the word discovery algorithm is exploited via a maximum motif coverage criterion to generate a thumbnail in an extractive manner. A limited number of relevant segments are chosen within the data so as to include the maximum number of motifs while remaining short enough and intelligible. Evaluation is performed on broadcast news reports with a panel of human listeners judging the quality of the thumbnails. Results indicate that motif-based thumbnails stand between random thumbnails and ASR-based keywords, however still far behind thumbnails and keywords humanly authored [34].

6.2.3. *Unsupervised video structure mining with grammatical inference*

Participants: Guillaume Gravier, Bingqing Qu.

In collaboration with Jean Carrive and Félicien Vallet, Institut National de l'Audiovisuel.

Unsupervised approaches were introduced a few years ago to analyze the structure of TV programs, relying on the discovery of repeated elements within a program or across multiple episodes of the same program. These methods can discover key repeating elements, such as jingles and separators, however they cannot infer the entire structure of a program. In [48], we studied a hierarchical use of grammatical inference to yield a temporal grammar of a program from a collection of episodes, discovering both the vocabulary of the grammar and the temporal organization of the words from the vocabulary. Using a set of basic event detectors and simple filtering techniques to detect repeating elements of interest, a symbolic representation of each episode is derived based on minimal domain knowledge. Grammatical inference based on multiple sequence alignment is then used in a hierarchical manner to provide a temporal grammar of the program at various levels of details.

6.2.4. *Efficient indexing for content retrieval*

Participants: Raghavendran Balu, Teddy Furon, Hervé Jégou.

In collaboration with Miajing Shi during, visiting Ph. D. student from Peking University.

Many nearest neighbor search algorithms rely on encoding real vectors into binary vectors. The most common strategy projects the vectors onto random directions and takes the sign to produce so-called sketches. In [22], we discuss the sub-optimality of this choice, and propose a better encoding strategy based on the quantization and reconstruction points of view. Our second contribution is a novel asymmetric estimator for the cosine similarity. Similar to previous asymmetric schemes, the query is not quantized and the similarity is computed in the compressed domain. We tackled the same similarity estimation problem with a rather different approach in [52], where we assume that only a few vectors of the database, so-called heavy hitters, have a similarity to the query that significantly deviates from 0. For this purpose, we have introduced a group testing framework for detecting large similarities between high-dimensional vectors, such as descriptors used in state-of-the-art description of multimedia documents. We produce a set of group representations that jointly encode several vectors into a single one, in the spirit of group testing approaches. By comparing a query vector to several of these intermediate representations, we screen the large values taken by the similarities between the query and all the vectors, at a fraction of the cost of exhaustive similarity calculation. Unlike concurrent indexing methods that suffer from the curse of dimensionality, our method exploits the properties of high-dimensional spaces.

6.3. Multimedia content description and structuring

6.3.1. *Linguistic knowledge extraction*

6.3.1.1. *Identifying events in texts*

Participant: Vincent Claveau.

In collaboration with Béatrice Arnulphy, former team member now with ANR, Xavier Tannier and Anne Vilnat, LIMSI.

Identifying events from texts is an information extraction task necessary for many NLP applications. Through the TimeML specifications and TempEval challenges, it has received some attention in the last years, yet, no reference result was available for French. In [20], we tried to fill this gap by proposing several event extraction systems, combining for instance Conditional Random Fields, language modeling and k-nearest-neighbors. These systems are evaluated on French corpora and compared with state-of-the-art methods on English. The very good results obtained on both languages validate our whole approach and set new standard for French.

6.3.1.2. Morpho-semantic analysis of terms

Participants: Vincent Claveau, Ewa Kijak.

In most Indo-European languages, many biomedical terms are rich morphological structures composed of several constituents mainly originating from Greek or Latin. The interpretation of these compounds are keystones to access information. Following our work on morphology in the biomedical domain, we proposed different techniques to generate probabilistic morph-semantic resources and we show how these alignment information can be used for segmenting compounds, attaching semantic interpretation to each part, proposing definitions (gloses) of the compounds [26]. When possible, these tasks are compared with state-of-the-art tools, and the results show the interest of our automatically built probabilistic resources.

6.3.1.3. Distributional semantics

Participants: Vincent Claveau, Ewa Kijak.

In collaboration with Olivier Ferret, CEA-LIST.

We addressed the issue of building and improving a distributional thesaurus. We first show that existing tools from the information retrieval domain can be directly used in order to build a thesaurus with state-of-the-art performance. Secondly, we focus more specifically on improving the obtained thesaurus, seen as a graph of k-nearest neighbors. By exploiting information about the neighborhood contained in this graph, we propose several contributions. 1) We show how the lists of neighbors can be globally improved by examining the reciprocity of the neighboring relation, that is, the fact that a word can be close to another and vice-versa. 2) We also propose a method to associate a confidence score to any lists of nearest neighbors (i.e., any entry of the thesaurus). 3) Last, we demonstrate how these confidence scores can be used to reorder the closest neighbors of a word. These different contributions are validated through experiments and offer significant improvement over the state-of-the-art [27], [60].

6.3.2. OCR and speech content-based description

6.3.2.1. Use of stress information for robust speech recognition

Participant: Guillaume Gravier.

In collaboration with S. Ziegler, PANAMA Inria team and Laboratoire de Sciences Cognitives et Psycholinguistique.

[44] presents a study on the robustness of stress information for automatic speech recognition in the presence of noise. The syllable stress, extracted from the speech signal, was integrated in the recognition process by means of a previously proposed decoding method. Experiments were conducted for several signal-to-noise ratio conditions and the results show that stress information is robust in the presence of medium to low noise. This was found to be true both when syllable boundary information was used for stress detection and when this information was not available. Furthermore, the obtained relative improvement increased with a decrease in signal quality, indicating that the stressed parts of the signal can be considered islands of reliability.

6.3.2.2. Boosting bonsai trees for handwritten/printed text discrimination

Participant: Christian Raymond.

In collaboration with Yann Ricquebourg, Baptiste Poirriez, Aurélie Lemaitre and Bertrand Couasnon, IRISA.

Boosting over decision-stumps proved its efficiency in natural language processing, essentially with symbolic features, and its good properties (fast, few and not critical parameters, not sensitive to overfitting) could be of great interest in the numeric world of pixel images. In [51], we investigated the use of boosting over small decision trees in image classification processing for the discrimination of handwritten/printed text. We conducted experiments to compare with usual SVM-based classification revealing convincing results with very close performance, but with faster predictions and behaving far less as a black-box. Those promising results tend to make use of this classifier in more complex recognition tasks like multiclass problems.

6.3.2.3. *Speaker role detection from spoken document*

Participant: Christian Raymond.

In collaboration with LIMSI and LIUM.

In [40] and [41], we tackle the problem of speaker role detection in broadcast news shows. In the literature, many proposed solutions are based on the combination of various features coming from acoustic, lexical and semantic information with a machine learning algorithm. Many previous studies mention the use of boosting over decision stumps to combine efficiently these features. We proposed a modification of this state-of-the-art machine learning algorithm changing the weak learner (decision stumps) by small decision trees, denoted bonsai trees. Experiments show that using bonsai trees as weak learners for the boosting algorithm largely reduces both system error rate and learning time.

6.3.3. *Image and video description and classification*

6.3.3.1. *Fine-grain image classification*

Participants: Teddy Furon, Philippe-Henri Gosselin, Hervé Jégou.

In collaboration with Xerox Research Center Europe.

We have addressed the problem of instance classification: our goal is to annotate images with tags corresponding to objects classes which exhibit small intra-class variations such as logos, products or landmarks. Our first contribution on image classification [13] describes the processing pipeline, which has won FGCOMP challenge associated with Imagenet. It improves a standard method based on Fisher vectors to adapt it to the context of fine-grained classes, where the difference between classes rely on few but typical visual differences. On the same task, we have proposed a novel algorithm [39] for the selection of class-specific *prototypes* which are used in a voting-based classification scheme.

6.3.3.2. *Aggregation of local descriptors*

Participants: Teddy Furon, Hervé Jégou, Giorgos Tolias.

In collaboration with the University of Oxford.

For unsupervised particular object and image recognition, we have considered the design of a single vector representation for an image that embeds and aggregates a set of local patch descriptors such as SIFT. In [36], we make two contributions, both aimed at regularizing the individual contributions of the local descriptors in the final representation. The first is a novel embedding method that avoids the dependency on absolute distances by encoding directions. The second contribution is a “democratization” strategy that further limits the interaction of unrelated descriptors in the aggregation stage. In [36], we addressed another issue inherent to existing encoding algorithms: Image search systems based on local descriptors typically achieve orientation invariance by aligning the patches on their dominant orientations. This choice introduces too much invariance because it does not guarantee that the patches are rotated consistently. To address this problem, we have introduced another aggregation strategy of local descriptors that achieves this covariance property by jointly encoding the angle in the aggregation stage in a continuous manner. It is combined with an efficient monomial embedding to provide a codebook-free method to aggregate local descriptors into a single vector representation.

6.3.3.3. *Action localization in videos*

Participants: Mihir Jain, Hervé Jégou.

In collaboration with the University of Amsterdam and the project-team SERPICO.

We have tackled the problem of action localization in videos [35], where the objective is to determine when and where certain actions appear. We introduce a sampling strategy, called tubelets and inspired a method recently introduced for image detection. It drastically reduces the number of hypotheses that are likely to include the action of interest. By using super-voxels and employing a criterion that reflects how action related motion deviates from background motion, the method is specifically adapted to 2D+t sequences and establishes the new state-of-the-art for action localization on the public datasets UCF Sports and MSR-II.

6.3.4. Text description for information retrieval

Participants: Vincent Claveau, Sébastien Le Maguer.

In collaboration with Natalia Grabar, STL UMR8163, and Thierry Hamon, LIMSI

Following previous work, we investigated the interest of “bag of bags of features” representation for texts in an vector-space information retrieval setting. Each text is thus represented as a bag of vector. With this representation, computing the similarity between two texts necessitates to aggregate every vector to vector similarity for the two bags. In [58], we examine the expected properties of such an aggregation function and show their influence through different experiments. When some specific conditions are met, we show that the gains over standard representation can be very important.

With a team composed with members of TEXMEX/LINKMEDIA, LIMSI and STL, we have participated to the biomedical information retrieval challenge proposed in the framework of CLEF eHealth [25]. For this first participation, our approach relies on a state-of-the-art IR system called Indri, based on statistical language modeling, and on semantic resources. The purpose of semantic resources and methods is to manage the term variation such as synonyms, morpho-syntactic variants, abbreviation or nested terms. Different combinations of resources and Indri settings are explored, mostly based on query expansion. We obtained good overall results (3rd in terms of MAP) and confirmed the interest of query expansion to retrieve a maximum of relevant documents.

6.4. Linking, navigation and analytics

6.4.1. NLP-driven hyperlink construction in broadcast videos

Participants: Rémi Bois, Vincent Claveau, Guillaume Gravier, Pascale Sébillot, Anca-Roxana Şimon.

In collaboration with Sien Moens, Katholieke Universiteit Leuven, Éric Jamet and Martin Ragot, Univ. Rennes 2.

The hyperlinking sub-task of the MediaEval Search and Hyperlinking task aims at creating hyperlinks between predefined anchor segments, i.e., fragments of videos, and short related video segments, called targets, that have to be automatically extracted from videos of a large collection. Capitalizing on the experience acquired in previous participations [54], we proposed a two step approach exploiting speech material: Potential target segments are first generated relying on a topic segmentation technique; For each anchor, the best 20 target segments are then selected according to two distinct strategies. The first strategy focuses on the identification of very similar targets using n-grams and named entities, while the second one makes use of an intermediate structure built from topic models, which offers the possibility to control serendipity and to explain the links created [53].

In 2014, we also initiated the CominLabs project “Linking media in acceptable hypergraphs” dedicated to the creation of explicit and meaningful links between multimedia documents or fragments of documents. Two main issues were adressed: The construction of a corpus, composed of audio and video news, reports and debates, newspapers and blog websites, as well as social networks; A preliminary study of the perceived usefulness of various types of links by end-users.

6.4.2. Analytics in collections of art critics

Participant: Vincent Claveau.

In collaboration with Fabienne Moreau and Nicolas Thély, Univ. Rennes 2.

We aim at exploiting text mining techniques in the service of digital humanities, and more precisely in the field of art criticism. It relies on a collaboration between our team, linguists and art and aesthetics specialists. In preliminary work [56], we adapted term extraction, named entity recognition and information retrieval techniques to this field to extract multiple linguistic clues from art review articles. Future work will make the most of these clues and clustering approaches to build a navigable and structured collection of the articles.

6.4.3. Data models for navigation

Participant: Laurent Amsaleg.

In collaboration with Björn Þór Jónsson, Grímur Tómasson, Hlynur Sigurþórsson, Áslaug Eriksdóttir and Marta Kristin Larusdóttir, School of Computer Science, Reykjavík University.

Digital photo collections—personal, professional, or social—have been growing ever larger, leaving users overwhelmed. It is therefore increasingly important to provide effective browsing tools for photo collections. Learning from the resounding success of multi-dimensional analysis (MDA) in the business intelligence community for On-Line Analytical Processing (OLAP) applications, we propose a multi-dimensional model for media browsing, called M^3 , that combines MDA concepts with concepts from faceted browsing. We present the data model and describe preliminary evaluations, made using server and client prototypes, which indicate that users find the model useful and easy to use [38]. A photo navigation prototype was demonstrated at the Intl. Conf. on Multimedia Modeling [37].

6.4.4. Exploiting k -nn graphs for image retrieval

Participants: Laurent Amsaleg, Hervé Jégou, Giorgos Tolias.

We have proposed two techniques exploiting the relationship between the images with a collection. In [29], we revisit how to exploit the k -reciprocal nearest neighbors to produce, for a given query, a neighborhood that improves over the one obtained with the original metric. This strategy is simpler than concurrent prior work, yet it is both effective and less sensitive to parameters. Second, we propose to employ measures defined on sets of shared nearest neighbors in order to re-rank the shortlist. Both methods are simple, yet they significantly improve the accuracy of image search engines on standard benchmarks. We also introduced a query expansion technique [18] for image search that is faster and more precise than the existing ones. The expansion generates an enriched representation which refines the initial local descriptors *individually* by aggregating those of the database, while new descriptors are produced from the images that are deemed relevant. The technique has two computational advantages over other query expansion techniques. First, the size of the enriched representation is comparable to that of the initial query. Second, the technique is effective even without using any geometry, in which case searching a database comprising 105k images typically takes 80 ms on a desktop machine. Overall, our technique significantly outperforms the visual query expansion state of the art on popular benchmarks.

6.5. Content-based information retrieval and security

6.5.1. Analysis of privacy preserving data aggregation for recommendation systems

Participants: Raghavendran Balu, Teddy Furon.

Work under the Alcatel-Lucent / Inria common Lab, in collaboration with Armen Aghasaryan, Dimitre Davidov, Makram Bouzid (ALU) and Sébastien Gambs (Inria Rennes Cidre team project).

We consider personalized recommendation systems in which before publication, the profile of a user is sanitized by a non-interactive mechanism compliant with the concept of differential privacy. We analyze two existing schemes offering a differentially private representation of profiles: BLIP (BLoom-and-FLIP) and JLT (Johnson-Lindenstrauss Transform). For assessing their security levels, we play the role of an adversary aiming at reconstructing a user profile [21]. We compare two inference attacks, namely single and joint decoding. The first one decides of the presence of a single item in the profile, and sequentially explores all the item set. The second one decides whether a subset of items is likely to be the user profile, and considers all the possible

subsets. Our contributions are a theoretical analysis as well as a practical implementation of both attacks, which were evaluated on datasets of real user profiles. The results obtained clearly demonstrates that joint decoding is the most powerful attack, while also giving useful insights on how to set the differential privacy parameter ϵ .

6.5.2. *Content based image retrieval with privacy*

Participants: Laurent Amsaleg, Teddy Furon, Li Weng.

Work initiated during a collaboration with A. Morton, L. Weng (with LINKMEDIA since May 2014) and S. Marchand-Maillet, Université de Genève.

We propose a privacy protection framework for large-scale content-based information retrieval. It offers two layers of protection. First, robust hash values are used as queries to prevent revealing original content or features. Second, the client can choose to omit certain bits in a hash value to further increase the ambiguity for the server. Due to the reduced information, it is computationally difficult for the server to know the client's interest. The server has to return the hash values of all possible candidates to the client. The client performs a search within the candidate list to find the best match. Since only hash values are exchanged between the client and the server, the privacy of both parties is protected. We introduce the concept of tunable privacy, where the privacy protection level can be adjusted according to a policy. It is realized through hash-based piece-wise inverted indexing. The idea is to divide a feature vector into pieces and index each piece with a sub-hash value. Each sub-hash value is associated with an inverted index list. The framework has been extensively tested using a large image database. We have evaluated both retrieval performance and privacy-preserving performance for a particular content identification application.

6.5.3. *Privacy protection in face recognition*

Participant: Teddy Furon.

In collaboration with B. Bhattarai, A. Mignon, F. Jurie, GREYC, Université of Caen.

We investigated a new approach for de-identifying face images, *i.e.* for preventing automatic matching with public face collections. The overall motivation is to offer tools for privacy protection on social networks. We address this question by drawing a parallel between face de-identification and oracle attacks in digital watermarking. In our case, the identity of the face is seen as the watermark to be removed. Inspired by oracle attacks, we forge de-identified faces by superimposing a collection of carefully designed noise patterns onto the original face. The modification of the image is controlled to minimize the probability of good recognition while minimizing the distortion. In addition, these de-identified images are by construction made robust to counter attacks such as blurring. We present an experimental validation in which we de-identify LFW faces and show that resulting images are still recognized by human beings while deceiving a state-of-the-art face recognition algorithm [23].

6.5.4. *Tardos code in practice*

Participant: Teddy Furon.

Joint work with the Inria supported start-up LAMARK.

We deal with active fingerprinting a.k.a. traitor tracing where a collusion of dishonest users merges their individual versions of a content to yield a pirated copy. The Tardos codes are one of the most powerful tools to fight against such collusion process by identifying the colluders. Instead of studying as usual the necessary and sufficient code length in a theoretical setup, we adopt the point of view of the practitioner. We call this the *operational mode*, *i.e.* a practical setup where a Tardos code has already been deployed and a pirated copy has been found. This new paradigm shows that the known bounds on the probability of accusing an innocent in the theoretical setup are way too pessimistic. Indeed the practitioner can resort to much tighter bounds because the problem is fundamentally much simpler under the operational mode. In the end, we benchmark under the operational mode several single decoders recently proposed in the literature [32].

7. Bilateral Contracts and Grants with Industry

7.1. Bilateral Contracts with Industry

- CIFRE Ph. D. contract with Institut National de l'Audiovisuel (Bingqing Qu)
- CIFRE Ph. D. contract with Institut National de l'Audiovisuel (Ludivine Kuznik)
- CIFRE Ph. D. contract with Orange (Mohamed-Haykel Boukadida)
- CIFRE Ph. D. contract with Technicolor (Himalaya Jain)

8. Partnerships and Cooperations

8.1. Regional Initiatives

8.1.1. *CominLabs Project CominWeb*

Participants: Vincent Claveau, Sébastien Le Maguer.

Duration: 1 year

Partners: Univ. Nantes

URL: <http://www.cominweb.cominlabs.ueb.eu>

In the framework of the CominWeb projet, a 50kE contract was granted by the Labex CominLabs to the team to carry a preliminary study about text similarity models in different contexts: information retrieval, content based recommendation, etc.

8.1.2. *CominLabs Project Linking Media in Acceptable Hypergraphs (LIMAH)*

Participants: Rémi Bois, Sébastien Campion, Vincent Claveau, Guillaume Gravier, Patrick Gros, Pascale Sébillot.

Duration: 4 years, started in April 2014

Partners: Telecom Bretagne (IODE), Univ. Rennes II (CRPCC, PREFics), Univ. Nantes (LINA/TAL)

URL: <http://limah.irisa.fr>

LIMAH aims at exploring hypergraph structures for multimedia collections, instantiating actual links reflecting particular content-based proximity—similar content, thematic proximity, opinion expressed, answer to a question, etc. Exploiting and developing further techniques targeting pairwise comparison of multimedia contents from an NLP perspective, LIMAH addresses two key issues of content-based graph-oriented multimedia collection structuring: How to automatically build from a collection of documents an hypergraph, i.e., graph combining edges of different natures, which provides exploitable links in selected use cases? How collections with explicit links modify usage of multimedia data in all aspects, from a technology point of view as well as from a user point of view? LIMAH studies hypergraph authoring and acceptability taking a multidisciplinary approach mixing ICT, law, information and communication science as well as cognitive and ergonomics psychology.

8.2. National Initiatives

8.2.1. *ANR Project FIRE-ID*

Participants: Sébastien Campion, Philippe-Henri Gosselin, Patrick Gros, Hervé Jégou.

Duration: 3 years, started in May 2012

Partner: Xerox Research Center Europe

The FIRE-ID project considers the semantic annotation of visual content, such as photos or videos shared on social networks, or images captured by video surveillance devices or scanned documents. More specifically, the project considers the fine-grained recognition problem, where the number of classes is large and where classes are visually similar, for instance animals, products, vehicles or document forms. We also assumed that the amount of annotated data available per class for the learning stage is limited.

8.2.2. ANR Project *Secular*

Participants: Laurent Amsaleg, Teddy Furon, Hervé Jégou, Ewa Kijak.

Duration: 3 years, started in September 2012

Partners: Morpho, Univ. Caen GREYC, Telecom ParisTech

Content-based retrieval systems (CBRS) are becoming the main multimedia security technology to enforce copyright laws or to spot illegal contents over the Internet. However, CBRS were not designed with privacy, confidentiality and security in mind. This comes in serious conflict with their use in these new security-oriented applications. Privacy is endangered due to information leaks when correlating users, queries and the contents stored-in-the-clear in the database. This is especially the case of images containing faces which are so popular in social networks. Biometrics systems have long relied on protection techniques and anonymization processes that have never been used in the context of CBRS. The project seeks to a better understanding of how biometrics related techniques can help increasing the security levels of CBRS while not degrading their performance.

8.3. European Initiatives

8.3.1. FP7 & H2020 Projects

8.3.1.1. VIAMASS

Type: FP7

Instrument: ERC Starting Grant

Duration: 04/2014 – 03/2019

Coordinator: Hervé Jégou

Inria contact: Hervé Jégou

Abstract: VIAMASS is a ERC Starting grant project coordinated by Hervé Jégou and with Teddy Furon as co-investigator. The goal of the project is to automatically discover visual links within a very large collection of images. These “visual hyper-links” will connect the objects across the images of the collection. This raises a major obstacle with respect to scalability: cross matching all the images is of quadratic complexity when performed with a brute-force approach. To this end, VIAMASS addresses issues at the frontier of the current state of the art in computer vision and signal processing.

8.3.1.2. Forensic Image Identifier and Analyzer

Program: Eurostars

Duration: 03/2011 – 07/2014

Coordinator: Videntifier Technologies

Other partners: Videntifier Technologies (Iceland), Forensic Pathways (UK)

Abstract: FIIA is an innovative software service for the Forensic market that automatically identifies and analyzes the content of images on web sites and seized computers. The service saves time and money, gathers better evidence, and builds stronger court cases. We are in charge of helping with the technology needed to identify the logos from terrorist organizations that are inserted in images or videos. Challenges are related to the poor resolution and small size of logos as well as to the very strict efficiency constraints that the logo detector must match.

8.4. International Initiatives

8.4.1. Inria Associate Teams

8.4.1.1. MOTIF

Title: Unsupervised motif discovery in multimedia content

International Partner (Institution - Laboratory - Researcher):

Pontifícia Universidade Católica de Minas Gerais - VIPLAB - Silvio Jamil Guimãraes

Universidade Federal Minas Gerais - NPDI - Arnaldo Albuquerque de Araújo

Duration: 2014 - 2017

See also: <http://www.icei.pucminas.br/projetos/viplab/projects/associate-team/index.html>

MOTIF aims at studying various approaches to unsupervised motif discovery in multimedia sequences, i.e., to the discovery of repeated sequences with no prior knowledge on the sequences. On the one hand, we will develop symbolic approaches inspired from work on bioinformatics to motif discovery in the multimedia context, investigating symbolic representations of multimedia data and adaptation of existing symbolic motif discovery algorithms. On the other hand, we will further develop cross modal clustering approaches to repeated sequence discovery in video data, building upon previous work.

8.4.2. Inria International Partners

8.4.2.1. Informal International Partners

- National Institute for Informatics, Japan
- Berkeley University, USA
- University of Amsterdam, The Netherlands
- Katholieke Universiteit Leuven, Belgium
- National Technical University of Athens, Greece
- Czech Technical University, Czech Republic

8.4.3. Other International Programs

- PICS CNRS MM-Analytics
 - Title: Fouille, visualisation et exploration multidimensionnelle de contenus multimédia ; Multi-Dimensional Multimedia Browsing, Mining, Analytics (num 6382).
 - International Partner (Institution - Laboratory - Researcher):
Reykjavík University, Iceland - Björn Þór Jónsson
 - Jan. 2014 – Dec. 2016
- STIC AmSud MAXIMUM Unsupervised Multimedia Content Mining
 - International coordinator: Guillaume Gravier, CNRS – IRISA, France
 - Scientific coordinators : Arnaldo de Albuquerque Araújo (Universidade Federal de Minas Gerais, Computer Science Department, NPDI); Benjamin Bustos (Universidad de Chile, Department of Computer Science, PRISMA); Silvio Jamil F. Guimarães (Pontifícia Universidade Católica de Minas Gerais, VIPLAB)
 - Jan. 2014 - Dec. 2015

8.5. International Research Visitors

8.5.1. Visits of International Scientists

Yannis Avrithis

Dates: October 2014 (2 weeks)

Institution: National Technical University of Athens (Greece)

8.5.2. Internships

Miaojing Shi

Dates: February 2014–January 2015 (1 year)

Subject: Large scale visual search

Institution: Pekin University (China)

8.5.3. Visits to International Teams

8.5.3.1. Research stays abroad

Petra Bosilj

Date: Sep. – Nov., 2014

Institution: Johann Bernoulli institute, Groningen, The Netherlands

Anca-Roxana Simon

Date: Apr. – June, 2014

Institution: Katholieke Universiteit Leuven, Belgium

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific events organisation

9.1.1.1. General chair, Scientific chair

Vincent Claveau was co-chair of the demonstration track at the Conf. Extraction et Gestion de Connaissances.

Guillaume Gravier is president of the steering committee of the ISCA/IEEE Workshop on Speech, Language and Audio in Multimedia series.

Patrick Gros is member of the steering board of the Content-Based Multimedia Indexing workshop series.

Pascale Sébillot is a member of the permanent steering committee of Conf. francophone en Traitement Automatique des Langues Naturelles.

Pascale Sébillot was a member of the steering committee of Conf. on Recherche d'Information et Applications.

9.1.1.2. Member of the organizing committee

Laurent Amsaleg and Li Weng organized a two-morning seminar on large-scale indexing at Irisa / Inria Rennes.

Andrei Bursuc and Hervé Jégou organized a micro-workshop in computer vision at Irisa / Inria Rennes.

Vincent Claveau was member of the organizing committee of the Conf. Extraction et Gestion de Connaissances.

Vincent Claveau, as ARIA financial head, participated in the financial organization of the Conférence en Recherche d'Information et Applications.

Vincent Claveau, as ARIA financial head, participated in the financial organization of the École d'été en Recherche d'Information et Applications.

Vincent Claveau was co-organizer and co-chair of a round table held during the Conf. francophone en Traitement automatique des langues naturelles.

9.1.2. Scientific events selection

9.1.2.1. Responsibility in conference program committee

Vincent Claveau was area chair at the Conf. francophone en Traitement Automatique des Langues Naturelles.

Guillaume Gravier was president of the program committee for Journées d'Études sur la Parole.

Guillaume Gravier was president of the program committee for ISCA/IEEE Workshop on Speech, Language and Multimedia.

Hervé Jégou was area chair for European Conf. on Computer Vision.

9.1.2.2. Membership of conference program committee

Laurent Amsaleg was a PC member of: Intl. Workshop on Content-Based Multimedia Indexing; IEEE Intl. Conf. on Multimedia and Expo; Intl. Conf. on MultiMedia Modeling; Intl. Conf. on Signal Image Technology and Internet Systems.

Vincent Claveau was a PC member of: Conférence en Recherche d'Information et Applications; IEEE/WIC/ACM Intl. Conf. on Web Intelligence.

Teddy Furon was a PC member of: IEEE Workshop on Information Forensics and Security.

Guillaume Gravier was a PC member of: ACM Intl. Conf. on Multimedia; Intl. Workshop on Content-Based Multimedia Indexing; Intl. Congress on Image and Signal Processing; Intl. Conf. on Computational Linguistics, Demo track; IEEE Intl. Conf. on Multimedia and Expo; ACM Intl. Conf. on Multimedia Information Retrieval; Annual Conf. of the Intl. Speech Communication Association; IEEE Intl. Workshop on Multimedia Signal Processing; Intl. Conf. on Language Resources and Evaluation.

Patrick Gros was a PC member of: Intl. Conf. on Creative Content Technologies; Conférence de Recherche d'information et Applications; European Signal Processing Conf.; Internation Symposium on Multimedia; Conf. on Machine Learning and Data Mining; Reconnaissance des Formes et Intelligence Artificielle.

Hervé Jégou was a PC member of: Computer Vision and Pattern Recognition; British Machine Vision Conf.; ACM Intl. Conf. on Multimedia Information Retrieval.

Ewa Kijak was a PC member of: Intl. Workshop on Content-Based Multimedia Indexing.

Christian Raymond was a PC member of: Annual Conf. of the Intl. Speech Communication Association; Intl. Conf. on Machine Learning and Applications.

Pascale Sébillot was a PC member of: Workshop on Distributional Semantics; Colloque Intl. Francophone sur l'Écrit et le Document; Congrès National de Reconnaissance des Formes et Intelligence Artificielle; Intl. Conf. on Language Resources and Evaluation.

9.1.2.3. Reviewer

Laurent Amsaleg reviewed projects for the Agence National de la Recherche.

Guillaume Gravier reviewed projects for the Agence National de la Recherche.

Christian Raymond reviewed for Conférence Traitement Automatique du Langage Naturel, Journées d'Étude sur la Parole,

9.1.3. Journal

9.1.3.1. Member of the editorial board

Vincent Claveau is member of the editorial committee of: Journal Traitement Automatique des Langues; Special issue of the Journal I3.

Teddy Furon is associate editor of IEEE Transactions on Information Forensics and Security

Hervé Jégou is associate editor of Computer Vision and Image Understanding

Christian Raymond is member of the editorial board of Revue Discours.

Pascale Sébillot is: editor of the Journal Traitement Automatique des Langues; member of the editorial committee of the Journal Traitement Automatique des Langues.

9.1.3.2. Reviewer

Laurent Amsaleg reviewed for Multimedia Tools and Applications, Mathematical Problems in Engineering.

Vincent Claveau was member of the reading committee for Multimedia Tools and Applications, for several issues of the Journal Traitement Automatique des Langues, for a special issue in Journal Documents Numériques

Teddy Furon reviewed for IEEE Trans. on Audio Speech and Language processing, IEEE Trans. on Information Forensics and Security, Elsevier Designs, Codes, and Cryptography journal, EURASIP Journal on Information Security.

Guillaume Gravier reviewed for IEEE Trans. on Audio Speech and Language, IEEE Transactions on Circuits and Systems for Video Technology, Intl. Journal of Multimedia Information Retrieval, Journal of Computer Science and Technology, Multimedia Tools and Applications, Traitement Automatique des Langues.

Hervé Jégou reviewed for IEEE Trans. Pattern Analysis and Machine Intelligence, IEEE Trans. Robotics, Computer Vision and Image Understanding.

Ewa Kijak reviewed for IEEE Transactions on Image Processing, Intl. Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, Journal of Applied Remote Sensing.

Pascale Sébillot was member of the reading committee for several issues of the Journal Traitement Automatique des Langues.

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

Master: Laurent Amsaleg, High-dimensional indexing, 13h, M2R, University Rennes 1, France

Master: Vincent Claveau, Data-Based Knowledge Acquisition: Symbolic Methods, 20h, M1, INSA de Rennes, France

Master: Vincent Claveau, Text Mining, 36h, M2, Univ. Rennes 1, France

Master: Vincent Claveau, Machine Learning for symbolic and sequential data, 7h, M2, Univ. Rennes 1, France

Master: Vincent Claveau, Information Retrieval, 15h, M2, ENSSAT, France

Master: Vincent Claveau, Information Retrieval, 13h, M2, Univ. Rennes 1, France

Licence: Teddy Furon, Probabilities, 40h, L1, Agrocampus Rennes, France

Licence: Guillaume Gravier, Databases, 30h, L2, INSA Rennes, France

Licence: Guillaume Gravier, C Programming, 10h, L3, INSA Rennes, France

Master: Guillaume Gravier, Data analysis and probabilistic modeling, 30h, M2, University Rennes 1, France

Master: Patrick Gros, Workshop on Mathematics, 12h, M2, Univ. Rennes 1, France

Master: Ewa Kijak, Image analysis and classification, 30h, M1, ESIR, France

Master: Ewa Kijak, Image processing, 64h, M1, ESIR, France

Master: Ewa Kijak, Supervised learning, 15h, M2R, University Rennes 1, France

Master: Ewa Kijak, Statistical data mining, 17h, M2, University Rennes 1, France

Master: Ewa Kijak, Indexing and multimedia databases, 15h, M2, ENSSAT, France

Master: Ewa Kijak, Computer vision, 8h, M2, ESIR, France

Master: Ewa Kijak, Automatic video summarization, Project with Technicolor, 15h, M2, ESIR, France

Master: Pascale Sébillot, Advanced Databases and Modern Information Systems, 70h, M2, INSA de Rennes, France

Master: Pascale Sébillot, Data-Based Knowledge Acquisition: Symbolic Methods, 18h, M1, INSA de Rennes, France

Master: Pascale Sébillot, Logic Programming, 12h, M1, INSA de Rennes, France

9.2.2. Supervision

HdR: Laurent Amsaleg, A Database Perspective on Large Scale High-Dimensional Indexing, Université de Rennes 1, defended on November 2014

HdR: Hervé Jégou, On visual recognition and similarity search, Université de Rennes 1, defended on October, 2014

PhD: Mihir Jain, Enhanced image and video representation for visual recognition, Université de Rennes 1, defended on April, 2014, Hervé Jégou, Patrick Bouthemy and Patrick Gros

PhD in progress: Abir Ncibi, Structuring TV steams, started November 2011, Vincent Claveau and Patrick Gros

PhD in progress: Mohammed-Haykel Boukadida, Video summarization based on constraint programming, started January 2012, Patrick Gros

PhD in progress: Petra Bosilj, Content based image indexing and retrieval using hierarchical image representations, started October 2012, Ewa Kijak and Sebastien Lefèvre (with OBELIX, IRISA team)

PhD in progress: Bingqing Qu, Structure discovery in collections of recurrent TV programs, started Oct. 2012, Guillaume Gravier

PhD in progress: Anca Roxana Simon, Hierarchical semantic structuring of video collections, started October 2012, Guillaume Gravier and Pascale Sébillot

PhD in progress: Raghavendran Balu, Privacy-preserving data aggregation and service personalization using highly-scalable data indexing techniques, started Oct. 2013, Teddy Furon and Laurent Amsaleg

PhD in progress: Raheel Kareem Qader, Phonology modeling for emotional speech synthesis, started January 2014, Gwénoél Lecorvé and Pascale Sébillot (with EXPRESSION, IRISA Team)

PhD in progress: Rémi Bois, Navigable directed multimedia hypergraphs: construction and exploitation, started October 2014, Guillaume Gravier and Pascale Sébillot

PhD in progress: Ahmet Iscen, Continuous memories for representing sets of vectors and image collections, started September 2014, Hervé Jégou and Teddy Furon

PhD in progress: Grégoire Jadi, Opinion mining in multimedia data, started October 2014, Vincent Claveau, Béatrice Daille (LINA, Nantes) and Laura Monceaux (LINA, Nantes)

PhD in progress: Vedran Vukotič, Deep neural architectures for automatic representation learning from multimedia multimodal data, started October 2014, Guillaume Gravier and Christian Raymond

9.2.3. Juries

Vincent Claveau

PhD, K. Abdulahhad, LIG - Univ. Joseph Fourier, Grenoble

PhD, M. Dupuch, INSERM, Paris

PhD, R. Harastani, LINA - Univ. de Nantes

Teddy Furon

HDR, Patrick Bas, Université de Lille

PhD, reviewer, Antonino Simone, University of Eindhoven

PhD, Farzad Farhadzadeh, Université de Genève

Guillaume Gravier

PhD, Alberto Pimentel, Universidade Federal de Minas Gerais

PhD, Nadia Derbas, reviewer, Université de Grenoble

Patrick Gros

HDR, reviewer, Sid-Ahmed Berrani, Univ. Rennes 1

PhD, Mohammed Derbel, Univ of Tours

Pascale Sébillot

HDR, reviewer and president, Xavier Tannier, Paris-Sud university

PhD, president, Guiyao Ke, South-Brittany university

PhD, president, Mohamed Hatmi, Nantes university

9.3. Community animation

Vincent Claveau is treasurer of the Association pour la Recherche d'Information et Applications.

Vincent Claveau is deputy head of the GDR MADICS.

Guillaume Gravier is president of the Association Francophone de la Communication Parlée.

Guillaume Gravier is co-founder and general chair of the ISCA SIG Speech, Language and Audio in Multimedia.

Guillaume Gravier is member of the Community Council of the Mediaeval Multimedia Evaluation series.

Pascale Sébillot is a member of the Conseil National des Universités.

9.4. Popularization

Laurent Amsaleg. Lettre d'information Émergence, issue 32, 11/07/2014, "Un outil pour fouiller plus vite les disques durs saisis".

Laurent Amsaleg. Sciences Ouest, issue 324, October 2014, "Ils repèrent les contenus illicites".

Vincent Claveau co-organized and co-animated a round table held during the TALN conference about the evolution of NLP, from the points of view of education, research and industry.

Hervé Jégou was invited panelist at Intl. Conf. on Multimedia Retrieval, April 2014.

Hervé Jégou gave an invited tutorial on large-scale image search at Intl. Conf. on Multimedia Retrieval, April 2014.

Hervé Jégou gave a tutorial on efficient matching at Conf. on Computer Vision and Pattern Recognition, June 2014.

Hervé Jégou was invited speaker at Conf. on Computer Vision and Pattern Recognition, June 2014.

Pascale Sébillot: Invited speaker (3h) NLP tackling large volumes of potentially damaged textual data: what is different?, IST Inria seminar on "Big data: new information partitions", October 2014.

10. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

- [1] L. AMSALEG. *A Database Perspective on Large Scale High-Dimensional Indexing*, Université Rennes 1, November 2014, Habilitation à diriger des recherches, <https://hal.inria.fr/tel-01097210>

- [2] M. JAIN. *Enhanced image and video representation for visual recognition*, Université Rennes 1, April 2014, <https://tel.archives-ouvertes.fr/tel-00996793>

Articles in International Peer-Reviewed Journals

- [3] C. BAI, J. ZHANG, Z. LIU, W.-L. ZHAO. *K-means based histogram using multiresolution feature vectors for color texture database retrieval*, in "Multimedia Tools and Applications", May 2014 [DOI : 10.1007/s11042-014-2053-8], <https://hal.inria.fr/hal-00997873>
- [4] A. BOURRIER, F. PERRONNIN, R. GRIBONVAL, P. PÉREZ, H. JÉGOU. *Explicit embeddings for nearest neighbor search with Mercer kernels*, in "Journal of Mathematical Imaging and Vision", January 2015, pp. 1-10 [DOI : 10.1007/s10851-015-0555-2], <https://hal.inria.fr/hal-00722635>
- [5] M. BÉCUE-BERTAUT, B. A. KOSTOV, A. MORIN, G. NARO. *Rhetorical Strategy in Forensic Speeches: Multidimensional Statistics-Based Methodology*, in "Journal of Classification", April 2014, vol. 31, n^o 1, pp. 85-106 [DOI : 10.1007/s00357-014-9148-9], <https://hal.inria.fr/hal-01026150>
- [6] V. CLAVEAU, S. LEFÈVRE. *Topic segmentation of TV-streams by watershed transform and vectorization*, in "Computer Speech and Language", 2014, forthcoming, <https://hal.archives-ouvertes.fr/hal-00998259>
- [7] J. D. CRUZ GOMEZ, C. BOTHOREL, F. POULET. *Analyse intégrée des réseaux sociaux pour la détection et la visualisation de communautés*, in "Technique et Science Informatiques (TSI)", 2014, vol. 33, n^o 4, pp. 399-427, <https://hal.archives-ouvertes.fr/hal-00937849>
- [8] C.-H. DEMARTY, C. PENET, M. SOLEYMANI, G. GRAVIER. *VSD, a public dataset for the detection of violent scenes in movies: design, annotation, analysis and evaluation*, in "Multimedia Tools and Applications", May 2014 [DOI : 10.1007/s11042-014-1984-4], <https://hal.inria.fr/hal-01026152>
- [9] T.-N. DOAN, T.-N. DO, F. POULET. *Classification d'images à grande échelle avec des SVMs*, in "Traitement du Signal", 2014, vol. 31, n^o 1-2, pp. 39-56, <https://hal.archives-ouvertes.fr/hal-00956405>
- [10] T.-N. DOAN, T.-N. DO, F. POULET. *Large Scale Classifiers for Visual Classification Tasks*, in "Multimedia Tools and Applications", 2014, 28 p. , <https://hal.archives-ouvertes.fr/hal-00952765>
- [11] T.-N. DOAN, T.-N. DO, F. POULET. *Parallel Incremental Power Mean SVM for Classification of Large Scale Visual Datasets*, in "International Journal of Multimedia Information Retrieval", 2014, vol. 3, n^o 2, pp. 89-96, <https://hal.archives-ouvertes.fr/hal-00943187>
- [12] K. ELAGOUNI, C. GARCIA, F. MAMALET, P. SÉBILLOT. *Text Recognition in Multimedia Documents: A Study of two Neural-based OCRs Using and Avoiding Character Segmentation*, in "International Journal on Document Analysis and Recognition, IJDAR", March 2014, vol. 17, n^o 1, pp. 19-31 [DOI : 10.1007/s10032-013-0202-7], <https://hal.archives-ouvertes.fr/hal-00867225>
- [13] P.-H. GOSSELIN, N. MURRAY, H. JÉGOU, F. PERRONNIN. *Revisiting the Fisher vector for fine-grained classification*, in "Pattern Recognition Letters", November 2014, vol. 49, pp. 92-98, <https://hal.archives-ouvertes.fr/hal-01056223>

- [14] O. KIHIL, D. PICARD, P.-H. GOSSELIN. *Local polynomial space–time descriptors for action classification*, in "Machine Vision and Applications", 2014, pp. 1-11 [DOI : 10.1007/s00138-014-0652-z], <https://hal.archives-ouvertes.fr/hal-01097536>
- [15] O. KIHIL, D. PICARD, P.-H. GOSSELIN. *A Unified framework for local visual descriptors evaluation*, in "Pattern Recognition", April 2015, vol. 48, pp. 1170-1180 [DOI : 10.1016/J.PATCOG.2014.11.013], <https://hal.archives-ouvertes.fr/hal-01089310>
- [16] V. F. MOTA, E. D. A. PEREZ, S. MACIEL LUIZ MAURÍLIO DA, M. B. VIEIRA, P.-H. GOSSELIN. *A tensor motion descriptor based on histograms of gradients and optical flow*, in "Pattern Recognition Letters", April 2014, vol. 39, pp. 85-91 [DOI : 10.1016/J.PATREC.2013.08.008], <https://hal.archives-ouvertes.fr/hal-00861395>
- [17] C. PENET, C.-H. DEMARTY, G. GRAVIER, P. GROS. *Variability Modelling for Audio Events Detection in Movies*, in "Multimedia Tools and Applications", July 2014 [DOI : 10.1007/s11042-014-2038-7], <https://hal.inria.fr/hal-01026151>
- [18] G. TOLIAS, H. JÉGOU. *Visual query expansion with or without geometry: refining local descriptors by feature aggregation*, in "Pattern Recognition", April 2014, <https://hal.inria.fr/hal-00971267>
- [19] L. WENG, L. AMSALEG, A. MORTON, S. MARCHAND-MAILLET. *A Privacy-Preserving Framework for Large-Scale Content-Based Information Retrieval*, in "Information Forensics and Security, IEEE Transactions on", 2014, 16 p. [DOI : 10.1109/TIFS.2014.2365998], <https://hal.inria.fr/hal-01083328>

International Conferences with Proceedings

- [20] B. ARNULPHY, V. CLAVEAU, X. TANNIER, A. VILNAT. *Techniques d'apprentissage supervisé pour l'extraction d'événements TimeML en anglais et français*, in "CORIA - Conférence en recherche d'information et applications", Nancy, France, March 2014, 16 p. , <https://hal.archives-ouvertes.fr/hal-01027563>

[21] *Best Paper*

R. BALU, T. FURON, S. GAMBS. *Challenging differential privacy: the case of non-interactive mechanisms*, in "European Symposium on Research in Computer Security", Wroclaw, Poland, LNCS, Springer-Verlag, September 2014, vol. 8657, Best Student Paper Award [DOI : 10.1007/978-3-319-11212-1_9], <https://hal.inria.fr/hal-01011346>.

- [22] R. BALU, T. FURON, H. JÉGOU. *Beyond "project and sign" for cosine estimation with binary codes*, in "ICASPP - International Conference on Acoustics, Speech, and Signal Processing", Florence, Italy, IEEE, May 2014, <https://hal.inria.fr/hal-00942075>

- [23] B. BHATTARAI, A. MIGNON, F. JURIE, T. FURON. *Puzzling Face Verification Algorithms for Privacy Protection*, in "IEEE Workshop on Information Forensics and Security", Atlanta, United States, Y. L. SUN, V. H. ZHAO (editors), IEEE, December 2014, <https://hal.inria.fr/hal-01066070>

- [24] M. H. BOUKADIDA, S.-A. BERRANI, P. GROS. *A Novel Modeling for Video Summarization Using Constraint Satisfaction Programming*, in "International Symposium On Visual Computing", Las Vegas, United States, G. BEBIS, R. BOYLE, B. PARVIN, D. KORACIN, R. MCMAHAN, J. JERALD, H. ZHANG, S. DRUCKER, C. KAMBHAMETTU, M. EL CHOUBASSI, Z. DENG, M. CARLSON (editors), Advances in Visual Computing,

- Springer International Publishing, December 2014, vol. 8888 [DOI : 10.1007/978-3-319-14364-4_20], <https://hal.inria.fr/hal-01100037>
- [25] V. CLAVEAU, T. HAMON, N. GRABAR, S. LE MAGUER. *RePaLi participation to CLEF eHealth IR challenge 2014: leveraging term variation*, in "Conference and Labs of the Evaluation Forum CLEF", Sheffield, United Kingdom, September 2014, 13 p. , <https://hal.archives-ouvertes.fr/hal-01027534>
- [26] V. CLAVEAU, E. KIJAK. *Generating and using probabilistic morphological resources for the biomedical domain*, in "LREC - 9th edition of the Language Resources and Evaluation Conference", Reykjavik, Iceland, May 2014, 7 p. , <https://hal.archives-ouvertes.fr/hal-01027778>
- [27] V. CLAVEAU, E. KIJAK, O. FERRET. *Improving distributional thesauri by exploring the graph of neighbors*, in "COLING - International Conference on Computational Linguistics", Dublin, Ireland, August 2014, 12 p. , <https://hal.archives-ouvertes.fr/hal-01027545>
- [28] V. CLAVEAU, A. NCIBI. *Knowledge discovery with CRF-based clustering of named entities without a priori classes*, in "Conference on Intelligent Text Processing and Computational Linguistics CICLing", Kathmandu, Nepal, A. GELBUKH (editor), LNCS, Springer, April 2014, vol. 8403, n^o 1-2, pp. 415-428, <https://hal.archives-ouvertes.fr/hal-01027520>
- [29] A. DELVINIOTI, H. JÉGOU, L. AMSALEG, M. E. HOULE. *Image Retrieval with Reciprocal and shared Nearest Neighbors*, in "VISAPP–International Conference on Computer Vision Theory and Applications", Barcelone, Portugal, January 2014, <https://hal.inria.fr/hal-00907481>
- [30] M. DOUZE, H. JÉGOU. *The Yael library*, in "ACM Multimedia", Orlando, United States, November 2014, <https://hal.inria.fr/hal-01020695>
- [31] J. FELLUS, D. PICARD, P.-H. GOSSELIN. *Dimensionality reduction in decentralized networks by Gossip aggregation of principal components analyzers*, in "European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning", Bruges, Belgium, April 2014, pp. 171-176, <https://hal.archives-ouvertes.fr/hal-00985721>
- [32] T. FURON, M. DESOUBEAUX. *Tardos codes for real*, in "IEEE Workshop on Information Forensics and Security", Atlanta, United States, IEEE, December 2014, <https://hal.inria.fr/hal-01066043>
- [33] O. GALIBERT, J. LEIXA, A. GILLES, K. CHOUKRI, G. GRAVIER. *The ETAPE Speech Processing Evaluation*, in "Conference on Language Resources and Evaluation", Reykyavik, Iceland, May 2014, <https://hal.archives-ouvertes.fr/hal-01083636>
- [34] G. GRAVIER, N. SOUVIRAA-LABASTIE, S. CAMPION, F. BIMBOT. *Audio thumbnails for spoken content without transcription based on a maximum motif coverage criterion*, in "Annual Conference of the International Speech Communication Association", Singapour, Singapore, September 2014, <https://hal.archives-ouvertes.fr/hal-01026402>
- [35] M. JAIN, J. VAN GEMERT, H. JÉGOU, P. BOUTHEMY, C. SNOEK. *Action localization with tubelets from motion*, in "CVPR - International Conference on Computer Vision and Pattern Recognition", Columbus, United States, June 2014, <https://hal.inria.fr/hal-00996844>

- [36] H. JÉGOU, A. ZISSERMAN. *Triangulation embedding and democratic aggregation for image search*, in "CVPR - International Conference on Computer Vision and Pattern Recognition", Columbus, United States, June 2014, <https://hal.inria.fr/hal-00977321>
- [37] B. Þ. JÓNSSON, Á. ERÍKSDÓTTIR, Ó. WAAGE, G. TÓMASSON, H. SIGURÞÓRSSON, L. AMSALEG. *$M^3 + P^3 + O^3 = Multi-D Photo Browsing$* , in "International Conference on MultiMedia Modeling", Dublin, Ireland, January 2014, <https://hal.inria.fr/hal-00925464>
- [38] B. T. JÓNSSON, G. TÓMASSON, H. SIGURÞÓRSSON, Á. ERÍKSDÓTTIR, L. AMSALEG, M. K. LARUS-DOTTIR. *A Multi-Dimensional Data Model for Personal Photo Browsing*, in "International Conference on Multimedia Modelling", Sydney, Australia, 2015, <https://hal.inria.fr/hal-01083344>
- [39] J. KRAPAC, F. PERRONNIN, T. FURON, H. JÉGOU. *Instance classification with prototype selection*, in "ICMR - ACM International Conference on Multimedia Retrieval", Glasgow, United Kingdom, April 2014, <https://hal.inria.fr/hal-00942275>
- [40] A. LAURENT, N. CAMELIN, C. RAYMOND. *Boosting bonsai trees for efficient features combination : application to speaker role identification*, in "Interspeech", Singapur, Singapore, September 2014, <https://hal.inria.fr/hal-01025171>
- [41] A. LAURENT, N. CAMELIN, C. RAYMOND. *Boosting de bonzaïs pour la combinaison efficace de descripteurs : application à l'identification du rôle du locuteur*, in "Journée d'Etude sur la Parole", Le Mans, France, June 2014, <https://hal.inria.fr/hal-01025142>
- [42] B. LUDUSAN, G. GRAVIER, E. DUPOUX. *Incorporating Prosodic Boundaries in Unsupervised Term Discovery*, in "International conference on Speech Prosody", Dublin, Ireland, May 2014, pp. 207-211, <https://hal.archives-ouvertes.fr/hal-01026421>
- [43] B. LUDUSAN, M. VERSTEEGH, A. JANSEN, G. GRAVIER, X.-N. CAO, M. JOHNSON, E. DUPOUX. *Bridging the gap between speech technology and natural language processing: an evaluation toolbox for term discovery systems*, in "Language Resources and Evaluation Conference", Reykyavik, Iceland, May 2014, <https://hal.inria.fr/hal-01026368>
- [44] B. LUDUSAN, S. ZIEGLER, G. GRAVIER. *Is Syllable Stress Information Robust for ASR in Adverse Conditions?*, in "International Conference on Speech Prosody", Dublin, Ireland, May 2014, pp. 939-943, <https://hal.archives-ouvertes.fr/hal-01026423>
- [45] R. NEGREL, D. PICARD, P.-H. GOSSELIN. *Efficient Dimension Reduction Of Global Signature With Sparse Projectors For Image Near Duplicate Retrieval*, in "IAPR International Conference on Pattern Recognition", Stockholm, Sweden, August 2014, 6 p. , <https://hal.archives-ouvertes.fr/hal-01064050>
- [46] R. NEGREL, D. PICARD, P.-H. GOSSELIN. *Evaluation of Second-order Visual Features for Land-Use Classification*, in "12th International Workshop on Content-Based Multimedia Indexing", France, June 2014, 5 p. , <https://hal.archives-ouvertes.fr/hal-01022971>
- [47] R. NEGREL, D. PICARD, P.-H. GOSSELIN. *Dimensionality reduction of visual features using sparse projectors for content-based image retrieval*, in "IEEE International Conference on Image Processing", Paris, France, October 2014, pp. 2192-2196, <https://hal.archives-ouvertes.fr/hal-01081770>

- [48] B. QU, F. VALLET, J. CARRIVE, G. GRAVIER. *Content-based inference of hierarchical structural grammar for recurrent TV programs using multiple sequence alignment*, in "IEEE International Conference on Multimedia and Expo", Chengdu, China, July 2014, <https://hal.archives-ouvertes.fr/hal-01026335>
- [49] B. QU, F. VALLET, J. CARRIVE, G. GRAVIER. *Using grammar induction to discover the structure of recurrent TV programs*, in "International Conferences on Advances in Multimedia", Nice, France, February 2014, pp. 112-117, <https://hal.archives-ouvertes.fr/hal-01026331>
- [50] B. QU, F. VALLET, J. CARRIVE, G. GRAVIER. *Content-based discovery of multiple structures from episodes of recurrent TV programs based on grammatical inference*, in "International Conference on Multimedia Modelling", Sydney, Australia, January 2015, <https://hal.archives-ouvertes.fr/hal-01089237>
- [51] Y. RICQUEBOURG, C. RAYMOND, B. POIRRIEZ, A. LEMAITRE, B. COÛASNON. *Boosting bonsai trees for handwritten/printed text discrimination*, in "Document Recognition and Retrieval (DRR)", San Francisco, United States, February 2014, <https://hal.archives-ouvertes.fr/hal-00910718>
- [52] M. SHI, T. FURON, H. JÉGOU. *A Group Testing Framework for Similarity Search in High-dimensional Spaces*, in "ACM Multimedia", Orlando, United States, November 2014, <https://hal.inria.fr/hal-01062531>
- [53] A.-R. SIMON, G. GRAVIER, P. SÉBILLOT, M.-F. MOENS. *IRISA and KUL at MediaEval 2014: Search and Hyperlinking Task*, in "MediaEval 2014 Workshop", Barcelona, Spain, Working Notes Proceedings of the MediaEval 2014 Workshop, October 2014, <https://hal.archives-ouvertes.fr/hal-01094850>
- [54] A.-R. SIMON, C. GUINAUDEAU, P. SÉBILLOT, G. GRAVIER. *Investigating domain-independent NLP techniques for precise target selection in video hyperlinking*, in "ISCA/IEEE Workshop on Speech, Language and Audio in Multimedia", Penang, Malaysia, September 2014, <https://hal.archives-ouvertes.fr/hal-01053698>
- [55] H. TABIA, H. LAGA, D. PICARD, P.-H. GOSSELIN. *Covariance Descriptors for 3D Shape Matching and Retrieval*, in "IEEE Conference on Computer Vision and Pattern Recognition", Columbus, Ohio, United States, June 2014, 8 p. , <https://hal.archives-ouvertes.fr/hal-01022970>
- [56] N. THÉLY, F. MOREAU, V. CLAVEAU, E. TOLONE. *La critique d'art au banc d'essai des humanités numériques*, in "DI - Digital intelligence", Nantes, France, September 2014, 5 p. , <https://hal.archives-ouvertes.fr/hal-01052806>
- [57] G. TOLIAS, T. FURON, H. JÉGOU. *Orientation covariant aggregation of local descriptors with embeddings*, in "European Conference on Computer Vision", Zurich, Switzerland, September 2014, <https://hal.inria.fr/hal-01020823>

National Conferences with Proceedings

- [58] V. CLAVEAU. *Agrégation de sac-de-sacs-de-mots pour la recherche d'information par modèles vectoriels*, in "EGC - 14 ème conférence Extraction et Gestion des Connaissances", Rennes, France, January 2014, 6 p. , <https://hal.archives-ouvertes.fr/hal-01027719>
- [59] V. CLAVEAU, P. GROS. *Clustering de données relationnelles pour la structuration de flux télévisuels*, in "EGC - 14 ème conférence Extraction et Gestion des Connaissances", Rennes, France, January 2014, 6 p. , <https://hal.archives-ouvertes.fr/hal-01027763>

- [60] V. CLAVEAU, E. KIJAK, O. FERRET. *Explorer le graphe de voisinage pour améliorer les thésaurus distributionnels*, in "TALN - 21ème conférence sur le Traitement Automatique des Langues Naturelles", Marseille, France, July 2014, 12 p. , <https://hal.archives-ouvertes.fr/hal-01027787>
- [61] S. LE MAGUER, E. DELAIS-ROUSSARIE, N. BARBOT, M. AVANZI, O. ROSEC, D. LOLIVE. *Algorithme de découpage en groupes prosodiques pour la dictée par l'usage de synthèse vocale*, in "JEP - 30ème édition des Journées d'Etudes sur la Parole", Le Mans, France, June 2014, <https://hal.inria.fr/hal-00977456>
- [62] H. YOO, S. LE MAGUER, E. DELAIS-ROUSSARIE, N. BARBOT, D. LOLIVE. *Evaluation d'un algorithme de chunking appliqué à la dictée*, in "JEP - 30ème édition des Journées d'Etudes sur la Parole", Le Mans, France, June 2014, <https://hal.inria.fr/hal-00977445>

Scientific Books (or Scientific Book chapters)

- [63] *Second IEEE/ISCA Workshop on Speech, Language and Audio in Multimedia*, Proceedings of the Second IEEE/ISCA Workshop on Speech, Language and Audio in Multimedia, September 2014, <https://hal.archives-ouvertes.fr/hal-01083630>
- [64] C.-H. DEMARTY, C. PENET, B. IONESCU, G. GRAVIER, M. SOLEYMANI. *Multimodal Violence Detection in Hollywood Movies: State-of-the-Art and Benchmarking*, in "Fusion in Computer Vision", B. IONESCU, J. BENOIS-PINEAU, T. PIATRIK, G. QUÉNOT (editors), Advances in Computer Vision and Pattern Recognition, Springer International Publishing, 2014, pp. 185-208 [DOI : 10.1007/978-3-319-05696-8_8], <https://hal.inria.fr/hal-00968515>
- [65] T.-N. DOAN, F. POULET. *Large Scale Support Vector Machines Algorithms for Visual Recognition*, Scholars' Press, May 2014, 164 p. , <https://hal.archives-ouvertes.fr/hal-01064443>
- [66] T.-N. DOAN, F. POULET. *Large Scale Image Classification: fast Feature Extraction, Multi-CodeBook Approach and SVM Training*, in "AKDM-4, Advances in Knowledge Discovery and Management, Vol.4", Studies in Computational Intelligence, Springer-Verlag, 2014, pp. 159-176, <https://hal.archives-ouvertes.fr/hal-00907533>
- [67] T. GAILLAT, P. SÉBILLOT, N. BALLIER. *Automated classification of unexpected uses of this and that in a learner corpus of English*, in "Recent Advances in Corpus Linguistics - Developing and Exploiting Corpora", L. VANDELANOTTE, K. DAVIDSE, C. GENTENS, D. KIMPS (editors), Language and Computers 78, Rodopi, 2014, pp. 309-324, <https://hal.archives-ouvertes.fr/hal-01058760>
- [68] P. SÉBILLOT. *Le traitement automatique des langues face aux données textuelles volumineuses et potentiellement dégradées : qu'est-ce que cela change ?*, in "Big data : nouvelles partitions de l'information. Actes du colloque de l'Inria, octobre 2014", L. CALDERAN, J. MILLET (editors), Information et stratégie, De Boeck, January 2015, pp. 35-49, <https://hal.archives-ouvertes.fr/hal-01056396>

Research Reports

- [69] R. QADER, G. LECORVÉ, D. LOLIVE, P. SÉBILLOT. *Phonology Modelling for Expressive Speech Synthesis: a Review*, July 2014, n^o PI-2020, <https://hal.inria.fr/hal-01021911>
- [70] L. WENG, L. AMSALEG, S. MARCHAND-MAILLET, A. MORTON. *A Privacy-Preserving Framework for Large-Scale Content-Based Information Retrieval*, February 2014, n^o 14.01, <https://hal.inria.fr/hal-01059560>

References in notes

- [71] M. AGOSTI, J. ALLAN. *Special issue on methods and tools for the automatic construction of hypertext*, in "Information Processing and Retrieval", 1997, vol. 33, n^o 2
- [72] C. A. BHATT, M. S. KANKANHALLI. *Multimedia data mining: State of the art and challenges*, in "Multimedia Tools and Applications", 2011, vol. 51, n^o 1, pp. 35–76
- [73] O. CHUM, J. MATAS. *Large-Scale Discovery of Spatially Related Images*, in "Pattern Analysis and Machine Intelligence", 2010, vol. 32, n^o 2, pp. 371-377
- [74] M. ESKEVICH, G. J. JONES, S. CHEN, R. ALY, R. ORDELMAN, D. NADEEM, C. GUINAUDEAU, G. GRAVIER, P. SÉBILLOT, T. D. NIES, P. DEBEVERE, R. V. DE WALLE, P. GALUŠCKOVÁ, P. PECINA, M. LARSON. *Multimedia Information Seeking through Search And Hyperlinking*, in "Intl. Conf. on Multimedia Retrieval", 2013
- [75] C. HERLEY. *ARGOS: Automatically extracting Repeating Objects from multimedia Streams*, in "IEEE Trans. on Multimedia", Feb. 2006, vol. 8, n^o 1, pp. 115–129
- [76] A. PARK, J. R. GLASS. *Unsupervised pattern discovery in speech*, in "IEEE Transaction on Acoustic, Speech and Language Processing", Jan. 2008, vol. 16, n^o 1, pp. 186–197
- [77] K. M. PUA, J. GAUCH, S. E. GUCH, J. Z. MIADOWICZ. *Real time repeated video sequence identification*, in "Computer Vision and Image Understanding", 2004, vol. 93, n^o 2004, pp. 310–327
- [78] C. G. M. SNOEK, M. WORRING. *Multimodal Video Indexing: A Review of the State-of-the-Art*, in "Multimedia Tools and Applications", 2005, vol. 25, n^o 1, pp. 5–35
- [79] O. DE ROOIJ, M. WORRING. *Browsing Video Along Multiple Threads*, in "IEEE Transactions on Multimedia", 2010, vol. 12, n^o 2, pp. 121–130