



IN PARTNERSHIP WITH:
CNRS

**Institut national des sciences
appliquées de Rennes**

Université Rennes 1

Activity Report 2015

Project-Team **LINKMEDIA**

Creating and exploiting explicit links between multimedia fragments

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. Description and structuring	4
3.2.3. Linking and collection data model	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. TermEx	5
5.2. Experimental platform	5
5.3. AllGO multimedia web services	5
6. New Results	6
6.1. Unsupervised motif and knowledge discovery	6
6.1.1. Estimation of continuous intrinsic dimension	6
6.1.2. Supervised multi-scale locality sensitive hashing	6
6.1.3. Rotation and translation covariant match kernels for image retrieval	6
6.1.4. Sequential pattern mining on audio data	7
6.1.5. Clustering by diverting supervised machine learning	7
6.1.6. Multimodal person discovery in TV broadcasts	7
6.1.7. Unsupervised video structure mining with grammatical inference	8
6.1.8. Information retrieval for distributional semantics, and vice-versa	8
6.2. Multimedia content description and structuring	8
6.2.1. Image description using component trees	8
6.2.2. Improved motion description for action classification	9
6.2.3. Word embeddings and recurrent neural networks for spoken language understanding	9
6.2.4. Hierarchical topic structuring	9
6.2.5. Partial least square hashing for large-scale face identification	9
6.2.6. Selection strategies for active learning in NLP	10
6.2.7. Tree-structured named entities extraction from competing speech transcripts	10
6.3. Content-based information retrieval	10
6.3.1. A comparison of dense region detectors for image search and fine-grained classification	10
6.3.2. Efficient large-scale similarity search using matrix factorization	10
6.3.3. Explicit embeddings for nearest neighbor search with Mercer kernels	11
6.3.4. Image search with selective match kernels: aggregation across single and multiple images	11
6.3.5. Early burst detection for memory-efficient image retrieval	11
6.3.6. Biomedical information retrieval	12
6.4. Linking, navigation and analytics	12
6.4.1. Sentiment analysis on social networks	12
6.4.2. A multi-dimensional data model for personal photo browsing	12
6.4.3. NLP-driven hyperlink construction in broadcast videos	13

6.4.4. Information extraction	13
6.5. Participation in benchmarking initiatives	13
7. Bilateral Contracts and Grants with Industry	14
8. Partnerships and Cooperations	14
8.1. Regional Initiatives	14
8.2. National Initiatives	14
8.2.1. ANR Project FIRE-ID	14
8.2.2. ANR Project Secular	14
8.2.3. ANR Project IDFRAud	15
8.2.4. FUI 19 NexGenTV	15
8.3. International Initiatives	15
8.3.1. Inria Associate Teams not involved in an Inria International Labs	15
8.3.2. Inria International Partners	16
8.3.3. Participation In other International Programs	16
8.4. International Research Visitors	16
8.4.1. Visits of International Scientists	16
8.4.2. Visits to International Teams	16
9. Dissemination	17
9.1. Promoting Scientific Activities	17
9.1.1. Scientific events organisation	17
9.1.2. Scientific events selection	17
9.1.2.1. Chair of conference program committees	17
9.1.2.2. Member of the conference program committees	17
9.1.2.3. Reviewer	18
9.1.3. Journal	18
9.1.3.1. Member of the editorial boards	18
9.1.3.2. Reviewer - Reviewing activities	18
9.1.4. Invited talks	18
9.1.5. Leadership within the scientific community	18
9.1.6. Scientific expertise	18
9.1.7. Research administration	19
9.2. Teaching - Supervision - Juries	19
9.2.1. Teaching	19
9.2.2. Supervision	20
9.2.3. Juries	20
9.3. Popularization	21
10. Bibliography	21

Project-Team LINKMEDIA

Creation of the Project-Team: 2014 July 01

Keywords:

Computer Science and Digital Science:

- 3.3.1. - On-line analytical processing
- 3.3.2. - Data mining
- 3.4.1. - Supervised learning
- 3.4.8. - Deep learning
- 5.3.3. - Pattern recognition
- 5.4.1. - Object recognition
- 5.4.3. - Content retrieval
- 5.8. - Natural language processing
- 8.2. - Machine learning
- 8.4. - Natural language processing

Other Research Topics and Application Domains:

- 9. - Society and Knowledge

1. Members

Research Scientists

Guillaume Gravier [Team leader, CNRS, Senior Researcher, HdR]
Laurent Amsaleg [CNRS, Researcher, HdR]
Vincent Claveau [CNRS, Researcher]
Teddy Furon [Inria, Researcher]
Hervé Jégou [Inria, Researcher, on leave since May 2015]

Faculty Members

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]

Engineers

Andrei Bursuc [Inria]
Ronan Sicre [Univ. Rennes I until Aug. 2016, then Inria]

PhD Students

Raghavendran Balu [Inria]
Rémi Bois [CNRS]
Petra Bosilj [Univ. Bretagne Sud]
Mohamed Haykel Boukadida [Orange Labs, until Jan 2015]
Ricardo Carlini Sperandio [UFMG, from Aug 2015]
Ahmet Iscen [Inria]
Grégoire Jadi [Univ. Nantes]
Cédric Maigrot [Univ. Rennes I, from Oct 2015]
Abir Ncibi [Inria, until Feb 2015]
Bingqing Qu [INA, until Sep 2015]

Anca-Roxana Şimon [Univ. Rennes I, until Nov 2015]

Vedran Vukotić [INSA Rennes]

Post-Doctoral Fellows

Gylfi Þor Guðmundsson [Inria]

Georgios Toliás [Inria, until May 2015]

Li Weng [Inria, until Aug 2015]

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:

- designing 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 [54], [52], [53].

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 dynamic time warping; 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 natural language processing (NLP) for spoken document.

3.2.2. Description 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 [55], [49]. 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: leveraging encryption and obfuscation; 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; conditional random fields.

3.2.3. Linking and collection data model

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 [48] and was recently extended to multimedia content [56], [51], [50]. 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. TermEx

Participant: Vincent Claveau [correspondent].

TermEx is a domain-independent terminology extraction system based on natural language processing and information retrieval concepts. This year, a new version (2.0) has been implemented that corresponds to a major rewriting in Python3 with support for English (in addition to French) and faster processing of documents in batch.

In 2015, TermEx has been licensed to a large company as a key component of the archiving process.

5.2. Experimental platform

Participant: Laurent Amsaleg [correspondent].

The experimental multimedia indexing platform (PIM) consists of dedicated equipments to experiment on very large collections of multimedia data. In 2015, no major evolution of PIM occurred and activities on the platform mainly consisted on maintenance. Due to the departure of Sébastien Campion, our former PIM manager, we have also initiated a reorganization of the responsibilities, in collaboration with SED.

5.3. AllGO multimedia web services

Participant: Guillaume Gravier [correspondent].

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. In 2015, the team hired a development engineer to revamp the web service offer, making services interoperable and broadening the scope of services made available.

6. New Results

6.1. Unsupervised motif and knowledge discovery

6.1.1. Estimation of continuous intrinsic dimension

Participants: Laurent Amsaleg, Teddy Furon.

In collaboration with Michael Houle, National Institute for Informatics (Japan).

Some of our research work was concerned with the estimation of continuous intrinsic dimension (ID), a measure of intrinsic dimensionality recently proposed by Houle. Continuous ID can be regarded as an extension of Karger and Ruhl's expansion dimension to a statistical setting in which the distribution of distances to a query point is modeled in terms of a continuous random variable. This form of intrinsic dimensionality can be particularly useful in search, classification, outlier detection, and other contexts in machine learning, databases, and data mining, as it has been shown to be equivalent to a measure of the discriminative power of similarity functions. In [11], we proposed several estimators of continuous ID that we analyzed based on extreme value theory, using maximum likelihood estimation, the method of moments, probability weighted moments, and regularly varying functions. Experimental evaluation was performed using both real and artificial data.

6.1.2. Supervised multi-scale locality sensitive hashing

Participants: Laurent Amsaleg, Li Weng.

LSH is a popular framework to generate compact representations of multimedia data, which can be used for content based search. However, the performance of LSH is limited by its unsupervised nature and the underlying feature scale. In [42], we proposed to improve LSH by incorporating two elements: supervised hash bit selection and multi-scale feature representation. First, a feature vector is represented by multiple scales. At each scale, the feature vector is divided into segments. The size of a segment is decreased gradually to make the representation correspond to a coarse-to-fine view of the feature. Then each segment is hashed to generate more bits than the target hash length. Finally the best ones are selected from the hash bit pool according to the notion of bit reliability, which is estimated by bit-level hypothesis testing. Extensive experiments have been performed to validate the proposal in two applications: near-duplicate image detection and approximate feature distance estimation. We first demonstrate that the feature scale can influence performance, which is often a neglected factor. Then we show that the proposed supervision method is effective. In particular, the performance increases with the size of the hash bit pool. Finally, the two elements are put together. The integrated scheme exhibits further improved performance.

6.1.3. Rotation and translation covariant match kernels for image retrieval

Participants: Andrei Bursuc, Teddy Furon, Hervé Jégou, Giorgos Tolias.

Most image encodings achieve orientation invariance by aligning the patches to their dominant orientations and translation invariance by completely ignoring patch position or by max-pooling. Albeit successful, such choices introduce too much invariance because they do not guarantee that the patches are rotated or translated consistently. In this work, we propose a geometric-aware aggregation strategy, which jointly encodes the local descriptors together with their patch dominant angle [38] and/or location [10]. The geometric attributes are encoded in a continuous manner by leveraging explicit feature maps. Our technique is compatible with generic match kernel formulation and can be employed along with several popular encoding methods, in particular bag of words, VLAD and the Fisher vector. The method is further combined with an efficient monomial embedding to provide a codebook-free method aggregating local descriptors into a single vector representation. Invariance is achieved by efficient similarity estimation of multiple rotations or translations, offered by a simple trigonometric polynomial. This strategy is effective for image search, as shown by experiments performed on standard benchmarks for image and particular object retrieval, namely Holidays and Oxford buildings.

6.1.4. Sequential pattern mining on audio data

Participants: Laurent Amsaleg, Guillaume Gravier, Simon Malinowski.

M. Sc. Internship of Corentin Hardy, in collaboration with René Quiniou, Inria Rennes, DREAM research team, within the framework of the STIC AmSud Maximum project and of the MOTIF Inria Associate Team.

Analyzing multimedia data is a challenging problem due to the quantity and complexity of such data. Mining for frequently recurring patterns is a task often ran to help discovering the underlying structure hidden in the data. This year, we have explored how data symbolization and sequential pattern mining techniques could help for mining recurring patterns in multimedia data. In [20], we have shown that even if sequential pattern mining techniques are very helpful in terms of computational efficiency, the data symbolization step is a crucial step to find for extracting relevant audio patterns.

6.1.5. Clustering by diverting supervised machine learning

Participants: Vincent Claveau, Teddy Furon, Guillaume Gravier.

M. Sc. Internship of Amélie Royer, ENS Rennes.

Clustering algorithms exploit an input similarity measure on the samples, which should be fine-tuned with the data format and the application at hand. However, manually defining a suitable similarity measure is a difficult task in case of limited prior knowledge or complex data structures for example. While supervised classification systems require a set of samples annotated with their ground-truth classes, recent studies have shown it is possible to exploit classifiers trained on an artificial annotation of the data in order to induce a similarity measure. In this work, we have proposed a unified framework, named similarity by iterative classifications (SIC), which explores the idea of diverting supervised learning for automatic similarity inference. We studied several of its theoretical and practical aspects. We also have implemented and evaluate SIC on three tasks of knowledge discovery on multimedia content. Results show that in most situations the proposed approach indeed benefits from the underlying classifier's properties and outperforms usual similarity measures for clustering applications.

6.1.6. Multimodal person discovery in TV broadcasts

Participant: Guillaume Gravier.

Work in collaboration with Cassio Elias dos Santos Jr. and William Robson Schwartz, in the framework of the Inria Associate Team MOTIF and of the STIC AmSud project Maximum.

Taking advantage of recent results on large-scale face comparison with partial least square, we developed various approaches for multimodal person discovery in TV broadcasts in the framework of the MediaEval 2015 international benchmark [30]. The task consists in naming the persons on screen that are speaking with no prior information, leveraging text overlays, speech transcripts as well as face and voice comparison. We investigated two distinct aspects of multimodal person discovery. One refers to face clusters, which are considered to propagate names associated with faces in one shot to other faces that probably belong to the

same person. The face clustering approach consists in calculating face similarities using partial least squares and a simple hierarchical approach. The other aspect refers to tag propagation in a graph-based approach where nodes are speaking faces and edges link similar faces/speakers. The advantage of the graph-based tag propagation is to not rely on face/speaker clustering, which we believe can be errorprone. The face clustering approach ranked among the top results in the international benchmark.

6.1.7. *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.

In [25], we addressed the problem of unsupervised program structuring with minimal prior knowledge about the program. We extended previous work to propose an approach able to identify multiple structures and infer structural grammars for recurrent TV programs of different types. The approach taken involves three sub-problems: i) we determine the structural elements contained in programs with minimal knowledge about which type of elements may be present; ii) we identify multiple structure for the programs if any and model the structures of programs; iii) we generate the structural grammar for each corresponding structure. Finally, we conducted use-case based evaluations on real recurrent programs of three different types to demonstrate the effectiveness of the proposed approach.

6.1.8. *Information retrieval for distributional semantics, and vice-versa*

Participants: Vincent Claveau, Ewa Kijak.

Distributional thesauri are useful in many tasks of natural language processing. In [33], [3], we address the problem of building and evaluating such thesauri with the help of information retrieval (IR) concepts. Two main contributions are proposed. First, in the continuation of previous work, we have shown how IR tools and concepts can be used with success to build thesauri. Through several experiments and by evaluating directly the results with reference lexicons, we show that some IR models outperform state-of-the-art systems. Secondly, we use IR as an application framework to indirectly evaluate the generated thesaurus. Here again, this task-based evaluation validate the IR approach used to build the thesaurus. Moreover, it allows us to compare these results with those from the direct evaluation framework used in the literature. The observed differences question these evaluation habits.

6.2. Multimedia content description and structuring

6.2.1. *Image description using component trees*

Participants: Petra Bosilj, Ewa Kijak.

In collaboration with Sébastien Lefèvre from Obelix Team (IRISA).

In this work, we explored the application of a tree-based feature extraction algorithm for the widely-used MSER features, and proposed a tree-of-shapes based detector of maximally stable regions. Changing an underlying component tree in the algorithm allows considering alternative properties and pixel orderings for extracting maximally stable regions. Performance evaluation was carried out on a standard benchmark in terms of repeatability and matching score under different image transformations, as well as in a large scale image retrieval setup, measuring mean average precision. The detector outperformed the baseline MSER in the retrieval experiments [37].

We also proposed a local region descriptor based on 2D shape-size pattern spectra, calculated on arbitrary connected regions, and combined with normalized central moments. The challenges when transitioning from global pattern spectra to the local ones were faced, and an exhaustive study on the parameters and the properties of the newly constructed descriptor was conducted. The descriptors were calculated on MSER regions, and evaluated in a simple retrieval system. Competitive performance with SIFT descriptors was achieved. An additional advantage of the proposed descriptors is their size which is less than half the size of SIFT [14], [15].

6.2.2. Improved motion description for action classification

Participant: Hervé Jégou.

In collaboration with Mihir Jain (University of Amsterdam, The Netherlands) and Patrick Bouthemy (Team-project SERPICO, Inria Rennes, France)

Even though the importance of explicitly integrating motion characteristics in video descriptions has been demonstrated by several recent papers on action classification, our current work concludes that adequately decomposing visual motion into dominant and residual motions, i.e., camera and scene motion, significantly improves action recognition algorithms. This holds true both for the extraction of the space-time trajectories and for computation of descriptors. We designed in [7] a new motion descriptor—the DCS descriptor—that captures additional information on local motion patterns enhancing results based on differential motion scalar quantities, divergence, curl and shear features. Finally, applying the recent VLAD coding technique proposed in image retrieval provides a substantial improvement for action recognition. These findings are complementary to each other and they outperformed all previously reported results by a significant margin on three challenging datasets: Hollywood 2, HMDB51 and Olympic Sports as reported in (Jain et al. (2013)).

6.2.3. Word embeddings and recurrent neural networks for spoken language understanding

Participants: Guillaume Gravier, Christian Raymond, Vedran Vukotić.

Recently, word embedding representations have been investigated for slot filling in spoken language understanding (SLU), along with the use of neural networks as classifiers. Neural networks, especially recurrent neural networks, which are adapted to sequence labeling problems, have been applied successfully on the popular ATIS database. In [29], we make a comparison of this kind of models with the previously state-of-the-art conditional random fields (CRF) classifier on a more challenging SLU database. We show that, despite efficient word representations used within these neural networks, their ability to process sequences is still significantly lower than for CRF, while also having a drawback of higher computational costs, and that the ability of CRF to model output label dependencies is crucial for SLU.

6.2.4. Hierarchical topic structuring

Participants: Guillaume Gravier, Pascale Sébillot, Anca-Roxana Şimon.

Topic segmentation traditionally relies on lexical cohesion measured through word re-occurrences to output a dense segmentation, either linear or hierarchical. We have proposed a novel organization of the topical structure of textual content [28]. Rather than searching for topic shifts to yield dense segmentation, our algorithm extracts topically focused fragments organized in a hierarchical manner. This is achieved by leveraging the temporal distribution of word re-occurrences, searching for bursts, to skirt the limits imposed by a global counting of lexical re-occurrences within segments. Comparison to a reference dense segmentation on varied datasets indicates that we can achieve a better topic focus while retrieving all of the important aspects of a text.

6.2.5. Partial least square hashing for large-scale face identification

Participants: Guillaume Gravier, Ewa Kijak.

Work performed with Cassio Elias dos Santos Jr. during his 3 months visit, in collaboration with William Robson Schwartz (UFMG, Brasil), in the framework of the Inria Associate Team MOTIF.

Face recognition has been largely studied in past years. However, most of the related work focus on increasing accuracy and/or speed to test a single pair probe-subject. In [31], we introduced a novel method inspired by the success of locality sensing hashing applied to large general purpose datasets and by the robustness provided by partial least squares analysis when applied to large sets of feature vectors for face recognition. The result is a robust hashing method compatible with feature combination for fast computation of a short list of candidates in a large gallery of subjects. We provided theoretical support and practical principles for the proposed hashing method that may be reused in further development of hash functions applied to face galleries. Comparative evaluations on the FERET and FRGCv1 datasets demonstrate a speedup of a factor 16 compared to scanning all subjects in the face gallery.

6.2.6. Selection strategies for active learning in NLP

Participants: Vincent Claveau, Ewa Kijak.

Nowadays, many NLP problems are modeled as supervised machine learning tasks, especially when it comes to information extraction. Consequently, the cost of the expertise needed to annotate the examples is a widespread issue. Active learning offers a framework to that issue, allowing to control the annotation cost while maximizing the classifier performance, but it relies on the key step of choosing which example will be proposed to the expert. In [3], we have examined and proposed such selection strategies in the specific case of conditional random fields which are largely used in NLP. On the one hand, we have proposed a simple method to correct a bias of certain state-of-the-art selection techniques. On the other hand, we have detailed an original approach to select the examples, based on the respect of proportions in the datasets. These contributions are validated over a large range of experiments implying several tasks and datasets, including named entity recognition, chunking, phonetization, word sense disambiguation.

6.2.7. Tree-structured named entities extraction from competing speech transcripts

Participant: Christian Raymond.

When real applications are working with automatic speech transcription, the first source of error does not originate from the incoherence in the analysis of the application but from the noise in the automatic transcriptions. In [41], we present a simple but effective method to generate a new transcription of better quality by combining utterances from competing transcriptions. We have extended a structured named entity (NE) recognizer submitted during the ETAPE challenge. Working on French TV and radio programs, our system revises the transcriptions provided by making use of the NEs it has detected. Our results suggest that combining the transcribed utterances which optimize the F-measure, rather than minimizing the WER scores, allows the generation of a better transcription for NE extraction. The results show a small but significant improvement of 0.9 % SER against the baseline system on the ROVER transcription. These are the best performances reported to date on this corpus.

6.3. Content-based information retrieval

6.3.1. A comparison of dense region detectors for image search and fine-grained classification

Participants: Hervé Jégou, Ahmet Iscen, Giorgos Tolias.

In collaboration with Philippe-Henri Gosselin (ETIS team, ENSEA, Cergy, France)

We consider a pipeline for image classification or search based on coding approaches like bag of words or Fisher vectors. In this context, the most common approach is to extract the image patches regularly in a dense manner on several scales. In [6], we propose and evaluate alternative choices to extract patches densely. Beyond simple strategies derived from regular interest region detectors, we propose approaches based on super-pixels, edges, and a bank of Zernike filters used as detectors. The different approaches are evaluated on recent image retrieval and fine-grain classification benchmarks. Our results show that the regular dense detector is outperformed by other methods in most situations, leading us to improve the state of the art in comparable setups on standard retrieval and fine-grain benchmarks. As a byproduct of our study, we show that existing methods for blob and super-pixel extraction achieve high accuracy if the patches are extracted along the edges and not around the detected regions.

6.3.2. Efficient large-scale similarity search using matrix factorization

Participants: Teddy Furon, Ahmet Iscen.

In collaboration with Michael Rabbat (McGill University, Montréal, Canada)

We considered the image retrieval problem of finding the images in a dataset that are most similar to a query image. Our goal is to reduce the number of vector operations and memory for performing a search without sacrificing accuracy of the returned images. We adopt a group testing formulation and design the decoding architecture using either dictionary learning or eigendecomposition. The latter is a plausible option for small-to-medium sized problems with high-dimensional global image descriptors, whereas dictionary learning is applicable in large-scale scenario. We evaluate our approach both for global descriptors obtained from SIFT and CNN features. Experiments with standard image search benchmarks, including the Yahoo100M dataset comprising 100 million images, show that our method gives comparable (and sometimes superior) accuracy compared to exhaustive search while requiring only 10 % of the vector operations and memory. Moreover, for the same search complexity, our method gives significantly better accuracy compared to approaches based on dimensionality reduction or locality sensitive hashing [43].

6.3.3. *Explicit embeddings for nearest neighbor search with Mercer kernels*

Participant: Hervé Jégou.

In collaboration with Anthony Bourrier and Patrick Pérez (Technicolor, Rennes, France), Florent Perronnin (Xerox, Grenoble, France) Rémi Gribonval (Team-project PANAMA, Inria Rennes, France).

Many approximate nearest neighbor search algorithms operate under memory constraints, by computing short signatures for database vectors while roughly keeping the neighborhoods for the distance of interest. Encoding procedures designed for the Euclidean distance have attracted much attention in the last decade. In the case where the distance of interest is based on a Mercer kernel, we propose a simple, yet effective two-step encoding scheme: first, compute an explicit embedding to map the initial space into a Euclidean space; second, apply an encoding step designed to work with the Euclidean distance. Comparing this simple baseline with existing methods relying on implicit encoding, we demonstrate better search recall for similar code sizes with the chi-square kernel in databases comprised of visual descriptors, outperforming concurrent state-of-the-art techniques by a large margin [2].

6.3.4. *Image search with selective match kernels: aggregation across single and multiple images*

Participants: Hervé Jégou, Giorgos Tolias.

In collaboration with Yannis Avrithis (National Technical University of Athens, Greece)

Our work [9] considers a family of metrics to compare images based on their local descriptors. It encompasses the VLAD descriptor and matching techniques such as Hamming Embedding. Making the bridge between these approaches leads us to propose a match kernel that takes the best of existing techniques by combining an aggregation procedure with a selective match kernel. The representation underpinning this kernel is approximated, providing a large scale image search both precise and scalable, as shown by our experiments on several benchmarks. We show that the same aggregation procedure, originally applied per image, can effectively operate on groups of similar features found across multiple images. This method implicitly performs feature set augmentation, while enjoying savings in memory requirements at the same time. Finally, the proposed method is shown effective for place recognition, outperforming state of the art methods on a large scale landmark recognition benchmark.

6.3.5. *Early burst detection for memory-efficient image retrieval*

Participant: Hervé Jégou.

In collaboration with Miaojing Shi, visiting Ph. D. student from Pekin University, and Yannis Avrithis (National Technical University of Athens, Greece)

Recent works show that image comparison based on local descriptors is corrupted by visual bursts, which tend to dominate the image similarity. The existing strategies, like power-law normalization, improve the results by discounting the contribution of visual bursts to the image similarity. We proposed to explicitly detect the visual bursts in an image at an early stage. We compare several detection strategies jointly taking into account feature similarity and geometrical quantities. The bursty groups are merged into meta-features, which are used as input to state-of-the-art image search systems such as VLAD or the selective match kernel. Then, we show the interest of using this strategy in an asymmetrical manner, with only the database features being aggregated but not those of the query. Extensive experiments performed on public benchmarks for visual retrieval show the benefits of our method, which achieves performance on par with the state of the art but with a significantly reduced complexity, thanks to the lower number of features fed to the indexing system [40], [44].

6.3.6. *Biomedical information retrieval*

Participants: Vincent Claveau, Ewa Kijak.

In collaboration with N. Grabar (STL), T. Hamon (LIMSI), and S. Le Maguer (Univ. Saarland).

The right of patients to access their clinical health record is granted by the code of Santé Publique. Yet, this piece of content remains difficult to understand. We propose different IR experiments in which we use queries defined by patients in order to find relevant documents [3], [16]. We use the Indri search engine, based on statistical language modeling, as well as semantic resources. More precisely, our approaches are chiefly based on the terminological variation (e.g., synonyms, abbreviations) to link between expert and patient languages. Various combinations of resources and Indri settings are explored, mostly based on query expansion.

6.4. Linking, navigation and analytics

6.4.1. *Sentiment analysis on social networks*

Participants: Vincent Claveau, Christian Raymond, Vedran Vukotić.

In the framework of our participation to the DeFT 2015 text-mining challenge, we have developed sentiment-analysis methods for tweets [34]. Several sub-tasks have been considered: i) valence classification of tweets and ii) fine-grained classification of tweets (which includes two sub-tasks: detection of the generic class of the information expressed in a tweet and detection of the specific class of the opinion/sentiment/emotion. For all three problems, we adopt a standard machine learning framework. More precisely, three main methods are proposed and their feasibility for the tasks is analyzed: i) decision trees with boosting (bonzaiboost), ii) naive Bayes with Okapi and iii) convolutional neural networks (CNNs). Our approaches are voluntarily knowledge free and text-based only, we do not exploit external resources (lexicons, corpora) or tweet metadata. It allows us to evaluate the interest of each method and of traditional bag-of-words representations vs. word embeddings. Methods using simple ML frameworks and IR-based similarity metrics have been demonstrated to yield the best results.

6.4.2. *A multi-dimensional data model for personal photo browsing*

Participant: Laurent Amsaleg.

Work performed in the framework of the CNRS PICS MMAAnalytics, and in collaboration with Marcel Worring, Univeristy of Amsterdam (The Netherlands)

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 proposed a multi-dimensional model for media browsing, called M3, that combines MDA concepts with concepts from faceted browsing [21]. 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.

6.4.3. NLP-driven hyperlink construction in broadcast videos

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

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

In the context of the the CominLabs project "Linking media in acceptable hypergraphs" dedicated to the creation of explicit and meaningful links between multimedia documents or fragments of documents, we have introduced a typology of possible links between contents of a multimedia news corpus [32]. While several typologies have been proposed and used by the community, we argue that they are not adapted to rich and large corpora which can contain texts, videos, or radio stations recordings. We have defined a new typology, as a first step towards automatically creating and categorizing links between documents' fragments in order to create new ways to navigate, explore, and extract knowledge from large collections.

We also investigated video hyperlinking based on speech transcripts, leveraging a hierarchical topical structure to address two essential aspects of hyperlinking, namely, serendipity control and link justification [26]. We proposed and compared different approaches exploiting a hierarchy of topic models as an intermediate representation to compare the transcripts of video segments. These hierarchical representations offer a basis to characterize the hyperlinks, thanks to the knowledge of the topics which contributed to the creation of the links, and to control serendipity by choosing to give more weights to either general or specific topics. Experiments have been performed on BBC videos from the Search and Hyperlinking task at MediaEval. Link precisions similar to those of direct text comparison have been achieved however exhibiting different targets along with a potential control of serendipity.

The Search and Anchoring in Video Archives task at MediaEval addressed two issues: The Search part aims at returning a ranked list of video segments that are relevant to a textual user query; The Anchoring part focuses on identifying video segments that would encourage further exploration within the archive. Capitalizing on the experience acquired in previous participations, we implemented a two step approach for both sub-tasks [27]. The first step, common to both, consists in generating a list of potential anchor segments and response-query segments relying on a hierarchical topical structuring technique. In the second step, for each query, the best 20 segments are selected according to content-based comparisons, while for the anchor detection sub-task, the segments are ranked based on a cohesion measure. The use of a hierarchical topical structure helps to propose segments of variable length at different levels of details with precise jump-in points for them. More, the algorithm deriving the structure relies on the burstiness phenomenon in word occurrences which gives an advantage over the classical bag-of-words model.

6.4.4. Information extraction

Participants: Vincent Claveau, Ewa Kijak.

In collaboration with X. Tannier (LIMSI), A. Vilnat (LIMSI) and B. Arnulphy (ANR).

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 is available for French. In [12], we try 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.

6.5. Participation in benchmarking initiatives

- Video hyperlinking, TRECVID
- Search and anchoring, Mediaeval Multimedia International Benchmark
- Multimodal person discovery in broadcast TV, Mediaeval Multimedia International Benchmark
- DeFT 2015 text-mining challenge

7. Bilateral Contracts and Grants with Industry

7.1. Bilateral Contracts with Industry

Teddy Furon spent 20 % of his time during 6 months to transfer research result to IRT B-com

CIFRE Ph. D. contract with Institut National de l'Audiovisuel (Bingqing Qu)

CIFRE Ph. D. contract with Technicolor (Himalaya Jain)

Ph. D. contract with Alcatel-Lucent Bell Labs (Raghavendran Balu) in the framework of the joint Inria-Alcatel Lucent lab.

8. Partnerships and Cooperations

8.1. Regional Initiatives

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

Participants: Rémi Bois, Vincent Claveau, Guillaume Gravier, Grégoire Jadi, 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: 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*

Participant: 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.

8.2.3. ANR Project IDFRAud

Participant: Teddy Furon.

Duration: 3 years, started in Feb. 2015

Partners: AriadNext, IRCGN, École Nationale Supérieure de Police

The IDFRAud project consists in proposing an automatic solution for ID analysis and integrity verification. Our ID analysis goes through three processes: classification, text extraction and ID verification. The three processes rely on a set of rules that are externalized in formal manner in order to allow easy management and evolving capabilities. This leads us to the ID knowledge management module. Finally, IDFRAud addresses the forensic link detection problem and to propose an automatic analysis engine that can be continuously applied on the detected fraud ID database. Cluster analysis methods are used to discover relations between false IDs in their multidimensional feature space. This pattern extraction module will be coupled with a suitable visualization mechanism in order to facilitate the comprehension and the analysis of extracted groups of inter-linked fraud cases.

8.2.4. FUI 19 NexGenTV

Participants: Vincent Claveau, Guillaume Gravier, Ewa Kijak, Pascale Sébillot.

Duration: 2.5 years, started in May 2015

Partners: Eurecom, Avisto Telecom, Wildmoka, Envivio

Television is undergoing a revolution, moving from the TV screen to multiple screens. Today's user watches TV and, at the same time, browses the web on a tablet, sends SMS, posts comments on social networks, searches for complementary information on the program, etc. Facing this situation, NexGen-TV aims at developing a generic solution for the enrichment, the linking and the retrieval of video content targeting the cost-cutting edition of second screen and multiscreen applications for broadcast TV. The main outcome of the project will be a software platform to aggregate and distribute video content via a second-screen edition interface connected to social media. The curation interface will primarily make use of multimedia and social media content segmentation, description, linking and retrieval. Multiscreen applications will be developed on various domains, e.g., sports, news.

8.3. International Initiatives

8.3.1. Inria Associate Teams not involved in an Inria International Labs

8.3.1.1. MOTIF

Title: Unsupervised motif discovery in multimedia content

International Partner (Institution - Laboratory - Researcher):

Pontificia Universidade Católica de Minas Gerais, Brasil - VIPLAB - Silvio Jamil Guimãraes

Universidade Federal Minas Gerais, Brasil - NPDI - Arnaldo Albuquerque de Araújo

Duration: 2014 - 2017

See also: <http://www-linkmedia.irisa.fr/motif>

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.

8.3.2. Inria International Partners

8.3.2.1. Informal International Partners

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

8.3.3. Participation In 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
- France Berkeley Fund Graph-NN: Computing and Manipulating Very Large Graphs of Nearest Neighbors
 - International coordinator: Laurent Amsaleg, CNRS – IRISA, France
 - Scientific coordinators : Michael Franklin (AMPLab, UC Berkeley)
 - Jun. 2015 - Dec. 2015

8.4. International Research Visitors

8.4.1. Visits of International Scientists

8.4.1.1. Internships

Bùi Văn Thạch (Ph.D. Student)
Date: Oct 2015 - Nov 2015
Institution: National University of Sokendai, Japan

8.4.2. Visits to International Teams

Ahmet Iscen
Date: Apr 2015 - Jun 2015
Institution: McGill University, Montreal, Canada

8.4.2.1. Explorer programme

Balu Raghavendran

Date: Jul 2015 - Sep 2015

Institution: [University of California Berkeley](#) (United States of America)

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific events organisation

9.1.1.1. General chair, scientific chair

Teddy Furon co-organized a GdR-ISIS workshop on Biometrics, Multimedia Indexing and Privacy.

Guillaume Gravier was general chair of the ACM Multimedia Third Workshop on Speech, Language and Audio in Multimedia (SLAM 2015) and is president of the steering committee of the workshop.

Pascale Sébillot was a member of the acting presidency of Conf. Francophone en Traitement Automatique des Langues Naturelles.

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

9.1.2. Scientific events selection

9.1.2.1. Chair of conference program committees

Vincent Claveau was area chair of Conf. Francophone en Traitement Automatique des Langues Naturelles.

Guillaume Gravier was chair of the program committee of the ACM Multimedia Third Workshop on Speech, Language and Audio in Multimedia (SLAM 2015).

Pascale Sébillot was area chair of Conf. Francophone en Traitement Automatique des Langues Naturelles.

9.1.2.2. Member of the conference program committees

Laurent Amsaleg was a PC member of: ACM Intl. Conf. on Multimedia; VISI, ACM Intl. Conf. on Multimedia Retrieval; IEEE Intl. Conf. on Multimedia and Exhibition; Intl. Conf. on Multimedia Modeling; Intl. Workshop on Content-Based Multimedia Indexing; Intl. Conf. on Similarity Search and Applications; Intl. Conf. on Signal Image Technology & Internet-based Systems.

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

Teddy Furon was a PC member of: IEEE Work. on Information Forensics and Security; Intl. Conf. on Acoustics, Speech and Signal Processing; Intl. Conf. on Multimedia, Communication and Computing; ACM Intl. Conf. on Multimedia; European Signal Processing Conf.

Guillaume Gravier was a PC member of: ACM Intl. Conf. on Multimedia; IEEE Intl. Conf. on Multimedia and Exhibition; Annual Conf. of the Intl. Speech Communication Association; IEEE Intl. Workshop on Multimedia Signal Processing; Intl. Workshop on Content-Based Multimedia Indexing; Intl. Conf. on Knowledge and Systems Engineering; Intl. Conf. on Statistical Language and Speech Processing.

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; Conf. Francophone en Traitement Automatique des Langues Naturelles.

Pascale Sébillot was a PC member of Intl. Conf. on Terminology and Artificial Intelligence.

9.1.2.3. Reviewer

Vincent Claveau was a reviewer of: Intl. Conf. on Machine Learning.

9.1.3. Journal

9.1.3.1. Member of the editorial boards

Vincent Claveau is member of the editorial board of the journal *Traitement Automatique des Langues*.

Teddy Furon was member of the editorial board of *IEEE Trans. on Information Forensics and Security* (up to March 2015).

Christian Raymond is member of the editorial board of the online journal *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 - Reviewing activities

Laurent Amsaleg reviewed for *Knowledge and Information Systems*.

Vincent Claveau reviewed for: *Multimedia Tools and Applications*, *Traitement Automatique des Langues*.

Teddy Furon reviewed for: *IEEE Trans. on Information Forensics and Security*; *IEEE Trans. on Multimedia, Data Mining and Knowledge Discovery*; *ACM Trans. on Information Systems*; *IEEE Trans. on Circuits and Systems for Video Technology*; *Digital Signal Processing Journal*; *Applied and Computational Harmonic Analysis Journal*; *IET Information Security Journal*.

Guillaume Gravier reviewed for: *IEEE Trans. on Audio Speech and Language*; *IEEE Trans. on Image Processing*; *EURASIP Journal on Audio, Speech, and Music Processing*; *Journal of Computer Science and Technology*; *Multimedia Tools and Applications*;

Christian Raymond reviewed for *Computer Speech and Language*.

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

9.1.4. Invited talks

Vincent Claveau gave an invited talk about biomedical NLP in Rennes University Hospital's computer science department.

9.1.5. Leadership within the scientific community

Vincent Claveau is finance head of the *Association pour la Recherche d'Informations et ses Applications (ARIA)*.

Vincent Claveau is deputy head of the *GdR MaDICS*, a CNRS inter-lab initiative to promote research about Big Data and Data Science.

Guillaume Gravier is president of the *Association Francophone de la Communication Parlée (AFPC)*, French-speaking branch of the *Intl. Speech Communication Association*.

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*.

Guillaume Gravier is the technical representative of Inria in the *cPPP Big Data Value Association*, actively working on technical aspects of data analytics.

9.1.6. Scientific expertise

Vincent Claveau served as expert for the *ERC Consolidator grant programme*, for the *FNRS (Belgian Funding agency)*, for the *Programme Hubert Curien*.

Teddy Furon reviewed projects for *Alpes Grenoble Innovation and French National Research Agency (ANR)*.

Teddy Furon is the scientific adviser for startup company Lamark (20 % of his time since July 2015).

Guillaume Gravier was vice-president of the Scientific Evaluation Committee of the National Research Agency for the theme 'HCI, Content, Knowledge, Big Data, Simulation, HPC'.

Pascale Sébillot reviewed projects for the Natural Sciences and Engineering Research Council of Canada.

9.1.7. Research administration

Guillaume Gravier is a member of the Board of the technology cluster Images & Réseaux.

Guillaume Gravier is a member of the Board of the Comité des Projets of Inria - Rennes Bretagne Atlantique.

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

Pascale Sébillot is a member of the theses advisory committee of the Matisse doctoral school.

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

For researchers, all activities are given. For professors and assistant professors, only courses at the M. Sc. level are listed.

Master: Laurent Amsaleg, Multidimensional 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, Probability and statistics, 10h, L3, INSA Rennes, France

Master: Guillaume Gravier, Data analysis and probabilistic modeling, 30h, M2, University Rennes 1, 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, 13h, M2, University Rennes 1, France

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

Master: Ewa Kijak, Computer vision, 15h, M2, ESIR, France

Master: Simon Malinowski, Short-term time series prediction, 29h, M1, Univ. Rennes 1

Master: Simon Malinowski, Supervised Learning, 10h, M2, Univ. Rennes 1

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

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

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

9.2.2. Supervision

PhD: Mohammed-Haykel Boukadida, Video summarization based on constraint programming, defended Dec. 2015, Patrick Gros

PhD: Bingqing Qu, Structure discovery in collections of recurrent TV programs, defended Dec. 2015, Guillaume Gravier

PhD: Anca Roxana Simon, Semantic structuring of video collections from speech: segmentation and hyperlinking, defended Dec. 2015, 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: Rémi Bois, Navigable directed multimedia hypergraphs: construction and exploitation, started October 2014, Guillaume Gravier and Pascale Sébillot

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: Ricardo Carlini Sperandio, Unsupervised motif mining in multimedia time series, started August 2015, Laurent Amsaleg and Guillaume Gravier

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

PhD in progress: Cédric Maigrot, Detecting fake information on social networks, started October 2015, Laurent Amsaleg, Vincent Claveau and Ewa Kijak

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

Laurent Amsaleg

PhD, Herwig Lejsek, Reykjavík University

Teddy Furon

PhD, Wei Fan, University of Grenoble

Guillaume Gravier

HDR, reviewer, Slim Essid, Telecom ParisTech

PhD, president, Grégor Dupuis, Université du Maine

Ewa Kijak

PhD, Cyrille Beaudry, Université de la Rochelle

Simon Malinowsky

PhD, Racha Khelif, Université de Franche-Comté

Pascale Sébillot

PhD, reviewer, Munshi Asadullah, Univ. Paris-Sud

PhD, reviewer, Sondes Bannour, Univ. Paris-Nord

PhD, president, Valéria Lelli Leitão Dantas, , Univ. Rennes 1

PhD, president, Bingqing Qu, Univ. Rennes 1

PhD, Mohamed Haykel Boukadida, Univ. Rennes 1

9.3. Popularization

Pascale Sébillot: Invited speaker (6h) Introduction to issues and solutions in NLP, URFIST Bretagne et Pays de la Loire seminar, June 2015.

10. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

- [1] A.-R. SIMON. *Semantic structuring of video collections from speech: segmentation and hyperlinking*, Université de Rennes 1, December 2015, <https://tel.archives-ouvertes.fr/tel-01253678>

Articles in International Peer-Reviewed Journals

- [2] 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>
- [3] V. CLAVEAU, E. KIJAK. *Thésaurus distributionnels pour la recherche d'information et vice-versa*, in "Revue des Sciences et Technologies de l'Information - Série Document Numérique", 2015, vol. 18, n^o 2-3 [DOI : 10.3166/DN.18.2-3.101-121], <https://hal.archives-ouvertes.fr/hal-01226551>
- [4] V. CLAVEAU, S. LEFÈVRE. *Topic segmentation of TV-streams by watershed transform and vectorization*, in "Computer Speech and Language", 2015, vol. 29, n^o 1, pp. 63-80 [DOI : 10.1016/J.CSL.2014.04.006], <https://hal.archives-ouvertes.fr/hal-00998259>
- [5] T. GUYET, S. MALINOWSKI, M.-C. BENYOUNÈS. *Extraction des zones cohérentes par l'analyse spatio-temporelle d'images de télédétection*, in "International Journal of Geomatics and Spatial Analysis / Revue Internationale de Géomatique", 2015, 22 p. , <https://hal.inria.fr/hal-01184095>
- [6] A. ISCEN, G. TOLIAS, P.-H. GOSSELIN, H. JÉGOU. *A comparison of dense region detectors for image search and fine-grained classification*, in "IEEE Transactions on Image Processing", 2015, 00 p. [DOI : 10.1109/TIP.2015.2423557], <https://hal.inria.fr/hal-01143201>
- [7] M. JAIN, H. JÉGOU, P. BOUTHEMY. *Improved motion description for action classification*, in "Frontiers in ICT", December 2015 [DOI : 10.3389/FICT.2015.00028], <https://hal.inria.fr/hal-01247605>
- [8] P. TIRILLY, V. CLAVEAU, P. GROS. *From Text to Images: Weighting Schemes for Image Retrieval*, in "Journal of Multimedia", January 2015, vol. 10, n^o 1, pp. 1-21 [DOI : 10.4304/JMM.10.01.1-21], <https://hal.archives-ouvertes.fr/hal-01122069>
- [9] G. TOLIAS, Y. AVRITHIS, H. JÉGOU. *Image search with selective match kernels: aggregation across single and multiple images*, in "International Journal of Computer Vision", 2015, 00 p. , <https://hal.inria.fr/hal-01131898>
- [10] G. TOLIAS, A. BURSUC, T. FURON, H. JÉGOU. *Rotation and translation covariant match kernels for image retrieval*, in "Computer Vision and Image Understanding", June 2015, 15 p. [DOI : 10.1016/J.CVIU.2015.06.007], <https://hal.archives-ouvertes.fr/hal-01168525>

International Conferences with Proceedings

- [11] L. AMSALEG, C. OUSSAMA, T. FURON, S. GIRARD, M. E. HOULE, K.-I. KAWARABAYASHI. *Estimating Local Intrinsic Dimensionality*, in "21st Conf. on Knowledge Discovery and Data Mining, KDD2015", Sidney, Australia, ACM (editor), ACM, August 2015, <https://hal.inria.fr/hal-01159217>
- [12] B. ARNULPHY, V. CLAVEAU, X. TANNIER, A. VILNAT. *Supervised Machine Learning Techniques to Detect TimeML Events in French and English*, in "20th International Conference on Applications of Natural Language to Information Systems, NLDB 2015", Passau, Germany, C. BEIMANN, S. HANDSCHUCH, A. FREITAS, F. MEZIANE, E. MÉTAIS (editors), Springer, June 2015, vol. 9103 [DOI : 10.1007/978-3-319-19581-0_2], <https://hal.archives-ouvertes.fr/hal-01226541>
- [13] A. BAILLY, S. MALINOWSKI, R. TAVENARD, T. GUYET, L. CHAPEL. *Bag-of-Temporal-SIFT-Words for Time Series Classification*, in "ECML/PKDD Workshop on Advanced Analytics and Learning on Temporal Data", Porto, Portugal, September 2015, <https://halshs.archives-ouvertes.fr/halshs-01184900>
- [14] P. BOSILJ, E. KIJAK, M. H. WILKINSON, S. LEFÈVRE. *Short Local Descriptors from 2D Connected Pattern Spectra*, in "IEEE International Conference on Image Processing", Quebec City, Canada, 2015, <https://hal.inria.fr/hal-01134071>
- [15] P. BOSILJ, M. H. WILKINSON, E. KIJAK, S. LEFÈVRE. *Local 2D Pattern Spectra as Connected Region Descriptors*, in "International Symposium on Mathematical Morphology", Reykjavik, Iceland, Lecture Notes in Computer Science, Springer, 2015, vol. 9082, pp. 182-193 [DOI : 10.1007/978-3-319-18720-4_16], <https://hal.archives-ouvertes.fr/hal-01168146>
- [16] V. CLAVEAU, T. HAMON, S. LE MAGUER, N. GRABAR. *Health consumer-oriented information retrieval*, in "Medical Informatics Europe conference, MIE 2015", Madrid, Spain, May 2015, <https://hal.archives-ouvertes.fr/hal-01226544>
- [17] V. CLAVEAU, E. KIJAK. *Strategies to select examples for Active Learning with Conditional Random Fields*, in "Conférence TALN 2015", Caen, France, June 2015, <https://hal.archives-ouvertes.fr/hal-01206847>
- [18] G. GRAVIER, G. J. JONES, M. LARSON, R. ORDELMAN. *Overview of the 2015 Workshop on Speech, Language and Audio in Multimedia*, in "ACM International Conference on Multimedia", Brisbane, Australia, 2015 [DOI : 10.1145/2733373.2806414], <https://hal.archives-ouvertes.fr/hal-01186433>
- [19] G. GRAVIER, M. RAGOT, L. AMSALEG, R. BOIS, G. JADI, E. JAMET, L. MONCEAUX, P. SÉBILLOT. *Shaping-Up Multimedia Analytics: Needs and Expectations of Media Professionals*, in "The 22nd International Conference on Multimedia Modelling, Special Session Perspectives on Multimedia Analytics", Miami, United States, January 2016, <https://hal.inria.fr/hal-01214829>
- [20] C. HARDY, L. AMSALEG, G. GRAVIER, S. MALINOWSKI, R. QUINIOU. *Sequential pattern mining on multimedia data*, in "European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Database Workshop on Advanced Analytics and Learning on Temporal Data", Porto, Portugal, 2015, <https://hal.archives-ouvertes.fr/hal-01186446>
- [21] B. T. JÓNSSON, G. TÓMASSON, H. SIGURÞÓRSSON, Á. ERÍKSDÓTTIR, L. AMSALEG, M. K. LARUS-DÓTTIR. *A Multi-Dimensional Data Model for Personal Photo Browsing*, in "International Conference on Multimedia Modelling", Sydney, Australia, 2015, <https://hal.inria.fr/hal-01083344>

- [22] B. P. JÓNSSON, M. WORRING, J. ZAHÁLKA, S. RUDINAC, L. AMSALEG. *Ten Research Questions for Scalable Multimedia Analytics*, in "The 22nd International Conference on Multimedia Modelling, Special Session Perspectives on Multimedia Analytics", Miami, United States, January 2016, <https://hal.inria.fr/hal-01214847>
- [23] S. LE MAGUER, T. HAMON, N. GRABAR, V. CLAVEAU. *Recherche d'information médicale pour le patient Impact de ressources terminologiques*, in "CONFérence en Recherche d'Information et Applications, CORIA 2015", Paris, France, March 2015, <https://hal.archives-ouvertes.fr/hal-01226537>
- [24] R. QADER, G. LECORVÉ, D. LOLIVE, P. SÉBILLOT. *Probabilistic Speaker Pronunciation Adaptation for Spontaneous Speech Synthesis Using Linguistic Features*, in "International Conference on Statistical Language and Speech Processing (SLSP)", Budapest, Hungary, November 2015, 12 p. , <https://hal.inria.fr/hal-01181192>
- [25] 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>
- [26] A.-R. SIMON, R. BOIS, G. GRAVIER, P. SÉBILLOT, E. MORIN, S. MOENS. *Hierarchical Topic Models for Language-based Video Hyperlinking*, in "Workshop on Speech, Language and Audio in Multimedia", Brisbane, Australia, 2015 [DOI : 10.1145/2802558.2814642], <https://hal.archives-ouvertes.fr/hal-01186429>
- [27] A.-R. SIMON, G. GRAVIER, P. SÉBILLOT. *IRISA at MediaEval 2015: Search and Anchoring in Video Archives Task*, in "Working Notes Proceedings of the MediaEval Workshop", Wurzen, Germany, 2015, <https://hal.archives-ouvertes.fr/hal-01196176>
- [28] A. SIMON, P. SÉBILLOT, G. GRAVIER. *Hierarchical topic structuring: from dense segmentation to topically focused fragments via burst analysis*, in "Recent Advances on Natural Language Processing", Hissar, Bulgaria, 2015, <https://hal.archives-ouvertes.fr/hal-01186443>
- [29] V. VUKOTIC, C. RAYMOND, G. GRAVIER. *Is it time to switch to Word Embedding and Recurrent Neural Networks for Spoken Language Understanding?*, in "InterSpeech", Dresde, Germany, September 2015, <https://hal.inria.fr/hal-01196915>
- [30] C. E. DOS SANTOS JR., G. GRAVIER, W. ROBSON SCHWARTZ. *SSIG and IRISA at Multimodal Person Discovery*, in "Working Notes Proceedings of the MediaEval Workshop", Wurzen, Germany, 2015, <https://hal.archives-ouvertes.fr/hal-01196171>
- [31] C. E. DOS SANTOS JR., E. KIJAK, G. GRAVIER, W. ROBSON SCHWARTZ. *Learning to Hash Faces Using Large Feature Vectors*, in "International Workshop on Content-based Multimedia Indexing", Prague, Czech Republic, 2015, <https://hal.archives-ouvertes.fr/hal-01186444>

National Conferences with Proceedings

- [32] R. BOIS, G. GRAVIER, P. SÉBILLOT, E. MORIN. *Towards a typology of links between journalistic content*, in "22e conférence Traitement automatique des langues naturelles, TALN 2015", Caen, France, June 2015, pp. 515-521, papier court, <https://hal.inria.fr/hal-01196052>

- [33] V. CLAVEAU, E. KIJAK. *Thésaurus distributionnels pour la recherche d'information et vice-versa*, in "Conférence en Recherche d'Information et Applications", Paris, France, Actes de la conférence CORIA 2015, March 2015, <https://hal.archives-ouvertes.fr/hal-01226532>
- [34] V. VUKOTIC, V. CLAVEAU, C. RAYMOND. *IRISA at DeFT 2015: Supervised and Unsupervised Methods in Sentiment Analysis*, in "DeFT, Défi Fouille de Texte, joint à la conférence TALN 2015", Caen, France, Actes de l'atelier DeFT, Défi Fouille de Texte, joint à la conférence TALN 2015, June 2015, <https://hal.archives-ouvertes.fr/hal-01226528>

Conferences without Proceedings

- [35] R. BALU, T. FURON, L. AMSALEG. *Sketching techniques for very large matrix factorization*, in "ECIR 2016 - 38th European Conference on Information Retrieval", Padoue, Italy, 2016, <https://hal.inria.fr/hal-01249621>
- [36] P. BOSILJ, E. APTOULA, S. LEFÈVRE, E. KIJAK. *Satellite image retrieval with pattern spectra descriptors*, in "International Conference on Image Information Mining: Geospatial Intelligence from Earth Observation (IIM)", Bucarest, Romania, 2015, <https://hal.archives-ouvertes.fr/hal-01253815>
- [37] P. BOSILJ, E. KIJAK, S. LEFÈVRE. *Beyond MSER: Maximally Stable Regions using Tree of Shapes*, in "British Machine Vision Conference", Swansea, United Kingdom, September 2015, <https://hal.archives-ouvertes.fr/hal-01194372>
- [38] A. BURSUC, G. TOLIAS, H. JÉGOU. *Kernel Local Descriptors with Implicit Rotation Matching*, in "ACM International Conference on Multimedia Retrieval", Shanghai, China, 2015, <https://hal.inria.fr/hal-01145656>
- [39] T. GAILLAT, P. SÉBILLOT, N. BALLIER. *Comparing corpora to identify learner-specific features of English: The case of this, that and it*, in "Learner Corpus Research Conference (LCR 2015)", Radboud, Netherlands, Radboud University, September 2015, pp. 68-69, <https://hal-univ-diderot.archives-ouvertes.fr/hal-01239837>
- [40] M. SHI, Y. AVRITHIS, H. JÉGOU. *Early burst detection for memory-efficient image retrieval*, in "Computer Vision and Pattern Recognition", Boston, United States, June 2015, <https://hal.inria.fr/hal-01146533>
- [41] D. WEISSENBACHER, C. RAYMOND. *Tree-Structured Named Entities Extraction from Competing Speech Transcriptions*, in "International Conference on Application of Natural Language to Information Systems", Passau, Germany, June 2015 [DOI : 10.1007/978-3-319-19581-0_22], <https://hal.inria.fr/hal-01196808>
- [42] L. WENG, I.-H. JHUO, M. SHI, M. SUN, W.-H. CHENG, L. AMSALEG. *Supervised multi-scale locality sensitive hashing*, in "Proc. of International Conference on Multimedia Retrieval (ICMR)", Shanghai, China, June 2015, <https://hal.inria.fr/hal-01141225>

Research Reports

- [43] A. ISCEN, M. RABBAT, T. FURON. *Efficient Large-Scale Similarity Search Using Matrix Factorization*, Inria ; MCGILL UNIVERSITY, December 2015, n° 8820, <https://hal.inria.fr/hal-01238242>
- [44] M. SHI, Y. AVRITHIS, H. JÉGOU. *Early burst detection for memory-efficient image retrieval*, Inria Rennes, June 2015, <https://hal.inria.fr/hal-01166239>

Scientific Popularization

- [45] P. SÉBILLOT. *Natural language processing faced with big and potentially impaired textual data: What difference does it make?*, in "Big data : nouvelles partitions de l'information. Actes du séminaire IST Inria, octobre 2014", L. CALDERAN, P. LAURENT, H. LOWINGER, J. MILLET (editors), Information et stratégie, De Boeck, February 2015, pp. 43-60, <https://hal.archives-ouvertes.fr/hal-01056396>

Other Publications

- [46] A. BAILLY, S. MALINOWSKI, R. TAVENARD, T. GUYET, L. CHAPEL. *Dense Bag-of-Temporal-SIFT-Words for Time Series Classification*, January 2016, working paper or preprint, <https://hal.archives-ouvertes.fr/hal-01252726>
- [47] L. WENG, M. KARSAI, N. PERRA, F. MENCZER, A. FLAMMINI. *Attention on Weak Ties in Social and Communication Networks*, September 2015, working paper or preprint, <https://hal.inria.fr/hal-01203175>

References in notes

- [48] 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
- [49] 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
- [50] 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
- [51] M. ESKEVICH, G. J. JONES, S. CHEN, R. ALY, R. ORDELMAN, D. NADEEM, C. GUINAUDEAU, G. GRAVIER, P. SÉBILLOT, T. DE 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
- [52] 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
- [53] 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
- [54] 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
- [55] 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
- [56] O. DE ROOIJ, M. WORRING. *Browsing Video Along Multiple Threads*, in "IEEE Transactions on Multimedia", 2010, vol. 12, n^o 2, pp. 121–130