IRISA
UMR

Activity Report 2017

Team LINKMEDIA

Creating and Exploiting Explicit Links between Multimedia Fragments

Joint team with Inria Rennes – Bretagne Atlantique

D6 – Media and Interactions

Table of contents

Project-Team LINKMEDIA

Creation of the Project-Team: 2014 July 01

Keywords:

Computer Science and Digital Science:

A3.3.2. - Data mining
A3.3.3. - Big data analysis
A5.3.3. - Pattern recognition
A5.4.1. - Object recognition
A5.4.3. - Content retrieval
A5.7. - Audio modeling and processing
A5.8. - Natural language processing
A9.2. - Machine learning
A9.4. - Natural language processing

Other Research Topics and Application Domains:

B9. - Society and Knowledge

1. Personnel

Research Scientists

Laurent Amsaleg [CNRS, Researcher, HDR]
Yannis Avrithis [Inria, Advanced Research Position]
Vincent Claveau [CNRS, Researcher]
Teddy Furon [Inria, Researcher]
Guillaume Gravier [Team leader, CNRS, Senior Researcher, HDR]
Miaojing Shi [Inria, Starting Research Position, from Dec 2017]

Faculty Members

Ewa Kijak [Univ de Rennes I, Associate Professor]
Simon Malinowski [Univ de Rennes I, Associate Professor]
Christian Raymond [INSA Rennes, Associate Professor]
Pascale Sébillot [INSA Rennes, Professor, HDR]

Post-Doctoral Fellow

Mateusz Budnik [CNRS, from Apr 2017]

PhD Students

Rémi Bois [CNRS, until Sep 2017]
Ricardo Carlini Sperandio [Univ de Rennes I]
Clément Dalloux [CNRS]
Mikail Demirdelen [Univ de Rennes I]
Cheikh Brahim El Vaigh [Inria, from Oct 2017]
Ahmet Iscen [Inria, until Aug 2017]
Gregoire Jadi [Univ de Nantes, until Oct 2017]
Cédric Maigrot [Univ de Rennes I]
Oriane Simeoni [Inria]
Vedran Vukotić [INSA Rennes, until Sep 2017]
Hanwei Zhang [East China Normal University, Shanghai, from Sep 2017]

Technical staff
    Firas Hmida [Univ de Rennes I, Temporary Assistant Professor, from Sep 2017]
    Gabriel Sargent [CNRS, until Aug 2017]
    Ronan Sicre [CNRS, until Aug 2017]
    Arnaud Touboulic [CNRS, until Jan 2017]

Administrative Assistant
    Aurélie Patier [Univ de Rennes I]

Visiting Scientist
    Gabriel Barbosa Da Fonseca [PUC Minas, Brazil, until Jan 2017]

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 direct 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 in multimedia 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 [56], [54], [55].

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 [57], [51]. 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 [50] and was recently extended to multimedia content [58], [53], [52]. 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 entertainement 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 consequently ever growing number of innovative applications 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. Highlights of the Year

5.1. Highlights of the Year

LINKMEDIA organizer of the IEEE Workshop on Information Forensics and Security in Dec. 2017.

5.1.1. Awards

Best demo award at ACM Multimedia 2017 for collaborative work within the FUI project NexGenTV.

Best poster award at Advances in Intelligent Data Analysis.

Best paper award with colleagues of IRISA's EXPRESSION team at 24e conférence sur le Traitement Automatique des Langues Naturelles.

BEST PAPERS AWARDS:

[12]
O. AHMED, G. SARGENT, F. GARNIER, B. HUET, V. CLAVEAU, L. COUTURIER, R. TRONCY, G. GRAVIER, P. BOUZY, F. LEMÉNOREL. *NexGenTV: Providing Real-Time Insight during Political Debates in a Second Screen Application*, in "MM 2017 - 25th ACM International Conference on Multimedia", Moutain View, United States, October 2017, https://hal.archives-ouvertes.fr/hal-01635966

[28]
A. LODS, S. MALINOWSKI, R. TAVENARD, L. AMSALEG. *Learning DTW-Preserving Shapelets*, in "IDA 2017 - 16th International Symposium on Intelligent Data Analysis", London, United Kingdom, Advances in Intelligent Data Analysis XVI, springer International Publishing, October 2017, vol. 10584, pp. 198-209 [*DOI :* 10.1007/978-3-319-68765-0_17], https://hal.archives-ouvertes.fr/hal-01565207

[40]
R. QADER, G. LECORVÉ, D. LOLIVE, P. SÉBILLOT. *Automatic disfluency insertion towards spontaneous TTS : formalization and proof of concept*, in "Traitement automatique du langage naturel (TALN)", Orléans, France, June 2017, https://hal.inria.fr/hal-01532031

6. New Software and Platforms

6.1. Platforms

6.1.1. A‖GO multimedia web services
Participants: Vincent Claveau, Clément Dalloux, Guillaume Gravier [correspondent], Gabriel Sargent.

Available at http://allgo.irisa.fr, the A‖GO platform allows for the easy deployment of the technology developed in the team as web services. Based on the A‖GO infrastructure, LINKMEDIA has continued making available a number of web services related to multimedia content analysis. In 2017, we continued our effort towards the interoperability of the services available (silence detection, face detection, text-based fragmentation) and added speaker diarization and negative sentence detection services.

7. New Results

7.1. Multimedia indexing, Motif and knowledge discovery

7.1.1. Towards engineering a web-scale multimedia service: a case study using SPARK
Participant: Laurent Amsaleg.

Joint work with Gylfi Þór Guðmundsson (Univ. Reykyavik), Björn Þór Jónsson (Univ. Copenhagen) and Michael J. Franklin (UC Berkeley).

Computing power has now become abundant with multi-core machines, grids and clouds, but it remains a challenge to harness the available power and move towards gracefully handling web-scale datasets. Several researchers have used automatically distributed computing frameworks, notably Hadoop and Spark, for processing multimedia material, but mostly using small collections on small clusters. We describe the engineering process for a prototype of a (near) web-scale multimedia service using the Spark framework running on the AWS cloud service. We present experimental results using up to 43 billion SIFT feature vectors from the public YFCC 100M collection, making this the largest high-dimensional feature vector collection reported in the literature. The design of the prototype and performance results demonstrate both the flexibility and scalability of the Spark framework for implementing multimedia services.

7.1.2. On competitiveness of nearest-neighbor based music classification: a methodological critique

Participant: Laurent Amsaleg.

Joint work with Haukur Pálmasson, Björn Þór Jónsson (Univ. Copenhagen), Markus Schedl (Johannes Kepler University), Peter Knees (TU Wien).

The traditional role of nearest-neighbor classification in music classification research is that of a straw man opponent for the learning approach of the hour. Recent work in high-dimensional indexing has shown that approximate nearest-neighbor algorithms are extremely scalable, yielding results of reasonable quality from billions of high-dimensional features. With such efficient large-scale classifiers, the traditional music classification methodology of reducing both feature dimensionality and feature quantity is incorrect; instead the approximate nearest-neighbor classifier should be given an extensive data collection to work with. We present a case study, using a well-known MIR classification benchmark with well-known music features, which shows that a simple nearest-neighbor classifier performs very competitively when given ample data. In this position paper, we therefore argue that nearest-neighbor classification has been treated unfairly in the literature and may be much more competitive than previously thought [30].

7.1.3. Unsupervised part learning for visual recognition

Participants: Ronan Sicre, Yannis Avrithis, Ewa Kijak.

Joint work with Frederic Jurie (Univ. Caen).

Part-based image classification aims at representing categories by small sets of learned discriminative parts, upon which an image representation is built. Considered as a promising avenue a decade ago, this direction has been neglected since the advent of deep neural networks. In this context, the work proposed here brings two contributions: first, this work proceeds one step further compared to recent part-based models (PBM), focusing on how to learn parts without using any labeled data. Instead of learning a set of parts per class, as generally performed in the PBM literature, the proposed approach constructs a partition of a given set of images into visually similar groups, and subsequently learns a set of discriminative parts per group in a fully unsupervised fashion. This strategy opens the door to the use of PBM in new applications where labeled data are typically not available, such as instance-based image retrieval. Second , we show that despite the recent success of end-to-end models, explicit part learning can still boost classification performance. We experimentally show that our learned parts can help building efficient image representations, which outperform state-of-the art deep convolutional neural networks on both classification and retrieval tasks [32].

7.1.4. Automatic discovery of discriminative parts as a quadratic assignment problem

Participants: Ronan Sicre, Yannis Avrithis, Teddy Furon, Ewa Kijak.

Joint work with Julien Rabin and Frédéric Jurie (Univ. Caen).

Part-based image classification consists in representing categories by small sets of discriminative parts upon which a representation of the images is built. This piece of work addresses the question of how to automatically learn such parts from a set of labeled training images. We propose to cast the training of parts as a quadratic assignment problem in which optimal correspondences between image regions and parts are automatically learned. We analyze different assignment strategies and thoroughly evaluates them on two public datasets: Willow actions and MIT 67 scenes [45].

7.1.5. Learning DTW-preserving shapelets

Participants: Laurent Amsaleg, Arnaud Lods, Simon Malinowski.

Joint work with Romain Tavenard (Univ. Rennes 2).

Dynamic time warping (DTW) is one of the best similarity measures for time series, and it has extensively been used in retrieval, classification or mining applications. It is a costly measure, and applying it to numerous and/or very long times series is difficult in practice. Recently, shapelet transform (ST) proved to enable accurate supervised classification of time series. ST learns small subsequences that well discriminate classes, and transforms the time series into vectors lying in a metric space. We adopt the ST framework in a novel way: we focus on learning, without class label information, shapelets such that Euclidean distances in the ST-space approximate well the true DTW. Our approach leads to an ubiquitous representation of time series in a metric space, where any machine learning method (supervised or unsupervised) and indexing system can operate efficiently [28].

7.1.6. *Tag propagation approaches within speaking face graphs for multimodal person discovery*

Participants: Guillaume Gravier, Gabriel Sargent, Ronan Sicre.

Joint work with Gabriel Barbosa Da Fonseca, Izabela Lyon Freire, Zenilton Patrocinio Jr and Silvio Jamil F. Guimaraes (PUC Minas, Brazil)

The indexing of broadcast TV archives is a current problem in multimedia research. As the size of these databases grows continuously, meaningful features are needed to describe and connect their elements efficiently, such as the identification of speaking faces. In this context, we focused on two approaches for unsupervised person discovery. Initial tagging of speaking faces is provided by an OCR-based method, and these tags propagate through a graph model based on audiovisual relations between speaking faces. Two propagation methods are proposed, one based on random walks and the other based on a hierarchical approach. To better evaluate their performances, these methods were compared with two graph clustering baselines. We also study the impact of different modality fusions on the graph-based tag propagation scenario. From a quantitative analysis, we observed that the graph propagation techniques always outperform the baselines. Among all compared strategies, the methods based on hierarchical propagation with late fusion and random walk with score-fusion obtained the highest MAP values. Finally, even though these two methods produce highly equivalent results according to Kappa coefficient, the random walk method performs better according to a paired t-test, and the computing time for the hierarchical propagation is more than 4 times lower than the one for the random walk propagation [22].

The tag propagation results were included in a large-scale comparison of systems for person discovery in broadcast videos resulting from the MediaEval 2016 international benchmark [27].

7.2. Multimedia content description and structuring

7.2.1. *The vulnerability of learning to adversarial perturbation increases with intrinsic dimensionality*

Participant: Laurent Amsaleg.

Joint work with James Bailey, Dominique Barbe, Sarah Erfani, Michael Houle, Vinh Nguyen amd Miloš Radovanovic.

Recent research has shown that machine learning systems, including state-of-the-art deep neural networks, are vulnerable to adversarial attacks. By adding to the input object an imperceptible amount of adversarial noise, it is highly likely that the classifier can be tricked into assigning the modified object to any desired class. It has also been observed that these adversarial samples generalize well across models. A complete understanding of the nature of adversarial samples has not yet emerged. Towards this goal, we present a novel theoretical result formally linking the adversarial vulnerability of learning to the intrinsic dimensionality of the data. In particular, our investigation establishes that as the local intrinsic dimensionality (LID) increases, 1-NN classifiers become increasingly prone to being subverted. We show that in expectation, a k-nearest neighbor of a test point can be transformed into its 1-nearest neighbor by adding an amount of noise that diminishes as the LID increases. We also provide an experimental validation of the impact of LID on adversarial perturbation for both synthetic and real data, and discuss the implications of our result for general classifiers [13].

7.2.2. *Efficient temporal kernels between feature sets for time series classification*
Participant: Simon Malinowski.

Joint work with Romain Tavenard, Adeline Bailly, Louis Chapel (Univ. Rennes 2), Benjamin Bustos and Heider Sanchez (Univ. of Chile).

In the time-series classification context, the majority of the most accurate core methods are based on the bag-of-words framework, in which sets of local features are first extracted from time series. A dictionary of words is then learned and each time series is finally represented by a histogram of word occurrences. This representation induces a loss of information due to the quantization of features into words as all the time series are represented using the same fixed dictionary. In order to overcome this issue, we have designed a kernel operating directly on sets of features. Then, we have extended it to a time-compliant kernel that allows one to take into account the temporal information. We applied this kernel in the time series classification context. Proposed kernel has a quadratic complexity with the size of input feature sets, which is problematic when dealing with long time series. However, we have shown that kernel approximation techniques can be used to define a good trade-off between accuracy and complexity. We experimentally demonstrated that the proposed kernel can significantly improve the performance of time series classification algorithms based on bag-of-words [33].

7.2.3. *Tampering detection and localization in images from social networks*
Participants: Cédric Maigrot, Ewa Kijak, Vincent Claveau.

Verifying the authenticity of an image broadcast on social networks is crucial to limit the dissemination of false information. In this work, we aim to provide information about tampering localisation on such images, in order to help either the user or automatic methods to discriminate truth from falsehood. These images may have been subjected to a large number of possible forgeries, which calls for the use of generic methods. Image forensics methods based on local features have proven to be effective for the specific case of copy-move forgery. By taking advantage of the number of images available on the internet, we propose a generic system based on image retrieval, and image comparison based on local features to localise any kind of tampering in images from social networks.

Images from social media are likely to have undergone a large variety of modifications, some being malicious, and some not. The proposed approach is evaluated on three dedicated datasets containing a variety of representative tamperings in images from social media, with difficult examples. This allows an analysis of the local-features approaches behavior in this context. The method is further compared to several state-of-the-art methods and proves to be superior. Finally, we propose a classification step to discriminate malicious modifications from the non-malicious ones.

We have also built and made publicly available a large and challenging adapted database of real case images for evaluation [29].

7.2.4. *Identity documents classification as an image classification problem*
Participants: Ronan Sicre, Teddy Furon.

Joint work with Ahmad Montaser Awal and Nabil Ghanni (AriadNext).

This works studies the classification of images of identification documents. More specifically, we address the classification of documents composed of few textual information and complex background (such as identity documents). Unlike most existing systems, the proposed approach simultaneously locates the document and recognizes its class. The latter is defined by the document nature (passport, ID, etc.), emission country, version, and the visible side (main or back). This task is very challenging due to unconstrained capturing conditions, sparse textual information, and varying components that are irrelevant to the classification, e.g. photo, names, address, etc. First, a base of document models is created from reference images.

This problem is critical in various security context where proposed system must offer high performances. We address this challenge as an image classification problem, which has received a large attention from the scientific community. We show that training images are not necessary and only one reference image is enough to create a document model. Then, the query image is matched against all models in the base. Unknown documents are rejected using an estimated quality based on the extracted document. The matching process is optimized to guarantee an execution time independent from the number of document models. Once the document model is found, a more accurate matching is performed to locate the document and facilitate information extraction. Our system is evaluated on several datasets with up to 3042 real documents (representing 64 classes) achieving an accuracy of 96.6 % in [14].

In a second step, several methods are evaluated and we report results allowing a better understanding of the specificity of identification documents. We are especially interested in deep learning approaches, showing good transfer capabilities and high performances [44], [49].

7.2.5. *Sentiment analysis*
Participants: Vincent Claveau, Christian Raymond.

In the framework of the NexGenTV project, we have participated to the text-mining challenge DeFT about sentiment analysis. We have proposed methods for the identification of figurative language (irony, humor...), and for the classification of figurative and non-figurative tweets according to their polarity. For these tasks, we explore the use of three methods of increasing complexity: i) k-nearest neighbors with information retrieval based techniques, ii) boosting of decision trees, iii) recurrent neural networks [36]. It allows us to evaluate the precise interest of each of our approach and the data representation that they use: bag-of-words for the first one, n-grams for the second and word embedding for the latest.

7.3. Content-based information retrieval

7.3.1. *Efficient diffusion on region manifolds: recovering small objects with compact CNN representations*
Participants: Yannis Avrithis, Teddy Furon, Ahmet Iscen.

Joint work with Giorgos Tolias and Ondrej Chum (Technical University of Prague).

Query expansion is a popular method to improve the quality of image retrieval with both conventional and CNN representations. It has been so far limited to global image similarity. This work focuses on diffusion, a mechanism that captures the image manifold in the feature space. The diffusion is carried out on descriptors of overlapping image regions rather than on a global image descriptor like in previous approaches. An efficient off-line stage allows optional reduction in the number of stored regions. In the on-line stage, the proposed handling of unseen queries in the indexing stage removes additional computation to adjust the precomputed data. We perform diffusion through a sparse linear system solver, yielding practical query times well below one second. Experimentally, we observe a significant boost in performance of image retrieval with compact CNN descriptors on standard benchmarks, especially when the query object covers only a small part of the image. Small objects have been a common failure case of CNN-based retrieval [25].

7.3.2. *Panorama to panorama matching for location recognition*
Participants: Yannis Avrithis, Teddy Furon, Ahmet Iscen.

Joint work with Giorgos Tolias and Ondrej Chum (Technical University of Prague).

Location recognition is commonly treated as visual instance retrieval on "street view" imagery. The dataset items and queries are panoramic views, i.e., groups of images taken at a single location. This work introduces a novel panorama-to-panorama matching process, either by aggregating features of individual images in a group or by explicitly constructing a larger panorama. In either case, multiple views are used as queries. We reach near perfect location recognition on a standard benchmark with only four query views [26].

7.3.3. *Memory vectors for similarity search in high-dimensional spaces*
Participants: Teddy Furon, Ahmet Iscen.

Joint work with Vincent Gripon (IMT Atlantique), Michael Rabbat (Mc Gill University), and Hervé Jégou (Facebook AI Research).

We study an indexing architecture to store and search in a database of high-dimensional vectors from the perspective of statistical signal processing and decision theory. This architecture is composed of several memory units, each of which summarizes a fraction of the database by a single representative vector. The potential similarity of the query to one of the vectors stored in the memory unit is gauged by a simple correlation with the memory unit's representative vector. This representative optimizes the test of the following hypothesis: the query is independent from any vector in the memory unit vs. the query is a simple perturbation of one of the stored vectors. Compared to exhaustive search, our approach finds the most similar database vectors significantly faster without a noticeable reduction in search quality. Interestingly, the reduction of complexity is provably better in high-dimensional spaces. We empirically demonstrate its practical interest in a large-scale image search scenario with off-the-shelf state-of-the-art descriptors [6].

7.3.4. *Exploiting multimodality in video hyperlinking to improve target diversity*
Participants: Rémi Bois, Guillaume Gravier, Christian Raymond, Pascale Sébillot, Ronan Sicre, Vedran Vukotić.

Video hyperlinking is the process of creating links within a collection of videos to help navigation and information seeking. Starting from a given set of video segments, called anchors, a set of related segments, called targets, must be provided. In past years, a number of content-based approaches have been proposed with good results obtained by searching for target segments that are very similar to the anchor in terms of content and information. Unfortunately, relevance has been obtained to the expense of diversity. In this paper, we study multimodal approaches and their ability to provide a set of diverse yet relevant targets. We compare two recently introduced cross-modal approaches, namely, deep auto-encoders and bimodal LDA, and experimentally show that both provide significantly more diverse targets than a state-of-the-art baseline. Bimodal autoencoders offer the best trade-off between relevance and diversity, with bimodal LDA exhibiting slightly more diverse targets at a lower precision [17].

7.3.5. *Generative adversarial networks for multimodal representation learning in video hyperlinking*
Participants: Guillaume Gravier, Christian Raymond, Vedran Vukotić.

Continuous multimodal representations suitable for multimodal information retrieval are usually obtained with methods that heavily rely on multimodal autoencoders. In video hyperlinking, a task that aims at retrieving video segments, the state of the art is a variation of two interlocked networks working in opposing directions. These systems provide good multimodal embeddings and are also capable of translating from one representation space to the other. Operating on representation spaces, they lack the ability to operate in the original spaces (text or image), which makes it difficult to visualize the crossmodal function, and do not generalize well to unseen data. Recently, generative adversarial networks (GANs) have gained popularity and have been used for generating realistic synthetic data and for obtaining high-level, single-modal latent representation spaces. In this work, we evaluate the feasibility of using GANs to obtain multimodal representations. We show that GANs can be used for multimodal representation learning and that they provide multimodal representations that are superior to representations obtained with multimodal autoencoders. Additionally, we illustrate the ability of visualizing crossmodal translations that can provide human-interpretable insights on learned GAN-based video hyperlinking models [35].

7.4. Linking, navigation and analytics

7.4.1. *Providing real-time insight during political debates in a second screen application*
Participants: Vincent Claveau, Guillaume Gravier, Gabriel Sargent.

Joint work with Institut Eurecom, Wildmoka and AVISTO Telecom in the framework of the FUI project NexGenTV.

Second screen applications are becoming key for broadcasters exploiting the convergence of TV and Internet. Authoring such applications however remains costly. Within the NexGenTV project, we developed a second screen authoring application that leverages multimedia content analytics and social media monitoring. A back-office is dedicated to easy and fast content ingestion, segmentation, description and enrichment with links to entities and related content. From the back-end, broadcasters can push enriched content to front-end applications providing customers with highlights, entity and content links, overviews of social network, etc. The demonstration operates on political debates ingested during the 2017 French presidential election, enabling insights on the debates [12].

http://www.nexgentv.fr/communication/events/discover-our-new-politics-debates-live-video-edition

7.4.2. *Information extraction in clinical documents*
Participants: Clément Dalloux, Vincent Claveau.

Joint work with Claudia Moro (Pontifícia Universidade Católica do Paraná, Brazil) and Natalia Grabar (Univ. Lille)

Extracting fine-grained information from clinical texts is a keystone for numerous medical applications. For instance, in clinical trial protocols eligibility criteria are expressed through texts in an unstructured way. This year, we have developed an annotated corpus of clinical trials and made it available to the community. Based on this corpus, we proposed automatic methods to extract numerical information [20] and to handle the variation of the units used [43]. In such medical applications, detecting negation, uncertainty, and the scope on which they apply is important. Thus, we have also developed an annotated corpus, made it available to the community, and we have proposed an automatic tool based on recurrent neural networks [37], [41] and made it available as a web service.

7.4.3. *Semi-supervision for information extraction*
Participants: Vincent Claveau, Ewa Kijak.

Many NLP problems are tackled as supervised machine learning tasks. 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. This year, we examined and proposed such selection strategies in the specific case of conditional random fields (CRF) which are largely used in NLP. On the one hand, we proposed a simple method to correct a bias of some state-of-the-art selection techniques. On the other hand, we built an original approach to select the examples, based on the respect of proportions in the datasets. These contributions were validated over a large range of experiments implying several datasets and tasks, including named entity recognition, chunking, phonetization, word sense disambiguation [19].

7.4.4. *Linking multimedia content for efficient news browsing via explorable news graphs*
Participants: Rémi Bois, Guillaume Gravier, Pascale Sébillot.

Joint work with Maxime Robert, Éric Jamet (Univ. Rennes 2) and Emmanuel Morin (Univ. Nantes) in the framework of the CominLabs project Linking Media in Acceptable Hypergraphs.

As the amount of news information available online grows, media professionals are in need of advanced tools to explore the information surrounding specific events before writing their own piece of news, e.g., adding context and insight. While many tools exist to extract information from large datasets, they do not offer an easy way to gain insight from a news collection by browsing, going from article to article and viewing unaltered original content. Such browsing tools require the creation of rich underlying structures such as graph representations. These representations can be further enhanced by typing links that connect nodes, in order to inform the user on the nature of their relation. We propose an efficient way to generate links between news items in order to obtain an easily navigable graph, and enrich this graph by automatically typing created

links. User evaluations are conducted on real-world data in order to assess for the interest of both the graph representation and link typing in a press reviewing task, showing a significant improvement compared to classical search engines [15], [16].

7.4.5. *Multimodal detection of fake news*
Participants: Vincent Claveau, Cédric Maigrot, Ewa Kijak.

Social networks make it possible to share rapidly and massively information, including fake news, hoaxes or rumors. Following our previous work in the frame of the Verification Multimedia Use task of Mediaeval 2016, we have explored the use of multimodal clues to detect fake news in social networks [38]. This year, we have studied the interest of combining and merging many approaches developed by the MediaEval participants in order to evaluate the predictive power of each modality. We have proposed several fusion strategies making the most of their potential complementarity [39].

7.5. Miscellaneous

In parallel with mainstream research activities, LINKMEDIA has a number of contribtions in other domains based on the expertise of the team members.

7.5.1. *One-step time-dependent future video frame prediction with a convolutional encoder-decoder neural network*
Participants: Guillaume Gravier, Christian Raymond, Vedran Vukotić.

Joint work with Silvia-Laura Pintea and Jan Van Gemert (TU Delft, The Netherlands).

There is an inherent need for autonomous cars, drones, and other robots to have a notion of how their environment behaves and to anticipate changes in the near future. In this work, we focus on anticipating future appearance given the current frame of a video. Existing work focuses on either predicting the future appearance as the next frame of a video, or predicting future motion as optical flow or motion trajectories starting from a single video frame. This work stretches the ability of convolutional neural networks (CNNs) to predict an anticipation of appearance at an arbitrarily given future time, not necessarily the next video frame. We condition our predicted future appearance on a continuous time variable that allows us to anticipate future frames at a given temporal distance, directly from the input video frame. We show that CNNs can learn an intrinsic representation of typical appearance changes over time and successfully generate realistic predictions at a deliberate time difference in the near future [34].

7.5.2. *About zero bit watermarking error exponents*
Participant: Teddy Furon.

This work aims to motivate more research works on the design of zero-bit watermarking schemes by showing an upper bound of the performances that known solutions failed to reach. To this end, an upper bound of error exponent characteristic is derived by translating Costa's rationale to zero-bit watermarking with side information. Three schemes are then considered: the dual-cone detection region originally proposed by Cox et al. and improved in Merhav et al. papers, ISS (Improved Spread Spectrum), and ZATT (Zero Attraction). It turns out that in certain conditions the latter performs better than the first one, which questions the optimality claimed Merhav et al. Nevertheless, the main conclusion is that these schemes are in general far away from the upper bound in the region of practical interest [23].

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

Quai des Apps: one-shot 2-day contract for scientific counseling on visual image retrieval.

9. Partnerships and Cooperations

9.1. Regional Initiatives

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

Participants: Rémi Bois, Vincent Claveau, Guillaume Gravier, Pascale Sébillot, Arnaud Touboulic.

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., a 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 ergonomy psychology.

9.1.2. *CominLabs Project BigCLIN*

Participants: Vincent Claveau, Ewa Kijak, Clément Dalloux.

Duration: 3 years, started in September 2016
Partners: STL-CNRS, Inserm/CHU Rennes, Inria
URL: http://www.bigclin.cominlabs.ueb.eu

Data collected or produced during clinical care process can be exploited at different levels and across different domains. Yet, a well-known challenge for secondary use of health big data is that much of detailed patient information is embedded in narrative text, mostly stored as unstructured data. The project proposes to address the essential needs when reusing unstructured clinical data at a large scale. We propose to develop new clinical records representation relying on fine-grained semantic annotation thanks to new NLP tools dedicated to French clinical narratives. To efficiently map this added semantic information to existing structured data for further analysis at big scale, the project also addresses distributed systems issues: scalability, management of uncertain data and privacy, stream processing at runtime, etc.

9.2. National Initiatives

9.2.1. *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.

9.2.2. FUI 19 NexGenTV

Participants: Vincent Claveau, Guillaume Gravier, Ewa Kijak, Gabriel Sargent, Ronan Sicre.

Duration: 2.5 years, started in May 2015
Partners: Eurecom, Avisto Telecom, Wildmoka, Envivio-Ericsson

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 domaine, e.g., sports, news.

9.2.3. Inria Project Lab Knowledge-driven data and content collaborative analytics (iCODA)

Participants: Laurent Amsaleg, Vincent Claveau, Cheikh Brahim El Vaigh, Guillaume Gravier, Pascale Sébillot.

Duration: 4.5 years, started in April 2017
Partners: Inria project-teams Linkmedia, CEDAR, GraphIK and ILDA, with Ouest France, Le Monde and AFP

One of today's major issues in data science is the design of algorithms that allow analysts to efficiently infer useful information and knowledge by collaboratively inspecting heterogeneous information sources, from structured data to unstructured content. Taking data journalism as an emblematic use-case, the goal of the project is to develop the scientific and technological foundations for knowledge- mediated user-in-the-loop collaborative data analytics on heterogenous information sources, and to demonstrate the effectiveness of the approach in realistic, high-visibility use-cases. The project stands at the crossroad of multiple research fields—content analysis, data management, knowledge representation, visualization—that span multiple Inria themes, and counts on a club of major press partners to define usage scenarios, provide data and demonstrate achievements.

9.3. European Initiatives

9.3.1. CHIST ERA ID_IOT

Participant: Teddy Furon.

Duration: 3 years, started in Oct. 2016
Partners: Eindhoven Univ. of Technology, Univ. of Geneva

The IoT will contain a huge number of devices and objects that have very low or nonexistent processing and communication resources, coupled to a small number of high-power devices. The weakest devices, which are most ubiquitous, will not be able to authenticate themselves using cryptographic methods. This project addresses these issues using physical unclonable functions (PUFs). PUFs, and especially quantum readout PUFs, are ideally suited to the IoT setting because they allow for the authentication and identification of physical objects without requiring any crypto or storage of secret information.

Furthermore, we foresee that back-end systems will not be able to provide security and privacy via cryptographic primitives due to the sheer number of IoT devices. Our plan is to address these problems using privacy-preserving database structures and algorithms with good scaling behaviour. Approximate nearest neighbour (ANN) search algorithms, which have remarkably good scaling behaviour, have recently become highly efficient, but do not yet have the right security properties and have not yet been applied to PUF data. Summarised in a nutshell, the project aims to improve the theory and practice of technologies such as PUFs and ANN search in the context of generic IoT authentication and identification scenarios.

9.3.2. Collaborations with Major European Organizations

Big Data Value Association (BDVA): LINKMEDIA is a co-founder and co-leader of the media group (TF7) within BDVA

9.4. International Initiatives

9.4.1. Inria International Partners

9.4.1.1. Informal International Partners

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

9.4.2. Participation in Other International Programs

- CNRS – CONFAP FIGTEM
    - Title: Fine-grained text-mining for clinical trials
    - International Partner (Institution - Laboratory - Researcher): Pontifícia Universidade Católica do Paraná - Health Informatics dept, Claudia Moro

        FIGTEM aims at developing natural language processing methods, including information extraction and indexing, dedicated to the clinical trial domain. The goal is to populate a formal representation of patients (via their electronic patient records) and clinical trial data in different languages (French, English, Portuguese).
    - Jan. 2016 – Dec. 2018

9.5. International Research Visitors

9.5.1. Visits of International Scientists

Giorgos Tolias

Date: Sept. 2017 (1 week)

Institution: Czech Technical University, Czech Republic

Vincent Oria

Date: July 2017 (2 weeks)

Institution: New Jersey Institute of Technology, Newark, USA

Michael Houle

Date: July 2017 (2 weeks)

Institution: National Institute of Informatics, Tokyo, Japan

9.5.1.1. Internships

Gabriel B. de Fonseca

Date: Nov. 2016 - Jan. 2017

Institution: PUC Minas, Brazil

9.5.2. Visits to International Teams

Laurent Amsaleg

Date: Oct. 2017 (2 days)

Institution: New Jersey Institute of Technology, Newark, USA

Laurent Amsaleg

Date: May. 2017 (1 week)

Institution: East China Normal University, Shanghai, PRC

Clément Dalloux

Date: Nov.-Dec 2017 (1 month)

Institution: Pontifícia Universidade Católica do Paraná, Brazil

Guillaume Gravier

Date: Nov. 2017 (1 week)

Institution: Universidad de Chile, Santiago, Chile

Guillaume Gravier

Date: May. 2017 (1 week)

Institution: East China Normal University, Shanghai, PRC

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Organisation

10.1.1.1. General Chair, Scientific Chair

Laurent Amsaleg was appointed general co-chair of ACM Intl. Conf. on Multimedia 2019.

Laurent Amsaleg was general chair of Intl. Conf. on Multimedia Modeling 2017.

Laurent Amsaleg and Vincent Claveau were general co-chairs of the Workshop on Computational Journalism collocated with EGC 2017.

Teddy Furon was general co-chair of IEEE Intl. Workshop on Information Forensics and Security 2017.

Teddy Furon was a co-chair of the GdR-ISIS national workshop "Comment concilier Big Data, identification des personnes, traçabilité des contenus et respect de la vie privée ?", 2017.

Ewa Kijak and Vincent Claveau organized and chaired a special session on Social Networks and User-Generated Content Verification at the IEEE Intl. Workshop on Information Forensics and Security.

10.1.1.2. Member of the Organizing Committees

Guillaume Gravier was special session chair at Intl. Conf. on Multimedia Modeling 2017.

Guillaume Gravier was nominated special session co-chair of IEEE Intl. Workshop on Content-Based Multimedia Indexing 2018.

10.1.2. Scientific Events Selection

10.1.2.1. Chair of Conference Program Committees

Yannis Avrithis was area chair of the ACM Intl. Conf. on Multimedia 2017.

Yannis Avrithis was area chair of the European Signal Processing Conf 2017.

Guillaume Gravier was apppointed area chair of ACM Intl. Conf. on Multimedia 2019.

10.1.2.2. Member of Conference Program Committees

Laurent Amsaleg was a PC member of: Base de Données AvancéesA; IEEE Intl. Workshop on Content-Based Multimedia Indexing, IEEE Intl. Conf. on Multimedia and Exhibition; Intl. Conf. on Similarity Search and Applications; ACM Intl. Conf. on Multimedia Retrieval.

Yannis Avrithis was a PC member of: IEEE Conf. on Computer Vision and Pattern Recognition; Intl. Conf. on Computer Vision; ACM Intl. Conf. on Multimedia; European Signal Processing Conf.; IEEE Intl. Conf. on Multimedia and Expo; IEEE Intl. Conf. on Acoustics, Speech, and Signal Processing.

Vincent Claveau was a PC member of: ACL demonstration track; EMNLP demonstration track; International Semantic Web Conference; Intl. Conf. on Multimedia Modeling; IEEE Intl. Workshop on Information Forensics and Security; Conférence en Recherche d'Information et Applications; TextMine workshop.

Teddy Furon was a PC member of ACM Information Hiding and Multimedia Security.

Guillaume Gravier was a PC member of: ACM Intl. Conf. on Multimedia; IEEE Intl. Conf. on Multimedia and Exhibition; ACM Intl. Conf. on Multimedia Retrieval; Annual Conf. of the Intl. Speech Communication Association; European Conf. on Information Retrieval; IEEE Intl. Conf. on Acoustics, Speech and Signal Processing; IEEE Intl. Symposium on Multimedia; Multimedia Modeling Conf.

Ewa Kijak was a PC member of: ACM Intl. Conf. on Multimedia; ACM Intl. Conf. on Multimedia Retrieval; Intl. Workshop on Content-Based Multimedia Indexing; IEEE Intl. Workshop on Information Forensics and Security.

Christian Raymond was a PC member of: Annual Conf. of the Intl. Speech Communication Association; IEEE Intl. Conf. on Acoustics Speech and Signal Processing; IEEE Intl. Conf. on Machine Learning And Applications; Conf. en Traitement Automatique des Langues Naturelles.

Pascale Sébillot was a PC member of: European Conf. on Information Retrieval; Annual Meeting of the Association for Computational Linguistics; Intl. Joint Conf. on Artificial Intelligence; Intl. Conf. on Multimedia Modeling; Conf. Traitement Automatique des Langues Naturelles; Traitement Automatique du Langage et Analyse de Documents.

10.1.2.3. Reviewer

Ewa Kijak reviewed for ACM Intl. Conf. on Multimedia.

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

Laurent Amsaleg was guest editor for the special issue of the Multimedia tools and applications Journal, best papers from the MMM2017 conference.

Vincent Claveau is editor of the journal Recherche d'Information, Document, Web Sémantique

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

Guillaume Gravier is associate editor of IEEE Trans. on Multimedia.

Guillaume Gravier was editor of the Working Notes Proc. of the MediaEval Multimedia Benchmark.

Christian Raymond is member of the editorial board of the electronic Journal Discours.

Pascale Sébillot is editor of the Journal Traitement Automatique des Langues.

Pascale Sébillot is member of the editorial board of the Journal Traitement Automatique des Langues.

10.1.3.2. Reviewer - Reviewing Activities

Laurent Amsaleg was a reviewer for: IEEE Trans. on Circuits and Systems for Video Technology; IEEE Trans. on Multimedia, Pattern Recognition Letters; IEEE Trans. on Information Forensics & Security; IEEE Trans. on Image Processing; IEEE Trans. on Information Systems; IEEE Trans. on Knowledge and Data Engineering.

Yannis Avrithis was a reviewer for Multimedia Tools and Applications, Elsevier Neurocomputing.

Vincent Claveau reviewed for reviewed for Multimedia Tools and Applications, Elsevier Information Sciences, Traitement Automatique des Langues.

Teddy Furon reviewed for IEEE Trans. on Information Forensics and Security, IEEE Trans. on Image Processing, and IEEE Trans. on Signal Processing.

Guillaume Gravier reviewed for Information Science, Multimedia Tools and Applications, IEEE Trans. on Multimedia.

Christian Raymond was a reviewer for Multimedia Tools and Applications.

Pascale Sébillot reviewed for Traitement Automatique des Langues.

10.1.4. Invited Talks

Laurent Amsaleg gave an invited talk at East China Normal University, Shanghai.

Yannis Avrithis gave an invited talk at Deep Learning in Vision, Language, and Multimedia Workshop on September 25 in Rennes.

Vincent Claveau gave an invited talk at Intelligence artificielle et Recherche d'Information : Journée commune AFIA - ARIA.

Guillaume Gravier gave an invited talk at Katholieke Universiteit Leuven.

Guillaume Gravier gave an invited talk at Universidad de Chile, Santiago de Chile.

Guillaume Gravier gave an invited talk at East China Normal University, Shanghai.

Pascale Sébillot gave an invited tutorial at symposium on Propriété intellectuelle et données dans l'environnement numérique, Rennes.

10.1.5. Leadership within the Scientific Community

Laurent Amsaleg is a member of the Steering Committee of SISAP for the 2016–2020 term.

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

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 was president of the Scientific Evaluation Committee of the National Research Agency for the theme 'Knowledge, data, content, big data, HPC, simulation' up to July 2017.

Since September 2017, Guillaume Gravier is president of the Scientific Evaluation Committee of the National Research Agency for the theme 'Knowledge, data, content, big data - AI'.

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

10.1.6. Scientific Expertise

Laurent Amsaleg was an expert evaluator for the French National Research Agency (ANR).

Yannis Avrithis was an expert evaluator for the French National Research Agency (ANR).

Vincent Claveau was an expert for was an expert for the Idex UCA-Jedi

Teddy Furon is scientific adviser for the company LAMARK.

10.1.7. Research Administration

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 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 27th section (computer science).

Pascale Sébillot is the director of the Computer Science Laboratory, INSA Rennes.

Pascale Sébillot is the deputy director of the Scientific Advisory Committee of IRISA UMR 6074.

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

Pascale Sébillot is a member of the board of the MathSTIC doctoral school.

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

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

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

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

Licence: Guillaume Gravier, Probability and statistics, 16h, L3, INSA Rennes, France

Licence: Guillaume Gravier, Natural Language Processing, 12h, L3 & M1, INSA Rennes, France

Master: Laurent Amsaleg, Multidimensional indexing, 13h, M2, University Rennes 1, France

Master: Yannis Avrithis, Deep Learning for Vision, 30h, M2, Univ. Rennes 1, France

Master: Vincent Claveau, Data-Based Knowledge Acquisition: Symbolic Methods, 20h, L3, INSA Rennes, France

Master: Vincent Claveau, Text Mining, 18h, 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

Master: Teddy Furon, Rare events, 20h, M2, Insa Rennes, France

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

Master: Ewa Kijak, Image processing, 67h, M1, ESIR, France

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

Master: Ewa Kijak, Supervised machine learning, 45h, M1, ESIR, France

Master: Ewa Kijak, Image indexing, 17h, M2, University Rennes 1, France

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

Master: Ewa Kijak, Computer vision, 22h, M2, ESIR, France

Master: Ewa Kijak, Image and text mining, 12h, M2, ENSAI, France

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

Master: Simon Malinowski, Supervised Learning, 24h, M2, Univ. Rennes 1

Master: Christian Raymond, Dialogue, 4H, M2, University Rennes 1, France

Master: Christian Raymond, Dialogue, 5H, M1, INSA Rennes, France

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

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

Master: Pascale Sébillot, Natural Language Programming, 6h, M1, INSA Rennes, France

10.2.2. Supervision

PhD: Rémi Bois, Navigable directed multimedia hypergraphs: construction and exploitation, defended December 2017, Guillaume Gravier and Pascale Sébillot

PhD: Ahmet Iscen, Continuous memories for representing sets of vectors and image collections, defended September 2017, Teddy Furon

PhD: Raheel Kareem Qader, Phonology modeling for emotional speech synthesis, defended March 2017, Gwénolé Lecorvé and Pascale Sébillot (with EXPRESSION, IRISA team)

PhD: Vedran Vukotić, Deep neural architectures for automatic representation learning from multimedia multimodal data, defended September 2017, Guillaume Gravier and Christian Raymond

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

PhD in progress: Clément Dalloux, Clinical text mining and indexing, started Dec. 2016, Vincent Claveau, Natalia Grabar (STL, Lille), Olivier Dameron (with DYLISS project-team)

PhD in progress: Mikail Demirdelen, User-adapted multi-document multimedia synthesis, started Oct. 2016, Guillaume Gravier and Pascale Sébillot

PhD in progress: Cheikh Brahim El Vaigh, Incremental content to data linking leveraging ontological knowledge in data journalism, started Oct. 2017, Guillaume Gravier, Pascale Sébillot and François Goasdoué (with CEDAR, Inria team)

PhD in progress: Mathieu Laroze, Active learning on adaptive representations for object detection in high-resolution imaging, started June 2016, Romain Dambreville, Chloe Friguet, Ewa Kijak and Sebastien Lefevre (with OBELIX, 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: Antoine Perquin, Universal speech synthesis through embeddings of massive heterogeneous data, started Oct. 2017, Laurent Amsaleg, Gwénolé Lecorvé, Damien Lolive (with EXPRESSION, IRISA team)

PhD in progress: Oriane Simeoni, Invariance and supervision in visual learning, started Oct. 2016, Yannis Avrithis and Guillaume Gravier

PhD in progress: Hanwei Zhang, Deep learning in adversarial contexts, started Oct. 2017, Laurent Amsaleg, Teddy Furon, Ewa Kijak

10.2.3. Juries

Guillaume Gravier

PhD, reviewer, Aparna Nurani Venkitasubramanian, Katholieke Universiteit Leuven

PhD, reviewer, Heider Sanchez, Universidad de Chile

PhD, reviewer, Mateusz Budnik, Univ. Grenoble Alpes

PhD, reviewer, Nhi Tran, CNAM

PhD, president, Gaël Le Lan, Univ. du Maine

Ewa Kijak

PhD, member, Yanwei Cui, Université Bretagne Sud

Pascale Sébillot

PhD, reviewer, Dialekti Valsamou, Université Paris Sud

PhD, reviewer, Anaïs Ollagnier, Aix-Marseille Université

PhD, president, Joseph Lark, Université de Nantes

PhD, member, Yoann Dupont, Université Sorbonne Nouvelle - Paris 3

10.3. Popularization

Vincent Claveau

participated to the collective book "Les Big Data à découvert", CNRS éditions

got an interview published in CNRS - Le journal about detecting fake news and image forgeries (joint work with C. Maigrot and E. Kijak).

11. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] A. ISCEN. *Continuous memories for representing sets of vectors and image collections*, Université Rennes 1, September 2017, https://tel.archives-ouvertes.fr/tel-01661318

[2] A. ISCEN. *Continuous memories for representing sets of vectors and image collections*, Université Rennes 1, September 2017, https://tel.archives-ouvertes.fr/tel-01661319

[3] V. VUKOTIC. *Deep neural architectures for automatic representation learning from multimedia multimodal data*, INSA Rennes, September 2017, https://hal.inria.fr/tel-01629669

[4] V. VUKOTIC. *Deep Neural Architectures for Automatic Representation Learning from Multimedia Multimodal Data*, INSA de Rennes, September 2017, https://tel.archives-ouvertes.fr/tel-01662594

Articles in International Peer-Reviewed Journals

[5] L. AMSALEG, V. CLAVEAU. *Notes about the Computational Journalism workshop 2017*, in "Recherche d'Information, Document et Web Sémantique", August 2017, vol. 1, n⁰ 1, 6 p. , https://hal.archives-ouvertes.fr/hal-01643444

[6] A. ISCEN, T. FURON, V. GRIPON, M. RABBAT, H. JÉGOU. *Memory vectors for similarity search in high-dimensional spaces*, in "IEEE Transactions on Big Data", 2017 [*DOI : 10.1109/TBDATA.2017.2677964*], https://hal.inria.fr/hal-01481220

[7] M. LARSON, M. SOLEYMANI, G. GRAVIER, B. IONESCU, G. J. F. JONES. *The Benchmarking Initiative for Multimedia Evaluation: MediaEval 2016*, in "IEEE MultiMedia", January 2017, vol. 24, n⁰ 1, pp. 93 - 96 [*DOI : 10.1109/MMUL.2017.9*], https://hal.archives-ouvertes.fr/hal-01635965

[8] C. MAIGROT, S. BRINGAY, J. AZÉ. *Concept drift vs suicide: How one can help prevent the other?*, in "International Journal of Computational Linguistics and Applications", October 2017, vol. 8, n⁰ 1, https://hal.archives-ouvertes.fr/hal-01617870

[9] B. QU, F. VALLET, J. CARRIVE, G. GRAVIER. *Content-based unsupervised segmentation of recurrent TV programs using grammatical inference*, in "Multimedia Tools and Applications", 2017 [*DOI : 10.1007/s11042-017-4816-5*], https://hal.archives-ouvertes.fr/hal-01522688

[10] G. SARGENT, F. BIMBOT, E. VINCENT. *Estimating the structural segmentation of popular music pieces under regularity constraints*, in "IEEE/ACM Transactions on Audio, Speech, and Language Processing", 2017, https://hal.inria.fr/hal-01403210

Articles in Non Peer-Reviewed Journals

[11] Y. KOMPATSIARIS, G. GRAVIER. *Special Issue on Content Based Multimedia Indexing*, in "Multimedia Tools and Applications", 2017 [*DOI :* 10.1007/S11042-017-5106-Y], https://hal.archives-ouvertes.fr/hal-01578925

International Conferences with Proceedings

[12] *Best Paper*
O. AHMED, G. SARGENT, F. GARNIER, B. HUET, V. CLAVEAU, L. COUTURIER, R. TRONCY, G. GRAVIER, P. BOUZY, F. LEMÉNOREL. *NexGenTV: Providing Real-Time Insight during Political Debates in a Second Screen Application*, in "MM 2017 - 25th ACM International Conference on Multimedia", Moutain View, United States, October 2017, https://hal.archives-ouvertes.fr/hal-01635966.

[13] L. AMSALEG, J. E. BAILEY, D. BARBE, S. ERFANI, M. E. HOULE, V. NGUYEN, M. RADOVANOVIC. *The Vulnerability of Learning to Adversarial Perturbation Increases with Intrinsic Dimensionality*, in "WIFS 2017 - 9th IEEE International Workshop on Information Forensics and Security", Rennes, France, Proceedings of the 9th IEEE Workshop on Information Forensics and Security, December 2017, https://hal.archives-ouvertes.fr/hal-01599355

[14] A.-M. A. AWAL, N. GHANMI, R. SICRE, T. FURON. *Complex Document Classification and Localization Application on Identity Document Images*, in "ICDAR 2017 - The 14th IAPR International Conference on Document Analysis and Recognition", Kyoto, Japan, November 2017, pp. 1-6, https://hal.inria.fr/hal-01660504

[15] R. BOIS, G. GRAVIER, E. JAMET, E. MORIN, M. ROBERT, P. SÉBILLOT. *Linking multimedia content for efficient news browsing*, in "ICMR'17 - ACM International Conference on Multimedia Retrieval", Bucharest, Romania, June 2017, https://hal.archives-ouvertes.fr/hal-01522413

[16] R. BOIS, G. GRAVIER, E. JAMET, M. ROBERT, E. MORIN, P. SÉBILLOT. *Language-based Construction of Explorable News Graphs for Journalists*, in "Empirical Methods in Natural Language Processing - Workshop on Natural Language Processing meets Journalism", Copenhague, Denmark, 2017, https://hal.archives-ouvertes.fr/hal-01578932

[17] R. BOIS, V. VUKOTIC, A.-R. SIMON, R. SICRE, C. RAYMOND, P. SÉBILLOT, G. GRAVIER. *Exploiting Multimodality in Video Hyperlinking to Improve Target Diversity*, in "MMM2017 - International Conference on Multimedia Modeling", Reykyavik, Iceland, L. AMSALEG, G. Þ. GUÐMUNDSSON, C. GURRIN, S. SATOH (editors), Lecture Notes in Computer Science, Springer, January 2017, vol. 10133 [*DOI :* 10.1007/978-3-319-51814-5_16], https://hal.archives-ouvertes.fr/hal-01498130

[18] V. CLAVEAU. *Clustering multi-relationnal TV data by diverting supervised ILP*, in "ILP 2017 - 27th International Conference on Inductive Logic Programming", Orléans, France, Proceedings of the 27th International Conference on Inductive Logic Programming, September 2017, pp. 1-6, https://hal.archives-ouvertes.fr/hal-01643977

[19] V. CLAVEAU, E. KIJAK. *Strategies to select examples for Active Learning with Conditional Random Fields*, in "CICLing 2017 - 18th International Conference on Computational Linguistics and Intelligent Text Processing", Budapest, Hungary, April 2017, pp. 1-14, https://hal.archives-ouvertes.fr/hal-01621338

[20] V. CLAVEAU, L. E. S. OLIVEIRA, G. BOUZILLÉ, M. CUGGIA, C. M. CABRAL MORO, N. GRABAR. *Numerical eligibility criteria in clinical protocols: annotation, automatic detection and interpretation*, in "AIME 2017 - 16th Conference in Artificial Intelligence in Medecine", Vienne, Austria, LNCS, Springer, June 2017, vol. 10259, pp. 203-208 [*DOI :* 10.1007/978-3-319-59758-4_22], https://hal.archives-ouvertes.fr/hal-01618930

[21] M. DINARELLI, V. VUKOTIC, C. RAYMOND. *Label-dependency coding in Simple Recurrent Networks for Spoken Language Understanding*, in "Interspeech", Stockholm, Sweden, August 2017, https://hal.inria.fr/hal-01553830

[22] G. B. D. FONSECA, G. SARGENT, I. L. FREIRE, R. SICRE, Z. J. PATROCINIO, G. GRAVIER, S. J. F. GUIMARÃES. *Tag Propagation Approaches within Speaking Face Graphs for Multimodal Person Discovery*, in "International workshop on Content-Based Multimedia Indexing (CBMI)", Firenze, Italy, Proceedings of the 15th international workshop on Content-Based Multimedia Indexing, June 2017, https://hal.archives-ouvertes.fr/hal-01551648

[23] T. FURON. *About zero bit watermarking error exponents*, in "ICASSP2017 - IEEE International Conference on Acoustics, Speech and Signal Processing", New Orleans, United States, IEEE, March 2017, https://hal.inria.fr/hal-01512705

[24] G. Þ. GUÐMUNDSSON, L. AMSALEG, B. Þ. JÓNSSON, M. J. FRANKLIN. *Towards Engineering aWeb-Scale Multimedia Service: A Case Study Using Spark*, in "Multimedia Systems Conference", Taipei, Taiwan, MMSys'17 Proceedings of the 8th ACM on Multimedia Systems Conference, acm, June 2017 [*DOI :* 10.1145/3083187.3083200], https://hal.archives-ouvertes.fr/hal-01565196

[25] A. ISCEN, G. TOLIAS, Y. AVRITHIS, T. FURON, O. CHUM. *Efficient Diffusion on Region Manifolds: Recovering Small Objects with Compact CNN Representations*, in "2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)", Honolulu, United States, July 2017, https://hal.inria.fr/hal-01505470

[26] A. ISCEN, G. TOLIAS, Y. AVRITHIS, T. FURON, O. CHUM. *Panorama to panorama matching for location recognition*, in "ACM International Conference on Multimedia Retrieval (ICMR) 2017", Bucharest, Romania, June 2017 [*DOI :* 10.1145/3078971.3079033], https://hal.inria.fr/hal-01511793

[27] N. LE, H. BREDIN, G. SARGENT, M. INDIA, P. LOPEZ-OTERO, C. BARRAS, C. GUINAUDEAU, G. GRAVIER, G. BARBOSA DA FONSECA, I. LYON FREIRE, Z. PATROCÍNIO JR, S. J. F. GUIMARÃES, G. MARTÍ, J. R. MORROS, J. HERNANDO, L. DOCIO-FERNANDEZ, C. GARCIA-MATEO, S. MEIGNIER, J.-M. ODOBEZ. *Towards large scale multimedia indexing: A case study on person discovery in broadcast news*, in "Content-Based Multimedia Indexing CBMI", Firenze, Italy, Proceedings of the 15th international workshop on Content-Based Multimedia Indexing, June 2017 [*DOI :* 10.1145/3095713.3095732], https://hal.archives-ouvertes.fr/hal-01551690

[28] *Best Paper*
A. LODS, S. MALINOWSKI, R. TAVENARD, L. AMSALEG. *Learning DTW-Preserving Shapelets*, in "IDA 2017 - 16th International Symposium on Intelligent Data Analysis", London, United Kingdom, Advances in Intelligent Data Analysis XVI, springer International Publishing, October 2017, vol. 10584, pp. 198-209 [*DOI :* 10.1007/978-3-319-68765-0_17], https://hal.archives-ouvertes.fr/hal-01565207.

[29] C. MAIGROT, E. KIJAK, R. SICRE, V. CLAVEAU. *Tampering detection and localization in images from social networks: A CBIR approach*, in "ICIAP 2017 - International Conference on Image Analysis and Processing", Catane, Italy, September 2017, pp. 1-11, https://hal.inria.fr/hal-01623105

[30] H. PÁLMASSON, B. Þ. JÓNSSON, L. AMSALEG, M. SCHEDL, P. KNEES. *On Competitiveness of Nearest-Neighbor Based Music Classification: A Methodological Critique*, in "10th International Conference on Similarity Search and Applications", munich, Germany, 10th International Conference on Similarity Search and Applications, October 2017, https://hal.archives-ouvertes.fr/hal-01565217

[31] R. QADER, G. LECORVÉ, D. LOLIVE, M. TAHON, P. SÉBILLOT. *Statistical Pronunciation Adaptation for Spontaneous Speech Synthesis*, in "Text, Speech and Dialogue (TSD)", Prague, Czech Republic, August 2017, https://hal.inria.fr/hal-01532035

[32] R. SICRE, Y. AVRITHIS, E. KIJAK, F. JURIE. *Unsupervised part learning for visual recognition*, in "CVPR 2017 - IEEE Conference on Computer Vision and Pattern Recognition", Honolulu, United States, July 2017, https://arxiv.org/abs/1704.03755 , https://hal.archives-ouvertes.fr/hal-01507379

[33] R. TAVENARD, S. MALINOWSKI, L. CHAPEL, A. BAILLY, H. SANCHEZ, B. BUSTOS. *Efficient Temporal Kernels between Feature Sets for Time Series Classification*, in "European Conference on Machine Learning and Principles and Practice of Knowledge Discovery", Skopje, Macedonia, September 2017, https://halshs.archives-ouvertes.fr/halshs-01561461

[34] V. VUKOTIC, S.-L. PINTEA, C. RAYMOND, G. GRAVIER, J. C. VAN GEMERT. *One-Step Time-Dependent Future Video Frame Prediction with a Convolutional Encoder-Decoder Neural Network*, in "International Conference of Image Analysis and Processing (ICIAP)", Catania, Italy, Proceedings of the 19th International Conference of Image Analysis and Processing, September 2017, https://hal.inria.fr/hal-01567742

[35] V. VUKOTIC, C. RAYMOND, G. GRAVIER. *Generative Adversarial Networks for Multimodal Representation Learning in Video Hyperlinking*, in "ACM International Conference on Multimedia Retrieval (ICMR) 2017", Bucharest, Romania, Proceedings of the 2017 ACM on International Conference on Multimedia Retrieval, ACM, June 2017 [*DOI :* 10.1145/3078971.3079038], https://hal.inria.fr/hal-01522419

National Conferences with Proceedings

[36] V. CLAVEAU, C. RAYMOND. *IRISA at DeFT2017 : classification systems of increasing complexity* , in "DeFT 2017 - Défi Fouille de texte", Orléans, France, Actes de l'Atelier Défi Fouille de Texte, DeFT, June 2017, pp. 1-10, https://hal.archives-ouvertes.fr/hal-01643993

[37] C. DALLOUX, V. CLAVEAU, N. GRABAR. *Détection de la négation : corpus français et apprentissage supervisé*, in "SIIM 2017 - Symposium sur l'Ingénierie de l'Information Médicale", Toulouse, France, November 2017, pp. 1-8, https://hal.archives-ouvertes.fr/hal-01659637

[38] C. MAIGROT, V. CLAVEAU, E. KIJAK. *Détection de fausses informations dans les réseaux sociaux : vers des approches multi-modales*, in "Extraction et Gestion des Connaissances (EGC)", Grenoble, France, January 2017, https://hal.inria.fr/hal-01548156

[39] C. MAIGROT, E. KIJAK, V. CLAVEAU. *Détection de fausses informations dans les réseaux sociaux : l'utilité des fusions de connaissances*, in "COnférence Recherche d'Information et Applications", Marseille, France,

Actes de la conférence CORIA 2017, March 2017, pp. 107-122 [*DOI :* 10.24348/CORIA.2017.11], https://hal.archives-ouvertes.fr/hal-01617878

[40] *Best Paper*
R. QADER, G. LECORVÉ, D. LOLIVE, P. SÉBILLOT. *Automatic disfluency insertion towards spontaneous TTS : formalization and proof of concept*, in "Traitement automatique du langage naturel (TALN)", Orléans, France, June 2017, https://hal.inria.fr/hal-01532031.

Conferences without Proceedings

[41] C. DALLOUX. *Identifying uncertainty and negation's cues and scope : State of the art*, in "RECITAL 2017 - 18ème Rencontre des Étudiants Chercheurs en Informatique en Traitement Automatique des Langues", Orléans, France, June 2017, pp. 1-14, https://hal.archives-ouvertes.fr/hal-01659646

[42] M. DEMIRDELEN, M. BUDNIK, G. SARGENT, R. BOIS, G. GRAVIER. *IRISA at TRECVid 2017: Beyond Crossmodal and Multimodal Models for Video Hyperlinking*, in "Working Notes of the TRECVid 2017 Workshop", Gettysburg, United States, 2017, https://hal.archives-ouvertes.fr/hal-01643232

[43] N. GRABAR, V. CLAVEAU. *Numerical criteria in clinical trials: annotating, detecting and normalizing*, in "Actes de la conférence TALN 2017", Orléans, France, June 2017, https://hal.archives-ouvertes.fr/hal-01643969

[44] R. SICRE, A. MONTASER AWAL, T. FURON. *Identity documents classification as an image classification problem*, in "ICIAP 2017 - 19th International Conference on Image Analysis and Processing", Catania, Italy, ICIAP 2017: Image Analysis and Processing - ICIAP 2017, Springer, September 2017, pp. 602-613 [*DOI :* 10.1007/978-3-319-68548-9_55], https://hal.archives-ouvertes.fr/hal-01660500

[45] R. SICRE, J. J. RABIN, Y. AVRITHIS, T. FURON, F. JURIE, E. KIJAK. *Automatic discovery of discriminative parts as a quadratic assignment problem*, in "ICCV Workshops – CEFRL", Venise, Italy, October 2017, https://hal.archives-ouvertes.fr/hal-01623148

Books or Proceedings Editing

[46] L. AMSALEG, V. CLAVEAU, X. TANNIER (editors). *Proceedings of the Computaional Journalism workshop 2017*, January 2017, https://hal.archives-ouvertes.fr/hal-01643634

[47] L. AMSALEG, G. Þ. GUÐMUNDSSON, C. GURRIN, B. Þ. JÓNSSON, S. SATOH (editors). *Proceedings of the 23rd International Conference on Multimedia Modeling, MMM 2017, Reykjavik, Iceland, January 4-6, 2017*, MultiMedia Modeling, Springer International Publishing, Reykjavík, Iceland, January 2017, vol. 10132, n$^{\text{o}}$ 0302-9743 [*DOI :* 10.1007/978-3-319-51811-4], https://hal.archives-ouvertes.fr/hal-01430555

[48] G. GRAVIER, B. BISHKE, C.-H. DEMARTY, M. ZAHARIEVA, M. A. RIEGLER, E. DELLANDRÉA, D. BOGDANOV, S. RICHARD, G. J. F. JONES, M. LARSON (editors). *Working Notes Proceedings of the MediaEval Workshop*, November 2017, https://hal.archives-ouvertes.fr/hal-01643258

Research Reports

[49] R. SICRE, A. M. AWAL, T. FURON. *Identity documents classification as an image classification problem*, Inria Rennes - Bretagne Atlantique ; IRISA ; AriadNext, April 2017, n^o RT-0488, https://hal.inria.fr/hal-01503541

References in notes

[50] 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° 2

[51] C. A. BHATT, M. S. KANKANHALLI. *Multimedia data mining: State of the art and challenges*, in "Multimedia Tools and Applications", 2011, vol. 51, n° 1, pp. 35–76

[52] O. CHUM, J. MATAS. *Large-Scale Discovery of Spatially Related Images*, in "Pattern Analysis and Machine Intelligence", 2010, vol. 32, n° 2, pp. 371-377

[53] M. ESKEVICH, G. J. F. 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

[54] C. HERLEY. *ARGOS: Automatically extracting Repeating Objects from multimedia Streams*, in "IEEE Trans. on Multimedia", Feb. 2006, vol. 8, n° 1, pp. 115–129

[55] A. PARK, J. R. GLASS. *Unsupervised pattern discovery in speech*, in "IEEE Transaction on Acoustic, Speech and Language Processing", Jan. 2008, vol. 16, n° 1, pp. 186–197

[56] 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° 2004, pp. 310–327

[57] 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° 1, pp. 5–35

[58] O. DE ROOIJ, M. WORRING. *Browsing Video Along Multiple Threads*, in "IEEE Transactions on Multimedia", 2010, vol. 12, n° 2, pp. 121–130