Regional Program STIC-AmSud
Project Proposal (Research – Innovation)

Basic Form

A. General Information

A1	Project title
	Transforming multimedia data for indexing and retrieval purposes

A2	Acronym
	TRANSFORM

A3	Research domain
	Multimedia; Data mining; Image, Speech and Audio Processing;

A4	Project goals
	⚼ Propose efficient transformations of multimedia data
	⚼ Design efficient indexing schemes based on these representations for retrieval purposes
	Applications to 3D-shapes, temporal and multimodal data indexing
	⚼ Assess the performance of proposed methods for tasks such as recurrence discovery in speech data, large-scale discovery of near duplicate sequences in video and audio tracks, multimodal video content linking and navigation

A5	Abstract
	Multimedia data are usually complex, possibly combining multiple channels and time information to convey a message. It is highly dimensional, multimodal, and involves variability and distortion. In this project, we aim at studying and designing different transformations of multimedia data that enables to facilitate its manipulation, while handling the different particularities stated above. **TRANSFORM** focuses on transforming multimedia data into compact representations that are suited for indexing and retrieval purposes. We will design particularly transformations adapted to 3D-shapes, temporal data and to multimodal data. The designed representations will be integrated into efficient indexing schemes for retrieval purposes. Targeted applications will include 3D-shapes indexing, discovery of audiovisually coherent fragments, of recurrence in speech data and multimodal content linking and navigation. This project will strengthen existing collaborations between Brazilian, Chilean and French labs, all having a strong background in multimedia content analysis. We will also establish exchange of students so that these collaborations can pursue in the future.

A6	**Scientific coordinators at each institution**			
	South America A		**South America B**	
	Institution	Pontifícia Universidade Católica de Minas Gerais - PUC Minas	Institution	Universidad de Chile Department of Computer Science (DCC)
	Project coordinator	Silvio Jamil F. Guimarães	Project coordinator	Benjamin Bustos
	Address	Rua Walter Ianni, São Gabriel, Belo Horizonte, Minas Gerais, Brasil	Address	Beauchef 851, Edificio Norte, 3er Piso, Santiago, Chile
	Phone/Fax	+55 31 3439 5204	Phone/Fax	+56 2 29 78 49 69
	Email	sjamil@pucminas.br	Email	bebustos@dcc.uchile.cl
	South America C		**South America D**	
	Institution	Universidade Federal de Minas Gerais - UFMG Computer Science Department	Institution	
	Project coordinator	Arnaldo de Albuquerque Araújo	Project coordinator	
	Address	Av. Antonio Carlos, 6627	Address	
	Tel/Fax	+55 31 3409 5873 +55 31 3409 5858	Tel/Fax	
	Email	arnaldo@dcc.ufmg.br	Email	
	France A		**France B**	
	Institution	CNRS, Institut de Recherche en Informatique et Systèmes Aléatoires - IRISA	Institution	
	Project coordinator	Simon Malinowski	Project coordinator	
	Address	Campus de Beaulieu 35042 Rennes Cedex	Address	

	Tel/Fax	+33 299 842 557	Tel/Fax	
	Email	Simon.malinowski@irisa.fr	Email	

A7	Other participating institutions

A8	List of expected participants (name and affiliation)
	Guillaume Gravier, CNRS – IRISA, France Simon Malinowski, Univ. Rennes 1- IRISA, France Laurent Amsaleg, CNRS – IRISA, France Benjamin Bustos, Universidad de Chile, Chile Silvio Jamil F. Guimarães, PUC Minas, Brazil Zenilton Kleber Gonçalves do Patrocínio Jr., PUC Minas, Brazil Wladmir Cardoso Brandão, PUC Minas, Brazil Arnaldo de Albuquerque Araújo, UFMG, Brazil Kleber Jacques Ferreira de Souza, PU C Minas, Brazil

A9	International Project Coordinator
	Simon Malinowski, Univ. Rennes 1- IRISA, France

B. Project Details

B1. Project guidelines

The amount of multimedia data has increased over the last decades as social networks and applications are becoming more and more popular and as cutting-edge digital equipment are accessible to a large population. This kind of data represents an enormous amount of information, from which different application domains (such as culture, environment, finance, social studies) can benefit.

Multimedia data are usually complex, possibly combining multiple channels and time information to convey a message. The inherent nature of multimedia data has to be taken into account when developing automatic systems to analyze it:
- Multimedia data is highly dimensional and multivariate
- Multimedia collections are huge: there is a strong need for efficient methods
- Multimedia data is multimodal (color, sound, texture, time): all the modalities brings its own information
- Multimedia data involves variability and distortion, particularly in spoken language or in images.

Tools and methods that will be developed and designed in the scope of this project will have to be adapted to such peculiarities. In this project, we aim at studying different transformations of multimedia data that enables to facilitate its manipulation, while handling the different particularities stated above. Hence, TRANSFORM **focuses on transforming multimedia data into compact representations that are suited for indexing and retrieval purposes**.

We are particularly interested in the retrieval task: given a query object (that can be any kind of multimedia data, such as image, video, sound, shape) and a database of objects, we want to find in this database object(s) that is (are) similar to the query. Of course, brute-force search is inappropriate here due to the nature of data. In order to design efficient schemes adapted to this task, we will focus on the design of transformations of multimedia data that will enable to reduce the complexity of the retrieval task while guaranteeing a sufficient level of accuracy.

TRANSFORM will contribute fundamental results in indexing and retrieval of multimedia data:
1. Design mid-level representations for 3D-shape and temporal data
2. Design similarity-preserving representation for temporal data
3. Design multimodal representation for multimedia data

Scope of the project

TRANSFORM will **strengthen collaborations between Brazilian, Chilean and French labs** with a strong background in multimedia content analysis, indexing and mining, to **propose, study and evaluate approaches to learn transformations for multimedia data**, both from a fundamental point of view and an application point of view. We will develop fundamental technology at the frontier of multimedia content analysis, multimedia indexing and database management to propose new representations for multimedia data. Achievements will be instantiated and demonstrated for different tasks:
- discovery of audiovisually coherent fragments in video contents
- recurrence discovery in speech data
- multimodal video content linking and navigation
- 3D-shape indexing

Challenges

The challenges that will be tackled by this project are mostly related to the specificities of multimedia data, like **scalability, variability and distortions**, which are stated above.

One of the first challenges is related to **scalability**. The techniques designed throughout this project will have to be able to handle large quantities of multimedia data. The rationale behind using transformed data rather than original data is mostly related to this scalability challenge. Transforming data into a vectorial representation in a Euclidean space has many advantages with respect to complexity issues: objects can be compared efficiently with the Euclidean distance and easily inserted in an indexing system.

Variability and distortions are inherent to multimedia data, particularly in audio data. It is hence essential for the learned transformations to handle such variability. In the case of audio, the Dynamic Time Warping measure is able to handle temporal distortions. However, its complexity is quadratic with the sequence lengths, which prevents from using it in applications where time series are long and data sets are huge. In addition, it is not a distance, which makes it not well suited to use it in kernel-based algorithms. In this project, we will aim at designing transformations that embed temporal data into an Euclidean space such that distances in this transformed space reflects the dynamic time warping measure in the original space.

Another issue related to this project concerns the **quality of the designed transformations**, since multimedia data involves different modalities, bringing complementary information. We will hence focus on designing transformations that capture information about each modality: color, shape, temporal space. Efficiently combining the different modalities will also be tackled in this project.

To summarize, TRANSFORM targets significant contributions, based on learning transformations for multimedia data, in the following challenges:

> ➢ Design of accurate representations for shape retrieval (task 1 below)

> ➢ Efficient indexing of temporal data (tasks 2 and 3 below)

> ➢ Design multimodal representation spaces (task 4 below)

State of the art and progress beyond

Multimedia data involves some specificities related to **scalability, variability, different modalities and distortions**. This proposal will cope with these issues by transforming the multimedia data into a representation adapted to indexing and retrieval purposes by considering, mainly, (i) mid-level representations for 3D-shape indexing and for temporal data; (ii) learning similarity preserving transformations of temporal data; and (iii) multimodal embedded representation spaces.

1) Mid-level representations for 3D-shapes and temporal data

Working on raw multimedia data is not an easy task due, mainly, to, its complexity like variability, distortions and different modalities, thus, instead of using raw data, TRANSFORM targets significant contributions in mid-level representation of the multimedia data in tasks as indexing, retrieval and classification. In an application point-of-view, TRANSFORM targets contributions by using two types of data: (i) 3D-shapes; and (ii) temporal data.

3D data has become an important data type for several real-world applications and domains. For example, 3D data is used in manufacturing processes, in models of proteins for biological studies, in cultural heritage projects, and on the entertainment industry. So, it is not surprising that in recent years there has been a lot of research work focused on this type of complex data.

A relevant problem involving 3D data is 3D shape matching. In this problem, one tries to find matches or correspondences between two 3D objects. This analysis of the 3D objects is made only using the geometrical (spatial) information of the objects. Usually, 3D objects are modeled as triangular meshes that define their surfaces. Also, 3D objects can be modeled as point clouds, where each point of the cloud corresponds to a point at the surface of the object. A point cloud is the usual output given by a 3D scanner after scanning an object, which can be later post-processed for obtaining a triangular mesh.

For implementing 3D shape matching algorithms, the standard approach is to compute descriptors (feature vectors) that describe the objects and allows them to be compared. This process is called 3D shape indexing. Descriptors for 3D objects can be organized in two main groups: global descriptors and local descriptors. A global descriptor is a high-dimensional feature vector that completely describes a 3D object. Thus, only one global descriptor per 3D object is required for matching and retrieval. A local descriptor is a feature vector that locally describes a region of a 3D object. Usually, a set of local features is computed from a single 3D object for describing it.

Global descriptors have been used for solving "whole-to-whole" matching (comparing a whole 3D object with another whole 3D object) [Bustos05]. Local descriptors are better suited for solving partial retrieval,

where for example one only has a part of the original 3D object and wants to match it with a whole object ("part-to-whole" matching). Once having the global/local descriptors and appropriate matching functions, it is possible to implement tasks like similarity search on 3D repositories. In this problem, one has a query object and wants to find all relevant objects in the repository. This type of search algorithms falls into the category of content-based information retrieval, as the aim of these algorithms is to find information (matches) on the repositories by using the 3D data itself.

A current open problem is multimodal **3D shape** matching, where additionally to geometry other modalities are used to perform the matching process. In particular, we are interested in 2D-to-3D matching. Content-based image retrieval is a research area with a long tradition, with surveys published already in 2000 [Smeulders00]. We plan to take advantage of sketch-based image retrieval methods [Eitz10, Saavedra14], as previous work on 3D shape matching have shown that view-based representation methods are effective for globally matching the geometry of 3D models [Bustos05, Tangelder08]. Also, there has been many recent research works focusing on 3D sketch-based retrieval [Li14].

Inspired by text mining, information retrieval and computer vision communities, recent works have investigated the use of Bag-of-Words for **temporal data** (mostly for time series classification though) [Baydogan13, Baydogan15, Lin12, Senin13, Wang13, Xie09, Bailly16]. These works are based on two main operations: converting time series into Bag-of-Words, and building a classifier upon this BoW representation. Usually, standard techniques such as random forests, SVM, neural networks or kNN are used for the classification step. Yet, many different ways of converting time series into Bag-of-Words have been introduced. Among them, [Baydogan13] proposes a framework to classify time series denoted TSBF where local features such as mean, variance and extremum values are computed on sliding windows. In [Wang13], discrete wavelet coefficients are extracted on sliding windows and then quantized into words using k-means. In [Lin13, Senin13], words are constructed using the Symbolic Aggregate approXimation (SAX) representation of time series. SAX symbols are extracted from time series and histograms of n-grams of these symbols are computed to form a Bag-of-Patterns (BoP). In [Senin13], authors combine SAX with Vector Space Model to form the SAX-VSM method. In [Baydogan15], authors design a symbolic representation of multivariate time series (MTS), called SMTS, where MTS are transformed into a feature matrix, whose rows are feature vectors containing a time index, the values and the gradient of time series at this time index (on all dimensions). Random samples of this matrix are given to decision trees whose leaves are seen as words. A histogram of words is output when the different trees are learned. In [Schafer14], Fourier coefficients are extracted and quantized into words. Histograms of n-grams of such words are then computed. Fourier coefficients are also used by the same author to speed-up time series similarity search [Schafer12]. For the same task, Chebyshev polynomials are used in [Cai04].

Local feature extraction has been investigated for long in the computer vision community. One of the most powerful local feature for image is SIFT [Lowe04]. It consists in detecting keypoints as extremum values of the Difference-of-Gaussians (DoG) function and describing their neighborhoods using histograms of gradients. Xie and Beigi [Xie09] use similar keypoint detection for time series. Scale-invariant features that characterize the shapes surrounding the extremum then describe keypoints. In [Bailly15], authors use quantizes 1-dimensional SIFT-descriptor to design a time series classifier based on Bag-of-Words.

In most of these approaches, temporal information is lost in the Bag-of-Word representation. In this project, we will be interested in extending this kind of representation so that temporal information can be included.

2) Temporal data indexing under Dynamic Time Warping

One popular similarity measure to compare time series is the Dynamic Time Warping (DTW), due to its capacity to cope with time shifts and warpings. Its complexity being quadratic with the length of time series, it is difficult to use DTW against very long time series and/or very large sets of time series. In turn, many research works have attempted to reduce that complexity have tried to run DTW onto a very limited subset of candidates.

In [Keogh05], authors proposed a lower bound of the DTW that is used during the search to prune sequences that cannot be the best match. An improved version of this lower bound is given in [Lemire09]. UCR Suite is designed in [Rak12]. This system combines 4 techniques to speed up the search including cascading lower bounds and early abandoning while computing distances. Authors of [Salvador07]

propose a multilevel approach to approximate the DTW in linear time with the sequence length. Tan et al [Pet17] have proposed an indexing scheme under DTW that combines a hierarchy of k-means and lower boundings of DTW.

In this project, we will adopt another point-of-view for indexing temporal data under Dynamic Time Warping. We will focus on learning different vectorial representations to embed time series in an Euclidean space such that the distances in that space reflects as well as possible the original DTW between time series. Working in such an Euclidean space is very useful for indexing purposes.

3) Multimodal embedded representation

Many situations require a common space to represent multiple modalities, e.g., texts and images. For instance, cross-modal information retrieval requires a common space to retrieve say an image for a text description (or conversely). Multimodal indexing also requires a common representation space, known as embedding, for the different modalities so that efficient indexing techniques can be used. Because of the inherently different nature of the representations coming from the different modalities, designing a representation space common to multiple modalities is highly challenging. Seminal approaches to this problem exploited canonical correlation analysis (CCA) to project the representation of each modality onto a common space where the projections of the different modalities of the same object are maximally correlated. While CCA initially considered linear projections, kernelized versions were also proposed. CCA is however known to have a number of strong limitations, clearly evidenced in [Tran16]. Recent advances in deep learning opened the door to neural-based approaches to multimodal embedding. Compact vector representations are now available for all modalities, e.g., skip-gram word embedding for texts, convolutional neural network embedding for images, long-short term memory embedding for videos. A number of recent pieces of work exploited these representations to learn a network—typically an auto-encoder multilayer perceptron—that learns how to combine vector representations from different modalities [Ngiam11, Srivasta12, Weston10]. We introduced a slightly different approach in [Vukotić16] based on the translation from one modality to the other to learn a multimodal embedded space rather than on the combination of the two modalities. Translation was shown to improve over combination and recent work on adversarial neural networks [Vukotić17] opened the door to a promising new family of translation approaches that we will push forward within the scope of the TRANSFORM project.

B2. Project description

TRANSFORM targets **significant scientific contributions** and **reinforcement of collaborations**, which can be outlined as follows:
1) Design of new transformations adapted to multimedia data
    a) Mid-level representations for 3D-shape
    b) Mid-level representations for temporal data
    c) Similarity-preserving representations for temporal data
2) Applications of these transformations
    a) Indexing of 3D-shapes
    b) Time series indexing and clustering
    c) Word retrieval in speech data
3) Design of multimodal representations for multimedia data
4) Reinforce the synergy between Chile, France and Brazil on this topic, targeting cross-fertilization between complementary skills

Expected results

TRANSFORM will contribute results both rom a fundamental point of view and an application point of view. The objectives and expected accomplishments resulting from the collaboration are as follows:
- Establish a technical cooperation between all labs involved in this project;
- Establish exchange of students;
- Study mid-level representations applied to multimedia sequences;
- Study mid-level representations applied to 3D-shapes;

- ⚲ Study multimodal representations for multimedia data;
- ⚲ Study similarity-preserving representations for temporal data;
- ⚲ Propose mid-level representations from 3D-shapes and temporal data for indexing, clustering and retrieval purposes;
- ⚲ Propose similarity-preserving representations from temporal data for indexing, clustering and retrieval purposes;
- ⚲ Propose multimodal representations for multimedia data for indexing and retrieval purposes
- ⚲ Design methods for word retrieval in speech data taking into account the proposed representations

We expect also to improve the qualification of the Brazilian, French and Chilean students (graduate and under-graduate levels), either discussing with all project members or going to other countries which integrates this project. Also, to share our results, we expect to publish some works and organize events, as follows:

- five (5) papers in international conferences;
- two (2) papers in national conferences;
- four (4) papers in international journals;
- three (3) master dissertations;
- four (4) undergraduate works;
- four (4) PhD thesis;
- organize an special session in one conference related to this subject, like CBMI
- organize a special issue in a journal;
- write a corresponding tutorial style "survey paper".

Methodology

In this proposal, we aim at designing and learning new representations of multimedia data that can handle the various specificities of this kind of data, such as temporal distortion and variability, multivariate data, multimodality.

TRANSFORM is divided in 4 tasks, all linked with the transformation of multimedia data to facilitate its manipulation: (i) mid-level representations for 3D-shape indexing; (ii) mid-level representations for temporal data; (iii) learning similarity preserving transformations of temporal data and (iv) multimodal representation spaces. A different partner will lead each of these tasks, but as these tasks are all related to a same field, we expect and will encourage interactions between them.

1) Mid-level representations for 3D-shape indexing

Leader: Benjamin Bustos (Univ. Chile/DCC)

In the area of 3D shape retrieval and indexing, much effort has been devoted to design and implement algorithms that search in the same object domain [Bustos05, Sipiran15]. Indeed, the typical query model has been the query-by-example, where an object of the same modality as the data set, a 3D model), is used as input for the retrieval system. The goal of the system is to obtain similar objects to the one posed as query. This works well on several applications contexts, but limits the domain where the retrieval process can be done.

An interesting problem is to match 3D objects using as query a different modality. This allows one to extend the search capabilities of the retrieval system and explore new applications. In particular, we are interested in the problem of using an image as query object, and perform the search in a 3D data set. For example, we may have a catalog with pictures of archaeological artifacts exhibited in a museum, and we may also have a 3D data set with all the scanned artifacts on the museum. Then, we would like to have automatic tools for searching which picture from the catalog corresponds to each 3D model in the data set. This particular problem needs to tackle multimodal shape retrieval (2D to 3D search).

In this research, our hypothesis is that a common representation in a feature space of both 2D to 3D objects will allow us to design a similarity search algorithm. This idea needs to solve two related problems: 1) we need to find a common a common feature space for both types of modality, and 2) we need to describe both 2D and 3D objects in this common space in a way that is suitable for implementing similarity search. For solving Point 1, our preliminary idea is to represent both 2D and 3D objects as sketches (like a hand drawn sketch). In the literature, there are techniques for representing 2D and 3D objects as sketches [Yoon10, Saavedra14, Li14], but we need to research which specific method is well suited for the task. Then, we need to choose an appropriate descriptor for characterizing sketches

computed from both modalities. For solving Point 2, our hypothesis is that mid-level representations will allow us to "merge" the descriptors computed from both modalities, for representing all objects in a way that improves the effectiveness of the retrieval step.

In summary, the proposed methodology for 2D-to-3D shape retrieval is the following:

- Generate a reference collection for evaluating 2D-to-3D matching. In a first approach, this could be done by obtaining objects from ShapeNet [Shape] and synthetically obtaining 2D projections for simulating the 2D data.

- Find adequate methods for describing 2D and 3D data as sketches.

- Compute sketch-based descriptors and research mid-level representation techniques for creating the common feature space.

- Implement similarity search algorithms on the common feature space and evaluate the effectiveness of the technique.

It is key to the success of this research that the common feature space allows one to implement an effective matching of the 2D query object and object from the 3D dataset. Therefore, this project will greatly benefit from a joint collaboration with other researchers from this STIC-AMSUD project that have previous experience working on mid-level representations.

2) Mid-level representations for temporal data

Leader: Silvio Jamil F. Guimarães, PUC Minas

The most discriminative low-level descriptors available in literature today rely on identifying regions of interest. Once these regions are identified, desirable features are then extracted around these regions. The output created by this process is a set of features, which are related to regions, representing the media. Facing this scenario, one popular approach is to map the set of local descriptors into one vector used as a global representation, so-called mid-level representation.

Among the methods for creating a mid-level representation, standout Bag-of-Words (BoW), Spatial Pyramids and Convolutional Networks for their notable results. Mid-level representations have three steps in common: (i) coding; (ii) pooling; and (iii) concatenation. Coding stands for the locally transformation applied into features vectors, extracting distribution characteristics. Pooling, in turn, explores the spatial relation between these characteristics; and concatenation constructs the final vector representation. One of the steps for computing mid-level representation is the identification of codewords that will characterize the most relevant features concerned to the data.

TRANSFORM will explore new strategies for the pooling step taking into account, for example, volumetric partition of the hypersphere centered at codewords. The goal is to maintain the same probability of assignment to a given hyper-region. We argue that this kind of pooling could decrease the quantization error that was created during codification. It is important to observe, in most of BoW approaches, temporal information is lost in those representations. In this project, we will be interested in extending this kind of representation so that temporal information can be included.

In summary, the proposed methodology for representing temporal data is the following:

- Identify the region of interest and/or use dense sampling;

- Compute the vocabulary for representing the most important features;

- Explore spatio-temporal relation between the features

- Create the final mid-level representation by concatenation

The key of the success of this representation is to understand how to aggregate the temporal information to the spatial data. Therefore, **TRANSFORM** will greatly benefit from a joint collaboration with member of the consortium that have previous experience working on mid-level representations and description of temporal data.

3) Learning representations to preserve similarity in temporal data: application to spoken content indexing

Leader: Simon Malinowski (IRISA, Univ. Rennes 1)

Spoken content can be seen as time series but with some peculiarities: it is highly dimensional (as opposed to multivariate time series that are usually considered) and may have high temporal variability (distortion in spoken content). We will hence investigate on the development of spoken content indexing approaches based on recent techniques about time series analysis that comply with such particularities.

Most of the time series indexing approaches in the literature rely on the Dynamic Time Warping (DTW) measure. As this measure has a quadratic complexity with the sequence length, it is not tractable as it is when dealing with very long time series and/or huge data sets. Another issue with DTW is that it is not a distance. Hence, it cannot be optimally used with classical kernel based algorithms (for classification for instance) or in classical indexing systems (designed for metric spaces).

Recently, we have designed a similarity-preserving transformation for time series [Lods17]. This method aims at learning a transformation of time series in a metric space such that the Euclidean distances between transformed series reflects as well as possible the DTW in the original time series space. Using such a transformation hence enables to work with time series in a metric space, which facilitates tasks such as clustering, indexing and classification. This transformation is based on the shapelet transform (adapted to the purpose of preserving similarity).

In this project, we will investigate on how to use such a transformation to build efficient time series indexing systems. We will aim in particular in designing a transformation such that the first components of the transformed series contain most of the power. This kind of transformation is particularly adapted to situations where limited amount of time is available at query time. A first rough approximate search can be made using only the first components of the transformed query. This search can then be improved if more time is available by using next transformed components.

The collaboration between the partners this project will help for these tasks, as we will benefit from the experiments of both partners in indexing high dimensional multimedia data.

4) Multimodal representation spaces

Leader: Guillaume Gravier (CNRS, IRISA)

Multimodal representations have recently gained interest with the upraising of neural network embedding of multimodal content. The model that we developed in [Vukotic2016] was successfully applied in a video hyperlinking competition, TRECVid, where we obtained the best results in 2016. The video hyperlinking task consists in created links within a large video collection (>2000 hours of videos) given a set of segments from which links are to be created. The task clearly benefits from a multimodal representation of the content of the video, mixing spoken content information and visual information. We recently extended the model to handle multimodal person recognition, combining face and voice into a single embedding, following up on previous work jointly implemented at PUC Minas and IRISA.

In **TRANSFORM**, we will pursue research activities on multimodal embedding and extend previous work in a number of ways. In particular, adversarial approaches to multimodal embedding have given promising results [Vukotic2017] that need to be generalized and strengthened. There are currently strong limitations to these approaches, which require fixed size images. In this project, we will seek to plug in more complex image generation neural networks into an adversarial architecture. Moreover, current approaches to multimodal embedding do not account for time, ignoring the order of the words in sentences and relying on a single key frame image rather than taking into account the temporal evolution of the visual content. We will leverage recurrent neural architectures such as long-short term memory (LSTM) networks to embed temporal multimodal sequences. The former were recently introduced in sentence embedding, accounting for the order of the words, e.g., in translation, and in image processing for classification tasks. As a first approximation, LSTM networks can be used independently in each modality before combining the resulting embedding. But we will also seek to design inherently multimodal LSTM networks that take input from textual, visual and audio modality simultaneously.

Results on multimodal embedding will be demonstrated in two tasks, namely video hyperlinking and multimodal person identification. IRISA has a strong background on the former, while PUC Minas and

UFMG jointly worked on the latter. In both cases, we will participate in the corresponding international benchmarks (TRECVid and MediaEval), offering international visibility to the teams and to the project.

5) Complementarity within the tasks

The 4 tasks described above will not be tackled independently. They have some connectivity that we will take into account to develop joint works.

Tasks 1 & 3: We will seek to bridge multimodal embedding with 3D shape modeling, considering each dimension as a modality.

Tasks 2 & 3: These tasks share the same objective. Comparison of the developed methods should enable to gain more insight on efficient representations for temporal data.

Tasks 3 & 4: LSTM networks (studied in task 4) can also be used as a similarity-preserving transformation. Closed work within these two tasks will hence take place.

B3. Schedule, with main execution stages

TRANSFORM will be developed in 2 years (divided by periods of two months), and can be divided in some steps taking into account the 4 tasks as follows: (i) mid-level representations for 3D-shape indexing; (ii) mid-level representations for temporal data; (iii) learning similarity preserving transformations of temporal data and (iv) multimodal representation spaces. To achieve the main expected results related to these tasks, in Table 1, we illustrate the major steps with the time to execution

A plenary kick-off meeting with representatives from all partners will take place at the very beginning of the project to strengthen the knowledge of each partner, foster new collaborations and ensure good collaborations by organizing the workflow and the contributions from the respective partners.
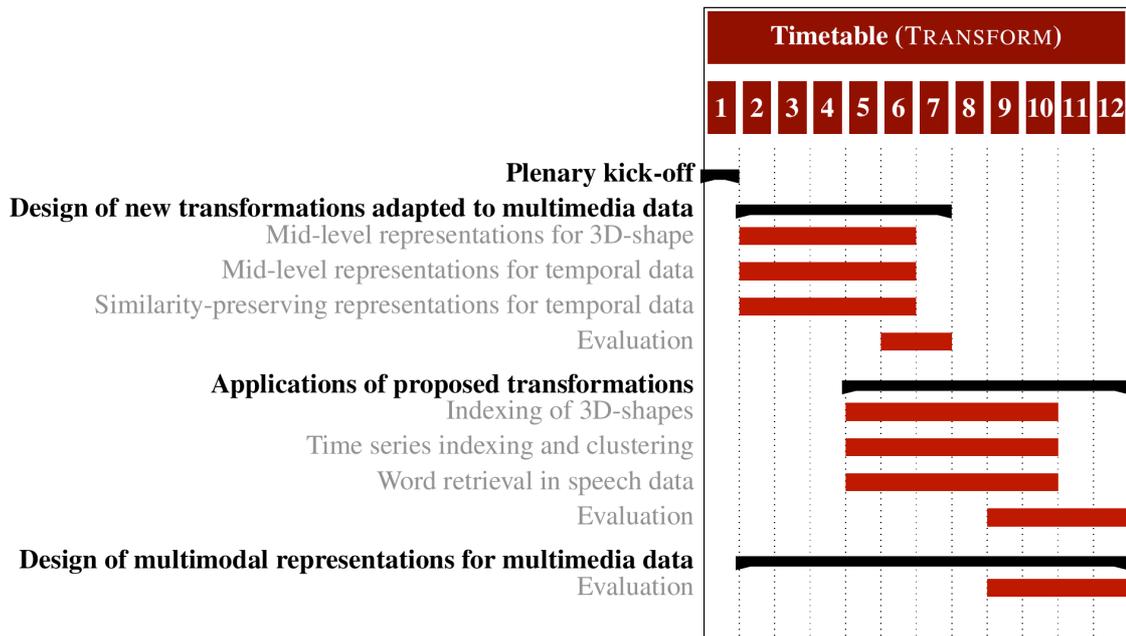

Table 1: Timetable for TRANSFORM

B4. Contributions

IRISA/LINKMEDIA
IRISA will contribute to all research axes with a focus on tasks 3 and 4. The main contributions are expected to be the following:
  ➢ Study and design efficient similarity preserving representations for temporal data
  ➢ Design indexing schemes for temporal data using such representations
  ➢ Design multimodal spaces for multimedia data

IRISA/LINKMEDIA will also be involved in tasks 1 and 2, by bringing insight on mid-level representations for 3D-shapes and temporal data and also on indexing systems.

PUCMINAS/VIPLAB

In this project, PUCMINAS/VIPLAB will lead task number 2, and will be involved in all tasks. The main contributions are expected to be the following:

➢ Study and design powerful mid-level representations for temporal data
➢ Design indexing structures: propose pooling and normalization adapted to the type of data processed and study similarity learning approaches for data comparison, in collaboration with IRISA

PUCMINAS/VIPLAB will hence also be involved in task 1 and 3 that shares same aims at task 2 but with different approaches. Techniques developed by PUCMINAS/VIPLAB will be considered also for task 4.

UCHILE/DCC

In this project, **UCHILE/DCC** will lead task 1, and will also be involved in tasks 2 and 3. The main contributions are expected to be the following:

➢ Study and design representations adapted to efficiently describe 3D-shapes
➢ Design retrieval and indexing schemes built on these representations for 3D-shapes.

UCHILE/DCC will also contribute to tasks 2 and 3, thanks to its experience in temporal data analysis and indexing.

UFMG/DCC

In this project, UFMG/DCC will be involved in all tasks. It will bring contributions in these different fields:

➢ Data representation: investigate new techniques in order to get compact and powerful data representations to be used in mid-level representations
➢ Indexing structures : study and propose metric data structures for the different tasks tackled by the project (3D-shapes, temporal data and multimodal information).
➢ Pooling: investigate new techniques in order to decrease the quantization errors preserving the temporal information

UFMG/DCC will hence also be involved in task 3 that shares same aims at task 2 but with different approaches. Techniques developed by UFMG-DCC will be considered also for task 4.

B5. Regional Aspects

TRANSFORM will reinforce the France-Brazil-Chile collaboration (visits, student exchanges, possibly joint Ph. D. theses). Fruitful collaborations between France, Brazil and Chili have already emerged thanks to a previous project so-called MAXIMUM. We expect this project to strengthen existing collaborations. This consortium can generate effective results since the teams are complementary. The experience of regional coordinators and several members of the consortium in cooperation projects will facilitate the discussions and progress of this proposal. We expect each partner to acquire skills in the area of expertise that are only covered by other partners. This will be made possible by regular visits of scientists and student exchanges and by a plenary kick-off meeting. We hence consider the collaboration established within TRANSFORM to last after the end of the project. Joint work, integration of fundamental results in a limited number of well-identified lab applications and regular visits and student exchanges will undoubtedly create strong links on which to build new proposals.

B6. Institutions and CVs of coordinators

IRISA / LINKMEDIA

IRISA is a public laboratory, common to CNRS, INRIA, Univ Rennes 1, INSA de Rennes, ENS Cachan, Télécom Bretagne, Univ. Bretagne Sud and Supélec, and devoted to research in information and communication science and technologies, with a first rank position in the following domains: i) Networks, Telecommunications and Multimedia, ii) Complex Systems and Software, iii) Modeling, Simulation and Visualization. Within IRISA, the LinkMedia team seeks to develop methodological, technological and scientific foundations to create, describe and exploit explicit links within multimedia collections on a large scale. We instantiate research activities along three major lines of research, namely 1. unsupervised motif discovery (improve multimedia indexing for efficient comparison of content items, develop scalable algorithms for the unsupervised discovery of repeated motifs, investigate motif

interpretation and exploitation) 2. description and structuring of multimedia content (improve technology for multimedia content modeling and for semantic interpretation, offering a better trade-off between accuracy, robustness and scalability) 3. linking of multimedia content (propose methodology for the creation and the description of explicit links in multimedia collections, study the impact of linked media, revisiting existing tasks and technology and developing new applications). LINKMEDIA has a strong experience in multimedia data analysis with numerous works and publications on this topic.

People involved at IRISA: Simon Malinowski (IRISA/ Univ. Rennes 1), Guillaume Gravier (DR CNRS, coordinator) and Laurent Amsaleg (CR CNRS)

PUCMINAS/VIPLAB

The Pontifical Catholic University of Minas Gerais – PUC Minas, founded in 1958, is one of the five largest universities in Brazil. Its activities concentrate mostly in the Belo Horizonte Campus, divided into three units – Coração Eucarístico, Barreiro and São Gabriel – and two university centres – Betim and Contagem. PUC Minas also comprises campi in four regions of the State of Minas Gerais: Arcos, Poços de Caldas, Serro and Guanhães. Altogether, PUC Minas has about 60,000 students, 2,000 faculty members and 1,800 staff members. VIPLAB is a research lab hosted at PUC Minas (Belo Horizonte Campus, Brazil) and it has started its operation in 2004. The main project areas are: digital image, video and audio processing; multimedia indexing and retrieval; multimedia content analysis. Many projects have been conducted with cooperation of other institutions and/or labs, such as NPDI/DCC/UFMG, IC/UNICAMP, LINKMEDIA/IRISA and ESIEE/FRANCE. And, besides the projects developed with the participation of several graduate and undergraduate students, VIPLAB has also been involved in bilateral programs for international cooperation such as FAPEMIG-INRIA, CAPES-COFECUB, STIC-AMSUD and CAPES-BRAFITEC.

People involved at PUCMINAS/VIPLAB: Silvio Jamil Ferzoli Guimarães, Wladmir Cardoso Brandão, Zenilton Kleber Gonçalves do Patrocínio Jr. and Kleber Jacques Ferreira de Souza.

UCHILE/DCC

The Department of Computer Science (DCC) at the University of Chile is one of the oldest Computer Science departments in Chile, with more than 40 years of tradition. Within DCC, and related to this project, the PRISMA research group focuses its research in the areas of data mining, social media and Web analysis, multimedia databases, pattern recognition, and similarity search. Its main purpose is twofold, 1) to develop efficient and effective methods for Multimedia Information Retrieval and 2) to extract valuable knowledge from social and multimodal data for improving human understanding of large data sets.

People involved at UCHILE/DCC: Benjamin Bustos.

UFMG/DCC

The Universidade Federal de Minas Gerais is one of the best-ranked federal (public) universities in Brazil. Our graduate course in Computer Science is one of oldest and well established PhD and MSc program in the country, being classified with grade 7 (in a scale 4-7) by CAPES-Ministerium of Education. The program aims at academic excellence in all dimensions to form human resources highly qualified, with a solid scientific and technological basis. The Computer Science Department, in the area of this project, develops considerable research on Image and Video Processing, Computer Vision, Machine Learning, Information Retrieval and Digital Libraries. The department has participated in main federal actions for excellence labels: PRONEX, INCT. In the areas of Image Processing and Visual Information Retrieval, we have been supported by three CAPES-COFECUB projects (1997-2000, 2002-2005, 2008-2011) involving UFMG, UNICAMP, PUC-Minas and UFOP, in the Brazilian side, and ESIEE Paris, ENSEA Cergy-Pontoise and UPMC-Lip6, in the French side. The department has been involved in many other bilateral programs for international cooperation (with France, Germany, Argentine, USA) like CNPq-CNRS, INRIA-FAPEMIG, CNPq-BMBF, BRAFITEC

People involved at UFMG-Computer Science Department is: Arnaldo de Albuquerque Araújo

Complementarity between the involved research teams

Well-established existing collaborations will facilitate synergy within TRANSFORM and help fostering new ones. The collaboration between UFMG, PUC and IRISA started 5 years ago on a project in

multimedia content analysis, the exchange of PhD Students is the best way for integrating the common research. Moreover, two international collaboration were supported by STIC-AMSUD (2014-2015) involving UFMG, IRISA, PUC Minas and UCHILE, and FAPEMIG/EQUIPE-ASSOCIE (2014-2016) involving UFMG, IRISA and PUC Minas.

During next year, we plan to co-advisor a PhD Student in a subject, which matches to the proposal project, more specifically, in "Tagging and indexing speaking faces". We expect another joint Ph. D. thesis in 1 or 2 years on representing temporal data for indexing purposes.

UFMG, UPMC, IRISA and UCHILE/DCC have also been in informal exchanges with joint participation in juries in 2013 for the doctoral program of UFMG. Moreover, IRISA and PUC Minas co-advisor one master student in PUC Minas, in 2017.

B7. Additional information

List all the complementary funding expected or already obtained.

CAPES/COFECUB, to be submitted in 2017 (People involved: Guillaume Gravier, Silvio Jamil F. Guimarães, Arnaldo de Albuquerque Araújo, Zenilton Kleber Gonçalves do Patrocínio Jr.)

Research Project approved at the CNPq Universal Edital 2017 (Level B - R$ 60,000.00) to support the acquisition of a new data processing server for the VIPLAB (People involved: Silvio Jamil F. Guimarães and Zenilton K. G. do Patrocínio Jr.).

GRANT for Researchers at the CNPq (Level 2) and at FAPEMIG (PPM) for the VIPLAB (People involved: Silvio Jamil F. Guimarães)

GRANT for Researchers at the CNPq (Level 1D) for the NPDI (People involved Arnaldo de Albuquerque Araújo)

Experience of the coordinators in similar projects

Guillaume Gravier leads the Linkmedia research group and the long history of international collaboration. He has participated in various international projects, among which the NoE Multimedia Understanding through Semantics, Computation and Learning (MUSCLE), the STIC-AmSud project MAXIMUM and the Inria Associate Team MOTIF. He has been coordinating these last two projects. Guillaume Gravier has also a history of informal international collaborations with National Institute for Informatics, Japan (joint participation to The Star Challenge in 2008), with INAOE, United States of Mexico (2 months stay), or with Delft Univeristy, Netherlands (joint project, exchange of students).

Simon Malinowski has participated in the previous STIC-AMSUD project MAXIMUM between France, Brazil and Chile (2 stays in Brazil, 1 in Chile, joint publications in progress) and in the Inria Associate Team MOTIF. Last year, he organized a workshop on temporal data analysis collocated with ECML/PKDD 2016.

Benjamin Bustos has been the head in the Chilean side of the international collaboration with Germany in the area of 3D Object Retrieval. The cooperation with Dr. Tobias Schreck (German side) started in 2003 and it is documented by a history of joint publications. This collaboration was formally funded during 2010 by the Deutsche Forschungsgemeinschaft (DFG), which supported one visit from Dr. Schreck to Chile and one visit from Prof. Bustos to Germany. In particular, the cooperation during that year produced one publication in an international conference and one publication in an indexed journal. Furthermore, two visits of Ph.D. students from Chile to Germany have taken place during 2010 and 2012, respectively. During the period 2013—2017, six more visits (three in Chile and three in Europe) where partially funded by CONICYT-FONDECYT (Chile) projects. This international collaboration is still ongoing and has produced so far 26 joint publications in the period 2003—2017.

Arnaldo de Albuquerque Araújo has been the head of the international cooperation with France in the areas of Image Processing and Visual Information Retrieval. This cooperation started in 1990, with a first visit to ESIEE Paris where one of his first graduate students was making his PhD thesis with a scholarship done by CAPES. Arnaldo made a second visit to ESIEE in April 1992 and invited Gilles Bertrand (head of the A3SI lab) to visit UFMG and to be one of the invited key speakers for SIBGRAPI 92. The main

next step to strengthen the cooperation was a post-doctorate mission (supported by CNPq) done by Arnaldo at ESIEE in the period of may 1994 to December 1995, involving three new PhD students (full time and sandwich study missions). During this period, Arnaldo visited many other research groups (ENSEA, ENST, ENSMP) and invited other colleagues (Mohamed Akil, Michel Couprie, Isabelle Bloch, Sylvie Philipp-Foliguet, Jean Pierre Cocquerez) to come to Brazil to participate of SIBGRAPI as invited speakers or give a tutorial course (followed by visits to UFMG and UNICAMP). After that, we have been supported by three CAPES-COFECUB projects (1997-2000, 2002-2005, 2008-2011) involving UFMG, UNICAMP, PUC-Minas and UFOP, in the Brazilian side, and ESIEE Paris, ENSEA Cergy-Pontoise and UPMC-Lip6, in the French side. We have had a strong formation of PhD students from UFMG and UNICAMP in the cooperation, including full time, sandwich program and double diploma theses. As a direct result of the human formation activity the team has been involved in many actions together: publications, participation in defense juries and congress committee programs, research workshops in UFMG and UNICAMP by occasion of the French work missions. We highlight the France-Brazil Digital Video Journey (together with SIBGRAPI 2009) organized by Arnaldo and Matthieu Cord (UPMC), by occasion of the Year of France in Brazil, and the Argentina-Chile-Brazil Workshop on Information Retrieval where we invited the team of Benjamin Bustos (U. de Chile) to meet our team (UFMG, PUC-Minas and UNICAMP) here at UFMG, in November 2012.

Silvio Jamil F. Guimarães has been the head of the international cooperation with France in the areas of Image Processing and Visual Information Retrieval. This cooperation started in 2008, with a first work mission to ESIEE Paris after his PhD. After that, we have been supported by one CAPES-COFECUB project (2008-2011) involving UFMG, UNICAMP, PUC-Minas and UFOP, in the Brazilian side, and ESIEE Paris, ENSEA Cergy-Pontoise and UPMC-Lip6, in the French side. In 2010, started a cooperation with IRISA/TEXMEX (supported by FAPEMIG-INRIA). During these last years, many work missions were done in order to well establish scientific relationship. The main next step to strengthen the cooperation was a post-doctorate mission (supported by CNPq) done by Silvio at ESIEE in the period of august 2011 to July 2012. During this period, Silvio visited many other research groups (ENSEA, EFREI, CNAM, IRISA/LinkMedia). In 2013, the first work mission to IRISA/ LinkMedia supported by FAPEMIG/INRIA in which technical discussion and an outline of the STIC-AMSUD were started. Moreover, in 2013 Silvio started a new cooperation with EFREI/France, which is supported by CAPES/BRAFITEC. In 2015 and 2016, two work missions to PRISMA supported by STIC-AMSUD helps to identify research lines for joint works that will be coped with the submitted STIC-AMSUD 2017

Present main activities and their relationship with the project's main goal.

IRISA/LinkMEDIA, UCHILE/DCC and PUC Minas/VIPLAB have current activities in several areas of multimedia and representation related to TRANSFORM. LinkMEDIA is a leader in multimedia indexing for efficient retrieval, which constitutes a key technology for the project. LinkMEDIA also has strong skills in image description and database management. LinkMEDIA has also been active, along with PUC Minas and UFMG, in a number of international multimedia benchmarks (TREVid, MediaEval) which will serve as a frame for the validation of the research developed within TRANSFORM. UCHILE/DCC is leader in similarity search, in particular for 3D content.

The project will build upon existing results at LinkMEDIA and VIPLAB, exploring new approaches in relations with the project's partners while benefiting from the experience acquired on the subject at hand. The collaboration between UFMG, PUC Minas and IRISA started 5 years ago on a project in multimedia content analysis, and the collaboration between all teams UCHILE, UFMG, PUC Minas and IRISA started 3 years ago on a project in motif discovery. The current project will undoubtedly strengthen the links between the teams and foster new common research directions.

Perspectives of continuing collaboration after project financing is over

In 2017, it will be submitted a project related to this proposal to "CAPES/COFECUB", in order to establish an agenda from 2018 to 2022 between Guillaume Gravier, Simon Malinowski, Silvio Jamil F. Guimarães, Arnaldo de Albuquerque Araújo. During this project, we expect co-supervision of PhD and Master students and organize some workshops.

Apart from existing collaborations between the teams, TRANSFORM targets significant contributions on an emerging topic in multimedia, which requires a wide range of competences. By fostering common research axes and promoting cross-fertilization, TRANSFORM will likely give birth to collaborations beyond the project. Upon success, a new project proposal will be considered.

Relevance to the call

In this section, we highlight some aspect of our proposal in order to show its relevance to the call.

Innovative nature

TRANSFORM targets indexing and retrieval from transformations of multimedia data taking into account temporal information. Thus, we tackle a new problem with high visibility. Moreover, the expertise of all teams, which can be considered a small group, will make it possible for a strong impact on the community with medium term potential for innovation and new applications.

The creation of new international collaborations

Reinforcement of the France-Brazil-Chile collaboration (visits, student exchanges, possibly joint Ph.D. theses). So far, the relationship involving this team could be verified in terms of some joint papers and ongoing joint PhD thesis.

The capacity to generate effective regional cooperation

Reinforcement of the France-Brazil-Chile collaboration (visits, student exchanges, possibly joint Ph. D. theses) can generate effective results since the teams are complementary. The experience of regional coordinators in cooperation projects will facilitate the discussions and progress of this proposal.

Complementarity of the participants

All participants have a strong background in multimedia content analysis. IRISA has background work on multimedia content indexing and time series, along with evaluation related expertise. PUC Minas/UFMG provide its expertise in mid-level representation and content based analysis. UCHILE/DCC provides expertise in multimedia indexing and information retrieval. Partners share a common core expertise on multimedia data thus favoring a quick start of the project. Yet, each partner brings its own complementary competences to the core expertise, the whole providing all the necessary knowledge that no single partner has.

The mutual transfer of expertise and innovation

We expect each partner to acquire skills in the area of expertise that are only covered by other partners. This will be made possible by regular visits of scientists and student exchanges.

Background of the research groups

All groups have a strong background in multimedia processing and are deeply rooted in the multimedia community.

The participation of postdoctoral students or young researchers in the exchanges

The Brazilian team is composed by senior and young researchers. Moreover, there are the presence of several PhD and master students in all labs. More specifically, for the PUCMINAS/VIPLAB, two master students will defense their dissertations for starting their PhD still 2017 in subjects directly related to TRANSFORM and they will stay sometime at IRISA. For the UCHILE/DCC one PhD is still in the project. Moreover, we expected that one of members of the Brazilian stay some time in UCHILE/DCC for a post-doctoral mission.

B8. International referees

Suggest names of at least 3 international referees to evaluate the project. These researchers should not be connected to people in the project.

1- Benoit Huet, EURECOM, France
2- Remco C. Veltkamp, Utrecht University, Netherlands
3- Alain Rakotomamonjy, INSA Rouen, France

Names of referees who should not review this project in your opinion (optional)

1- Eamon Keogh

Bibliography

[Bailly16] Adeline Bailly, Simon Malinowski, Romain Tavenard, Laetitia Chapel, Thomas Guyet. Dense Bag-of-Temporal-SIFT-Words for Time Series Classification. Advanced Analysis and Learning on Temporal Data, Springer, 2016, 978-3319444116.

[Baydogan15] Mustafa G. Baydogan and George Runger. Learning a symbolic representation for multivariate time series classification. Data Mining and Knowledge Discovery, 29(2):400–422, 2015.

[Baydogan13] Mustafa G. Baydogan, George Runger, and Eugene Tuv. A Bag-of-Features Framework to Classify Time Series. IEEE Transactions on Pattern Analysis and Machine Intelligence, 35(11):2796–2802, 2013.

[Bustos05] Benjamin Bustos, Daniel Keim, Dietmar Saupe, Tobias Schreck, and Dejan Vranic. Feature-based similarity search in 3D object databases. ACM Computing Surveys, 37(4):345-387, 2005.

[Cai04] Cai, Y., Ng, R.: Indexing spatio-temporal trajectories with chebyshev polynomials. In: proceedings of the 2004 ACM SIGMOD International Conference on Management of Data. SIGMOD '04, 2004

[Eitz10] Mathias Eitz, Kristian Hildebrand, Tamy Boubekeur, and Marc Alexa. An evaluation of descriptors for large-scale image retrieval from sketched feature lines. Computer & Graphics 34(5):482—498, 2010.

[Keogh05] Eamonn Keogh and Chotirat Ann Ratanamahatana. Exact indexing of dynamic time warping Knowl Inf Syst (2005) 7: 358. doi:10.1007/s10115-004-0154-9

[Lemire09] Daniel Lemire, Faster retrieval with a two-pass dynamic-time-warping lower bound, Pattern Recognition, Volume 42, Issue 9, September 2009

[Li14] Bo Li, Yijuan Lu, Afzal Godil, Tobias Schreck, Benjamin Bustos, Alfredo Ferreira, Takahiko Furuya, Manuel Fonseca, Henry Johan, Takahiro Matsuda, Ryutarou Ohbuchi, Pedro Pascoal, and Jose Saavedra. A comparison of methods for sketch-based 3D shape retrieval. Computer Vision and Image Understanding 119:57-80, 2014.

[Lin12] Jessica Lin, Rohan Khade, and Yuan Li. Rotation-invariant similarity in time series using bag-of-patterns representation. International Journal of Information Systems, 39:287–315, 2012.

[Lods17] A. Lods, S. Malinowski, R. Tavenard and L. Amsaleg. Learning DTW-preserving shapelets. Submitted to Intelligent Data Analysis Symposium (IDA 17)

[Lowe04] David G. Lowe. Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision, 60(2):91–110, 2004

[Ngiam11] J. Ngiam, A. Khosla, M. Kim, J. Nam, H. Lee, and A. Y. Ng. Multimodal deep learning. In Intl. Conf. on Machine Learning, 2011

[Pet17] Tan, C.W., Webb, G.I., Petitjean, F.: Indexing and classifying gigabytes of time series under time warping. In: Proceedings of the SIAM International Conference on Data Mining. SDM '17

[Rak12] Rakthanmanon, T., Campana, B., Mueen, A., Batista, G., Westover, B., Zhu, Q., Zakaria, J., Keogh, E.: Searching and mining trillions of time series subsequences under dynamic time warping. In: Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '12

[Saavedra14] Jose Saavedra and Benjamin Bustos. Sketch-based image retrieval using keyshapes. Multimedia Tools and Applications 73(3):2033-2062, 2014

[Salvador07] Stan Salvador and Philip Chan. Toward accurate dynamic time warping in linear time and space. *Intell. Data Anal.* 11, 5 (October 2007), 561-580.

[Schafer12] Patrick Schäfer and Mikael Högqvist, SFA: a symbolic fourier approximation and index for similarity search in high dimensional datasets. EDBT, ACM. ISBN: 978-1-4503-0790-1 , 2012

[Schafer14] Patrick Schäfer: The BOSS is concerned with time series classication in the presence of noise. Data Mining and Knowledge Discovery 29(6), 2014

[Senin13] Pavel Senin and Sergey Malinchik. SAX-VSM: Interpretable Time Series Classifi- cation Using SAX and Vector Space Model. Proceedings of the IEEE International Conference on Data Mining, pages 1175–1180, 2013.

[Sipiran15] Ivan Sipiran, Benjamin Bustos, Tobias Schreck, Alex M. Bronstein, Sungbin Choi, Long Lai, Haisheng Li, Roee Litman, and Li Sun. Scalability of non-rigid 3D shape retrieval. In Proc. 8th Eurographics Workshop on 3D Object Retrieval (3DOR'15), pages 121-128. Eurographics Association, 2015.

[Smeulders00] Arnold W.M. Smeulders, Marcel Worring, Simone Santini, Amarnath Gupta, and Ramesh Jain. Content-Based Image Retrieval at the End of the Early Years. IEEE Transactions on Pattern Analysis and Machine Intelligence , 2000

[Shape] ShapeNet. https://www.shapenet.org

[Srivastava12] N. Srivastava and R. Salakhutdinov. Learning representations for multimodal data with deep belief nets. In Intl. Conf. on Machine Learning, 2012.

[Tangelder08] Johan W. H. Tangelder, Remco C. Veltkamp. A survey of content based 3D shape retrieval methods. Multimedia Tools and Applications 39:441—471, 2008.

[Tran16] Ti Quynh Nhi Tran, Hervé Le Borgne, and Michel Crucianu. Aggregating image and text quantized correlated components. In *IEEE Intl. Conf. on Computer Vision and Pattern Recognition*, 2016.

[Vukotić16] V. Vukotić, C. Raymond and G. Gravier. Bi-directional joint representation learning with symmetrical deep neural networks for multimodal and crossmodal applications. In ACM Intl. Conf. in Multimedia Retrieval, 2016.

[Vukotić17] V. Vukotić, C. Raymond and G. Gravier. GANs for Multimodal Representation Learning in Video Hyperlinking. In ACM Intl. Conf. in Multimedia Retrieval, 2017.

[Wang13] Jim Wang, Ping Liu, Mary F.H. She, Saeid Nahavandi, and Addas Kouzani. Bagof-Words Representation for Biomedical Time Series Classification. Biomedical Signal Processing and Control, 8(6):634–644, 2013.

[Weston10] J. Weston, S. Bengio, and N. Usunier. Large scale image annotation: learning to rank with joint word-image embeddings. Machine learning, 81(1):21–35, 2010.

[Xie09] Jierui Xie and Mandis Beigi. A Scale-Invariant Local Descriptor for Event Recognition in 1D Sensor Signals. In Proceedings of the IEEE International Conference on Multimedia and Expo, pages 1226–1229, 2009.

[Yoon10] Sang Min Yoon, Maximilian Scherer, Tobias Schreck, Arjan Kuijper. Sketch-based 3D model retrieval using diffusion tensor fields of suggestive contours. In Proc. ACM Multimedia, pages 193-200. ACM, 2010.

C. Project Budget

Project title: TRANSFORM

Participating institutions: IRISA/LINKMEDIA, UFMG/DCC/NPDI, PRISMA/DCC, PUC/VIPLAB

The STIC-AmSud program **funds travel expenses** (air tickets and *per diem*) to researchers in research missions and workshops. Moreover, according to call for proposed published by CAPES/BRASIL, it is possible to ask funds for PhD Students. In this proposal, we ask for two internships for 12-month (one for each year). This demand is inserted in other funding.

C1. First year (2018)

Planned missions – Year 1

Researcher	Status (student, junior, senior)	Institution	Origine	Destination	Planned date	Duration (max. 30d)	Estimated cost of the trip (€)	Trip funding institution[1]	Estimate of total *per diem* (€)	Mission funding institution (per diem)[2]	Mission objectives
Guillaume Gravier	Senior	IRISA/LINKMEDIA	Rennes	Belo Horizonte	Mar. 2018	12	1800	CNRS	2592	CNRS	Plenary kick-off meeting
Laurent Amsaleg	Senior	IRISA/LINKMEDIA	Rennes	Belo Horizonte	Mar. 2018	5	1800	CNRS	1080	CNRS	Plenary kick-off meeting
Simon Malinowski	Senior	IRISA/LINKMEDIA	Rennes	Belo Horizonte	Mar. 2018	12	1800	CNRS	2592	CNRS	Plenary kick-off meeting
Benjamin Bustos	Senior	UCHILE/DCC	Santiago	Belo Horizonte	Mar. 2018	10	550	CONICYT	1000	CONICYT	Plenary kick-off meeting
Silvio Guimarães	Senior	VIPLAB/PUC Minas	Belo Horizonte	Santiago	Jun. 2018	7	550	CAPES	1155	CAPES	Joint work on mid-level representation for temporal data
Wladmir Brandão	Junior	VIPLAB/PUC Minas	Belo Horizonte	Santiago	Nov. 2018	7	550	CAPES	1155	CAPES	Joint work on indexing and retrieval
Zenilton K. G. do Patrocínio Jr.	Senior	VIPLAB/PUC Minas	Belo Horizonte	Rennes	Oct. 2018	7	1500	CAPES	2166	CAPES	Joint work on indexing and retrieval and mid-level representation

PhD Student 1	Student	UFMG/DCC/NPDI	Belo Horizonte	Santiago	Oct. 2018	-	-	Other	-	Other	Joint work on time series
PhD Student 2	Student	VIPLAB/PUC Minas	Belo Horizonte	Rennes	Oct. 2018	-	-	Other	-	Other	Joint work on multimodal representation spaces

CONSOLIDATED BUDGET: Year 1

Funding requested to the STIC-AmSud Program
Estimated costs (€)

	CNRS France	CAPES Brazil	CONICYT Chile	Total requested funding to STIC-AmSud	Other funding	TOTAL
A- Travel costs (air tickets)	5400	2600	550	8550	4400	12950
B- Maintenance costs (*per diem)*	6264	4476	1000	11740	31200	42940
TOTAL	11664	7076	1550	20290	35600	55890

C2. Second year (2019)

Second year funding depends on approval of intermediate progress report.

Planned missions – Year 2

Researcher	Status (student, junior, senior)	Institution	Origin	Destination	Planned date	Duration (max. 30d)	Estimated cost of the trip (€)	Trip funding institution[1]	Estimate of total *per diem* (€)	Mission funding institution (per diem)[2]	Mission objectives
Guillaume Gravier	Senior	IRISA/LINKMEDIA	Rennes	Santiago	May 2019	14	1900	CNRS	2795	CNRS	Joint work on symbolic method and efficient algorithms
Laurent Amsaleg	Senior	IRISA/LINKMEDIA	Rennes	Santiago	May 2019	14	1900	CNRS	2795	CNRS	Joint work on temporal sequence indexing
Simon Malinowski	Senior	IRISA/LINKMEDIA	Rennes	Santiago	Sep. 2019	14	1900	CNRS	2795	CNRS	Joint work on indexing structures
Benjamin Bustos	Senior	UCHILE/DCC	Santiago	Rennes	Mar. 2019	14	1900	CONICYT	1529	CONICYT	Joint work on indexing structures
Silvio Guimarães	Senior	PUCMINAS/VIPLAB	Belo Horizonte	Rennes	Jan. 2019	10	1500	CAPES	3094	CAPES	Joint work on indexing structures and efficient algorithms
Kleber Jacques Ferreira de Souza	Junior	VIPLAB/PUC Minas	Belo Horizonte	Santiago	Nov. 2018	7	550	CAPES	1155	CAPES	Joint work on mid-level representation for temporal data
Arnaldo de Albuquerque Araújo	Senior	UFMG/DCC/NPDI	Belo Horizonte	Rennes	Jul. 2019	7	1500	CAPES	2166	CAPES	Join work on mid-level representation

Post-doc 1	Junior	PUCMINAS/VIPLAB	Belo Horizonte	Santiago	Dec. 2019	-	-	Other	-	Other	Joint work on indexing and retrieval
PhD Student 3	Student	PUCMINAS/VIPLAB	Belo Horizonte	Rennes	Fev. 2019	-	-	Other	-	Other	Joint work on mid-level representation for 3D shape

CONSOLIDATED BUDGET: Year 2

Funding requested to the STIC-AmSud Program
Estimated costs (€)

	CNRS France	CAPES Brazil	CONICYT Chile	Total requested funding to STIC- AmSud	Other funding	TOTAL
A- Travel costs (air tickets)	5700	3550	1900	11150	4400	15550
B- Maintenance costs (*per diem)*	8385	6415	1529	16328	24000	40328
TOTAL	14085	9965	3429	27478	28400	55878

C3. BUDGET TOTALS

	Year 1	Year 2	Total
Funding requested to MAE (France)	**0**	**0**	**0**
Funding requested to INRIA (France)	**0**	**0**	**0**
Funding requested to CNRS (France)	**11664**	**14085**	**25749**
Funding requested to Institut MINES-TELECOM (France)	**0**	**0**	**0**
Funding requested to MINCYT (Argentina)	**0**	**0**	**0**
Funding requested to CAPES (Brazil)	**7076**	**9965**	**17041**
Funding requested to CONICYT (Chile)	**1550**	**3429**	**4979**
Funding requested to CONACYT (Paraguay)	**0**	**0**	**0**
Funding requested to CONCYTEC (Peru)	**0**	**0**	**0**
Funding requested to ANII (Uruguay)	**0**	**0**	**0**
Matching funds from the partners	**0**	**0**	**0**
Other sources	**35600**	**28400**	**64000**
TOTAL	**55890**	**55878**	**111768**

ANNEX

1/ Personal data

Name: Guillaume Gravier

Birth date: Feb. 12, 1971

Professional address (with telephone and e-mail): IRISA, 35042 Rennes Cedex

Current job title and size of the research group: Directeur de recherche (senior research scientist), CNRS, head of the Linkmedia research group (9 permanent researchers)

2/ Highest obtained degree (with indication of place and date)

Ph. D., Telecom Paris, Jan. 2000.

HDR, Univ. Rennes 1, Sep. 2009.

3/ Professional activity in the last 5 years

Since July 2014: CNRS Senior Research scientist, IRISA, head of the Linkmedia team – Multimedia analytics, multimedia data mining

From Jan. 2011 to June 2014: CNRS Research scientist, IRISA, Texmex team – Multimedia, multimodal modeling, spoken content processing, data mining, indexing

Areas of expertise: multimedia; audio, speech and language processing; probabilistic modeling; pattern recognition; machine learning; classification; fusion; multimedia indexing; information retrieval; data mining;

4/ Other duties/ positions

Cofounder and president of the ISCA Special Interest Group on Speech and Language in Multimedia
Associate editor of IEEE Trans. On Multimedia
Member of the steering committee of the MediaEval international benchmark

5/ Awards, fellowships and external recognition

Steering committee and Program coordinator of Interspeech 2013
Technical program chair IEEE Workshop in Content-Based Multimedia Indexing 2016, ACM Intl. Conf. on Multimedia 2019
Area chair ACM Intl. Conf. on Multimedia 2016
General chair of the ISCA/IEEE Workshop on Speech, Language and Audio in Multimedia, 2013 and 2015

6/ Ongoing funded research projects with dates, titles, sources of funding

Linking media in acceptable hypergraphs (LABEX CominLabs, 2014-2018)
Knowledge-mediated integrated data and content analytics: The case of data journalism (Inria Project Lab, 2017-2021)

7/ Projects approved in the least 5 years

Quaero (OSEO,2008-2013)
STIC AmSud Maximum (CNRS, 2014-2014)
Inria Associate Team MOTIF (Inria, 2015-2016)

8/ Publications

8.1 – Highlight the most important publications related to the project theme

R. Bois, V. Vukotić, A.-R. S imon, R. Sicre, C. Raymond, P. Sébillot, G. Gravier. Exploiting multimodality in video hyperlinking to improve target diversity. Proc. Intl. Workshop on Multimedia Modeling, 2017
V. Vukotić, C. Raymond, G. Gravier. Generative Adversarial Networks for Multimodal Representation Learning in Video Hyperlinking. ACM Intl.\ Conf.\ on Multimedia Retrieval, 2017
V. Vukotić, C. Raymond and G. Gravier. Bidirectional joint representation learning with symmetrical deep neural networks for multimodal and crossmodal applications. ACM Intl. Conf. on Multimedia Retrieval, 2016
R. Bois, V. Vukotic, R. Sicre, C. Raymond, G. Gravier. IRISA at TRECVid2016: Crossmodality, Multimodality and Monomodality for Video Hyperlinking. Working Notes of the TRECVid 2016 Workshop, 2016
M. Larson, M. Soleymani, G. Gravier, B. Ionescu, G. J.F. Jones. The Benchmarking Initiative for Multimedia Evaluation : MediaEval 2016. Invited report in IEEE Multimedia, 24(1) :93–96, 2017

8.2 – Publications in cooperation with the project partners

G. Fonseca, G. Sargent, S. J. Guimaraes and G. Gravier. Tag Propagation Approaches within Speaking Face Graphs for Multimodal Person Discovery. IEEE Workshop on Content-Based Multimedia Retrieval, 2017
Nam Le, H. Bredin, G. Sargent et al. Towards large scale multimedia indexing: A case study on person discovery in broadcast news. IEEE Workshop on Content-Based Multimedia Retrieval, 2017
C. Dos Santos, E. Kijak, G. Gravier, W. R. Schwartz. Partial Least Squares for Face Hashing. Neurocomputing, Elsevier, 213, pp.34-47, 2016.
H. Batista Da Silva, Z. Patrocino Jr., G. Gravier, L. Amsaleg, A. de A. Araújo, et al. Near-Duplicate Video Detection Based on an Approximate Similarity Self-Join Strategy. IEEE International Workshop on Content-based Multimedia Indexing, 2016.
G. Sargent, G. Barbosa de Fonseca, I. Lyon Freire, R. Sicre, Z. do Patrocínio Jr., et al.. PUC Minas and IRISA at Multimodal Person Discovery. Working Notes Proceedings of the MediaEval Workshop, 2016
C. dos Santos Jr., G. Gravier, W. R. Schwartz. SSIG and IRISA at Multimodal Person Discovery. Working Notes Proceedings of the MediaEval Workshop, 2015

9/ Theses oriented and post-doctoral fellows supervised

I supervised 20 Ph. D. students in total, among which 15 already defended. I am currently supervising 5 PhDs and 2 post-docs.

9.1 – Finished/defended in the last 5 years

C. Penet. Multimodal video content-based analysis for VoD. Ph. D. 2013.
S. Ziegler. Speech recognition driven by phonetic islands. Ph. D. 2014.
A. Simon. Hierarchical topic segmentation of spoken content. Ph. D. 2015.
B. Qu. Discovery of a TV program structure from a collection. Ph. D. 2015.

9.2 – Ongoing

R. Bois. Navigable multimedia hypergraphs: construction and exploitation. Defense scheduled for Sep. 2017.

Vedran Vukotić. Deep neural architectures for multimodal representation learning. Defense scheduled for Sep. 2017.

Ricardo Carlini Sperandio. Unsupervised motif discovery in multimedia time series. Defense scheduled in 2019.

Mikail Demirdelen. Multimedia multimodal multidocument summarization adaptive to user needs. Defense scheduled in 2019.

Oriane Simeoni. Invariance and supervision in visual learning. Defense scheduled in 2019.

ANNEX

1/ Personal data

Name: Simon Malinowski

Birth date: Jun. 29, 1982

Professional address (with telephone and e-mail):

IRISA, 35042 Rennes Cedex - +33299842557 – simon.malinowski@irisa.fr

Current job title and size of the research group: Assistant Professor, Univ. Rennes 1, LINKMEDIA research group (15 permanent researchers)

2/ Highest obtained degree (with indication of place and date)

Ph. D., Univ. Rennes 1, Dec. 2008.

3/ Professional activity in the last 5 years

Since Sep. 2014: Assistant professor, Univ. Rennes 1/ IRISA, Linkmedia Team – Data mining and machine learning applied to temporal data – Time series classification and indexing

From Sep. 2013 to Aug. 2014: Post-doctoral researcher, ENSMM, Besançon, France – Temporal models for remaining useful life prediction

From Sep. 2012 to Aug. 2013: Assistant Professor (temporary) , AgroCampus Ouest, France – Symbolic time series representation for time series mining

Areas of expertise: pattern recognition; machine learning; time series classification; time series indexing; data mining;

4/ Other duties/ positions

N/A

5/ Awards, fellowships and external recognition

General chair of the 2nd Workshop on Advanced Analysis and Learning on Temporal Data, 2016, collocated with ECML/PKDD

6/ Ongoing funded research projects with dates, titles, sources of funding

N/A

7/ Projects approved in the least 5 years

N/A

8/ Publications

8.1 – Highlight the most important publications related to the project theme

1. A. Bailly,S. Malinowski, R. Tavenard, T. Guyet and L. Chapel Dense Bag-of-Temporal-SIFT-Words for time series classification - Lecture Notes in Artificial Intelligence - 2016
2. A. Le Guennec, S. Malinowski and R. Tavenard - Data Augmentation for Time Series Classification using Convolutional Neural Networks - ECML/PKDD Workshop on Advanced Analytics and Learning on Temporal Data - 2016
3. R. Tavenard, S. Malinowski, L. Chapel, A. Bailly, H. Sanchez and B. Bustos – Efficient temporal kernels between feature sets for time series classification, submitted to ECML/PKDD 2017
4. A. Lods, S. Malinowski, R. Tavenard and L. Amsaleg – Learning DTW-preserving shapelets – submitted to Inetlligence Data Analysis Symposium 2017

8.2 – Publications in cooperation with the project partners

1. R. Tavenard, S. Malinowski, L. Chapel, A. Bailly, H. Sanchez and B. Bustos – Efficient temporal kernels between feature sets for time series classification, submitted to ECML/PKDD 2017

9/ Theses oriented and post-doctoral fellows supervised

I supervised 1Ph. D. student during my post-doctoral research in Besancon.

9.1 – Finished/defended in the last 5 years

R. Khelif. Estimation du RUL par des approches basées sur l'expérience: de la donnée vers la connaissance.

9.2 – Ongoing

A. Lods – Master Student – Time series indexing

1/ Personal data

Name: Benjamin Bustos
Birth date: 02/06/1977
Professional address (with telephone and e-mail):
Beauchef 851, Edificio Norte, 3er piso, Santiago, Chile
Phone: +56 229784969, E-mail: bebustos@dcc.uchile.cl
Current job title and size of the research group:
- Associate Professor, Department of Computer Science, University of Chile.
- PRISMA Research Group (two senior researchers, six Ph.D. Students).

2/ Highest obtained degree (with indication of place and date)

Doctor of Natural Sciences, Universität Konstanz, Germany, October 2006.

3/ Professional activity in the last 5 years

Academic at the Department of Computer Science, University of Chile.

4/ Other duties/ positions

N/A.

5/ Awards, fellowships and external recognition

- (2011) Yahoo! Faculty Research and Engagement award in the form of an Unrestricted Gift (USD 10,000.00) for the Project "Exploiting visual-semantic graphs for Web image retrieval and automatic image tagging" (Yahoo! collaborator: Dr. Barbara Poblete).

- (2010) Distinguished Teaching Award 2010, given by the Department of Computer Science, University of Chile.

- (2010) Yahoo! Faculty Research and Engagement award in the form of an Unrestricted Gift (USD 10,000.00) for the Project "Web multimedia content" (Yahoo! collaborator: Dr. Barbara Poblete).

- (2007) Annual award given by the Chilean Computer Science Society (SCCC) to a Chilean researcher with a recent Ph.D. in Computer Science. It included a plenary talk at Jornadas Chilenas de Computación (Chilean Congress of Computer Science) with the title "Efficient content search in multi-metric spaces".

6/ Ongoing funded research projects with dates, titles, sources of funding

- (2014-2017) Young researcher at the Millennium Nucleus Center for Semantic Web Research, Grant NC 120004, Chile.

7/ Projects approved in the least 5 years

- (2014-2016) Main researcher in CONICYT-FONDECYT (Chile) Project 1140783, Efficient 3D Shape Matching Using Local Descriptors.
- (2014-2015) Main researcher in Stic-AmSud Project 14STIC-04, Unsupervised Multimedia Content Mining.
- (2011-2013) Main researcher in CONICYT-FONDECYT (Chile) Project 1110111, Interest Points Detection Algorithms for 3D Shape Matching.
- (2011-2013) Main researcher in FONDEF (Chile) Project D09I1185, Real-Time Scalable Observatories in the Web.

8/ Publications

8.1 – Highlight the most important publications related to the project theme

1. Ivan Sipiran, Jakub Lokoc, Benjamin Bustos, and Tomas Skopal. Scalable 3D shape retrieval using local features and the Signature Quadratic Form Distance. To appear in The Visual Computer. Springer.
2. Rodrigo Schulz, Pablo Guerrero, and Benjamin Bustos. Directed Curvature Histograms for Robotic Grasping. In Proc. 10th Eurographics Workshop on 3D Object Retrieval (3DOR'17), pages 113-120. Eurographics Association, 2017.
3. Robert Gregor, Andreas Lamprecht, Ivan Sipiran, Tobias Schreck, and Benjamin Bustos. Empirical evaluation of dissimilarity measures for 3D object retrieval with application to multi-feature retrieval. In Proc. 13th International Workshop on Content-Based Multimedia Indexing (CBMI'15), pages 1-6. IEEE, 2015.
4. Ivan Sipiran, Benjamin Bustos, Tobias Schreck, Alex M. Bronstein, Sungbin Choi, Long Lai, Haisheng Li, Roee Litman, and Li Sun. Scalability of non-rigid 3D shape retrieval. In Proc. 8th Eurographics Workshop on 3D Object Retrieval (3DOR'15), pages 121-128. Eurographics Association, 2015.
5. Ivan Sipiran, Rafael Meruane, Benjamin Bustos, Tobias Schreck, Bo Li, Yijuan Lu, and Henry Johan. A benchmark of simulated range images for partial shape retrieval. The Visual Computer 30(11):1293-1308. Springer, 2014.
6. Bo Li, Yijuan Lu, Afzal Godil, Tobias Schreck, Benjamin Bustos, Alfredo Ferreira, Takahiko Furuya, Manuel Fonseca, Henry Johan, Takahiro Matsuda, Ryutarou Ohbuchi, Pedro Pascoal, and Jose Saavedra. A comparison of methods for sketch-based 3D shape retrieval. Computer Vision and Image Understanding 119:57-80. Elsevier, 2014.
7. Ivan Sipiran and Benjamin Bustos. Key-components: Detection of salient regions on 3D meshes. The Visual Computer 29(12):1319-1332. Springer, 2013.
8. Ivan Sipiran, Benjamin Bustos, and Tobias Schreck. Data-aware 3D partitioning for generic shape retrieval. Computer & Graphics 37(5):460-472. Elsevier, 2013.
9. Zhouhui Lian, Afzal Godil, Benjamin Bustos, Mohamed Daoudi, Jeroen Hermans, Shun Kawamura, Yukinori Kurita, Guillaume Lavoue, Hien Van Nguyen, Ryutarou Ohbuchi, Yuki Ohkita, Yuya Ohishi, Fatih Porikli, Martin Reuter, Ivan Sipiran, Dirk Smeets, Paul Suetens, Hedi Tabia, Dirk Vandermeulen. A comparison of methods for non-rigid 3D shape retrieval. Pattern Recognition, 46(1):449-461. Elsevier Science, 2013.
10. Ivan Sipiran and Benjamin Bustos. A fully hierarchical approach for finding correspondences in non-rigid shapes. In Proc. 14th IEEE International Conference on Computer Vision (ICCV'13), pages 817-824. IEEE, 2013.
11. Benjamin Bustos, Tobias Schreck, Michael Walter, Juan Manuel Barrios, Matthias Schaefer, and Daniel Keim. Improving 3D similarity search by enhancing and combining 3D descriptors. Multimedia Tools and Applications, 58(1):81-108. Springer, 2012.
12. Jose Saavedra, Benjamin Bustos, Tobias Schreck, Sang Min Yoon, and Maximilian Scherer. Sketch-based 3D model retrieval using keyshapes for global and local representation. In Proc. 5th Eurographics Workshop on 3D Object Retrieval (3DOR'12), pages 47-50. Eurographics Association, 2012.
13. Tobias Schreck, Maximilian Scherer, Michael Walter, Benjamin Bustos, Sang Min Yoon, and Arjan Kuijper. Graph-based combinations of fragment descriptors for improved 3D object retrieval. In Proc. ACM Multimedia Systems (MMSYS'12), pages 23-28. ACM, 2012.
14. Ivan Sipiran and Benjamin Bustos. Harris 3D: A robust extension of the Harris operator for interest point detection on 3D meshes. The Visual Computer, 27(11):963-976. Springer, 2011.

8.2 – Publications in cooperation with the project partners
1. Carlos Alberto Fraga Pimentel Filho, Benjamin Bustos, Arnaldo de Albuquerque Araujo, Silvio Jamil Ferzoli Guimaraes. Combining pixel domain and compressed domain index for sketch based image retrieval. To appear in Multimedia Tools and Applications. Springer.

9/ Theses oriented and post-doctoral fellows supervised

9.1 – Finished/defended in the last 5 years

Ph.D. theses:
1. Guillermo Cabrera (co-guided with Nancy Hitschfeld). Extraction and Classication of Objects from Astronomical Images in the Presence of Labeling Bias. University of Chile, 2015.
2. Ivan Sipiran. Local Features for Shape Matching and Retrieval. University of Chile, 2014.
3. Juan Manuel Barrios. Content-based Video Copy Detection. University of Chile, 2013.
4. José Saavedra. Image Descriptions for Sketch Based Image Retrieval. University of Chile, 2013.

M.Sc. theses:
1. Rafael Meruane. Evaluación de Algoritmos de Matching Parcial en Objetos 3D Utilizando un Escáner 3D Óptico Portátil. University of Chile, 2015.
2. Victor Sepúlveda. Indexamiento en Espacios No-Métricos. University of Chile, 2012.

9.2 – Ongoing

Ph.D. theses:
1. Heider Sánchez. Event Detection in Streaming Multivariate Time Series. University of Chile.
2. Arniel Labrada. Retrieval of 3D Objects in 2D Images Using Sketch Abstraction. University of Chile.

M.Sc. theses:
1. Camila Alvarez (co-guided with Jose Saavedra). Automatic clothing labeling of outdoor images.
3. Catalina Espinoza (co-guided with Marcela Munizaga). Caracterización y reconocimiento de usuarios a través de la observación de su movilidad en transporte público. University of Chile.
4. Felipe Hernandez (co-guided with Marcela Munizaga). Estructura Espacial Urbana de Movilidad desde Datos Masivos de Transporte Público en Santiago de Chile. University of Chile.
5. Rodrigo Schulz (co-guided with Pablo Guerrero). Síntesis de Agarres para Grasping Robótico a Partir de Nubes de Puntos 3D. University of Chile.
6. Sebastian Ferrada (co-guided with Aidan Hogan). IMGpedia: A LArge-Scale Knowledge-Base to Perform Visuo-Semantic Queries over Wikimedia Common Images. University of Chile.

1/ Personal data

Name: Silvio Jamil Ferzoli Guimarães
Birth date: 23/06/1974
Professional address (with telephone and e-mail):
Rua Walter Ianni, 255 – São Gabriel – Belo Horizonte – Minas Gerais - Brasil
Phone: +55 31 3439 5204, E-mail: sjamil@pucminas.br
Current job title and size of the research group:
Associate Professor, Department of Computer Science, PUC Minas, Brazil

VIPLAB Research Group (three senior researchers, nine MSc. Students, several undergraduate students).

2/ Highest obtained degree (with indication of place and date)

Doctor in Computer Science, UFMG, Brazil, Mars 2003 (co-tutelle with Université Marne-la-Vallee, France)

3/ Professional activity in the last 5 years

Academic at the Department of Computer Science, PUC Minas, Brazil.

4/ Other positions

Associate researcher at ESIEE/France.

5/ Awards, fellowships and external recognition

(2012) Grant Santander Universidades. (5,000.00 euros) for work mission to UPC/Spain to develop the Project "Image and video analysis using hierarchical clustering methods" (UPC collaborator: Prof. Philippe Salembie and Prof. Ferran Marques).

(2011) Post-doctoral stage at ESIEE/France supported by CAPES/Brazil

(2010) Best paper in WEBMEDIA with the work "Identification and Analysis of Video Subsequence Using Bipartite Graph Matching".

(2017) Grant of researcher at the CNPq/Brazil

6/ Ongoing funded research projects with dates, titles, sources of funding

(2011-2013) Main researcher in FAPEMIG/INRIA (Cooperation project) Project 126, "Video summarization using multimodal Information (TVSum)" in collaboration to Dr. Patrick Gros and Dr. Guillaume Gravier

(2012-2014) Senior researcher in CNPq (Brazil) Project,"Information retrieval in large image and video datasets" in collaboration to Prof. Arnaldo de Albuquerque Araújo.

(2012-2014) Senior researcher in CAPES/CDEFI (Cooperation project) Project BRAFITEC BRasil France Ingénieur TEChnologie "FOR TIC! FORmação em Tecnologia da Informação e da Comunicação".

7/ Projects approved in the least 5 years

(2008-2011) Senior researcher in CAPES/COFECUB Project intitled "Interactive analysis for content based multimedia data"

(2013-2014) Senior researcher in STIC-AMSUD Project intitled "Unsupervised Multimedia Content Mining"

(2017-2018) Senior researcher in FAPEMIG-PPM Project intitled "Representação hierárquica de informações multimídia para análise de imagem e vídeo"

(2015-2017) Senior researcher in CAPES/PVE Project intitled "Segmentação hierárquica de imagem e vídeo baseado em análise de grafos"

8/ Publications

8.1 – Highlight the most important publications related to the project theme

ALMEIDA, R. P. ; PATROCÍNIO JR., ZENILTON K.G. ; GUIMARÃES, SILVIO JAMIL F. . Exploring Quantization Error to Improve Human Action Classification. In: International Joint Conference on Neural Networks (IJCNN 2017), 2017, Alaska. 2017 International Joint Conference on Neural Networks (IJCNN 2017), 2017 (to appear).

PIMENTEL, C. A. F. ; ARAÚJO, Arnaldo de Albuquerque ; BUSTOS, B. ; GUIMARAES, SILVIO J.F. . Combining pixel domain and compressed domain index for sketch based image retrieval. MULTIMEDIA TOOLS AND APPLICATIONS (DORDRECHT. ONLINE), 2017. (to appear).

BELO, LUCIANA DO SANTOS ; CAETANO JR., CARLOS ANTÔNIO ; PATROCÍNIO JR., ZENILTON KLEBER GONÇALVES DO ; GUIMARÃES, SILVIO JAMIL FERZOLI . Summarizing video sequence using a graph-based hierarchical approach. Neurocomputing (Amsterdam), v. 173, p. 1001-1016, 2016.

CAETANO, CARLOS ; AVILA, SANDRA ; SCHWARTZ, WILLIAM ROBSON ; GUIMARÃES, SILVIO JAMIL F. ; Araújo, Arnaldo de A. . A mid-level video representation based on binary descriptors: A case study for pornography detection. Neurocomputing (Amsterdam), v. 213, p. 102-114, 2016.

SILVA, H. B. ; CAETANO JR., C. A. ; ARAÚJO, Arnaldo de Albuquerque ; CONCEICAO, D. V. ; ALMEIDA, R. P. ; FONSECA, G. B. ; PATROCINIO JR, Zenilton Kleber G Do ; GUIMARÃES, SILVIO JAMIL F. . Video similarity search by using compact representations. In: The 31st ACM/SIGAPP Symposium on Applied Computing, 2016, Pisa. The 31st ACM/SIGAPP Symposium on Applied Computing, 2016. p. 80-83.

SILVA, H. B. ; ARAÚJO, Arnaldo de Albuquerque ; GRAVIER, G. ; AMSALEG, L. ; PATROCINIO JR, Zenilton Kleber G Do ; GUIMARÃES, SILVIO JAMIL F. . Near-Duplicate Video Detection Based on an Approximate Similarity Self-Join Strategy. In: 14th International Workshop on Content-based Multimedia Indexing - CBMI'16, 2016, Bucarest. 14th International Workshop on Content-based Multimedia Indexing - CBMI'16, 2016.

SARGENT, G. ; FONSECA, G. B. ; FREIRE, I. L. ; SICRE, R. ; PATROCINIO JR, Zenilton Kleber G Do ; GUIMARÃES, SILVIO JAMIL F. ; GRAVIER, G. . PUC Minas and IRISA at Multimodal Person Discovery. In: Working Notes Proceedings of the MediaEval 2016 Workshop, 2016, Amsterdam. Working Notes Proceedings of the MediaEval 2016 Workshop, 2016. v. 1. p. 1-3.

SPERANDIO, RICARDO C. ; PATROCÍNIO, ZENILTON K. G. ; DE PAULA, HUGO B. ; Guimarães, Silvio J. F. . An efficient access method for multimodal video retrieval. Multimedia Tools and Applications (Dordrecht. Online), v. 74, p. 1357-1375, 2015.

PATROCINIO JR., Z. ; GUIMARÃES, SILVIO JAMIL FERZOLI . A Two-Step Video Subsequence Identification based on Bipartite Graph Matching. In: 2012 IEEE International Conference on Systems, Man, and Cybernetics, 2012, Seul

GUIMARÃES, SILVIO JAMIL FERZOLI ; Patrocínio, Zenilton Kleber Gonçalves . Identification of video subsequence using bipartite graph matching. Journal of the Brazilian Computer Society (Impresso), v. 17, p. 175-192, 2011

8.2 – Publications in cooperation with the project partners
OLIVEIRA, LEONARDO S. DE ; PATROCINIO, ZENILTON K.G. DO ; GUIMARAES, SILVIO JAMIL F. ; GRAVIER, GUILLAUME . Searching for Near-Duplicate Video Sequences from a Scalable Sequence Aligner. In: 2013 IEEE International Symposium on Multimedia (ISM), 2013, Anaheim. 2013 IEEE International Symposium on Multimedia. p. 223-226.

SILVA, H. B. ; ARAÚJO, Arnaldo de Albuquerque ; GRAVIER, G. ; AMSALEG, L. ; PATROCINIO JR, Zenilton Kleber G Do ; GUIMARÃES, SILVIO JAMIL F. . Near-Duplicate Video Detection Based on an Approximate Similarity Self-Join Strategy. In: 14th International Workshop on Content-based Multimedia Indexing - CBMI'16, 2016, Bucarest. 14th International Workshop on Content-based Multimedia Indexing - CBMI'16, 2016.

PIMENTEL, C. A. F. ; ARAÚJO, Arnaldo de Albuquerque ; BUSTOS, B. ; GUIMARAES, SILVIO J.F. . Combining pixel domain and compressed domain index for sketch based image retrieval. MULTIMEDIA TOOLS AND APPLICATIONS (DORDRECHT. ONLINE), 2017 (to appear).

SARGENT, G. ; FONSECA, G. B. ; FREIRE, I. L. ; SICRE, R. ; PATROCINIO JR, Zenilton Kleber G Do ; GUIMARÃES, SILVIO JAMIL F. ; GRAVIER, G. . PUC Minas and IRISA at Multimodal Person Discovery. In: Working Notes Proceedings of the MediaEval 2016 Workshop, 2016, Amsterdam. Working Notes Proceedings of the MediaEval 2016 Workshop, 2016. v. 1. p. 1-3.

Gabriel Barbosa da Fonseca, Izabela Lyon Freire, Zenilton Patrocínio Jr and Silvio Jamil F. Guimarãe, Gabriel Sargent, Ronan Sicre and Guillaume Gravier. Tag Propagation Approaches within Speaking Face Graphs for Multimodal Person Discovery. CBMI 2017 (accepted).

9/ Theses oriented and post-doctoral fellows supervised

9.1 – Finished/defended in the last 5 years
- Kleber Jacques (MSc. student) – Graph-based color image simplification
-Aniceto Junior (MSc. student) – Evolutionary computing for improving image segmentation
- Ângelo Magno (MSc. student) – Text extraction in image sequences
- Willer Gomes (MSc. student) – A graph-based static vídeo summarization using global descriptor
- Kleber Jacques (Phd – co-advisor) – Hierarchical graph-based video segmentation
- Luciana Belo (MSc. student) – Graph-based video summarization using multimodal information

- Carlos Antônio Caetano Jr (MSc. student) – Detecção de pornografia em vídeos digitais
- Renata Diniz (MSc. student) – Avaliação de segmentações de vídeo calculadas a partir de mapas de saliência
- Carlos Alberto Fraga Pimentel Filho (Post-doc) - Stochastic watershed cut based on noised topographical surface

9.2 – Ongoing

- Filipe Tório (MSc. student) – Evaluation of morpholocal hierarchies for supervised video segmentation
- Raquel Almeida (MSc. student) – Exploring mid-level representation for human action classification
- Gabriel Fonseca (MSc. student) – Tag Propagation Approaches within Speaking Face Graphs for Multimodal Person Discovery
- Henrique Silva (Phd – co-advisor) – Indexing structures for fast sequence join discovery

ANNEX: CV

1/ Personal data

Name: Wladmir Cardoso Brandão
Birth date: 24/06/1973
Professional address (with telephone and e-mail):
Rua Walter Ianni, 255 – São Gabriel – Belo Horizonte – Minas Gerais - Brasil
Phone: +55 31 3439 5204, E-mail: wladmir@pucminas.br
Current job title and size of the research group:
Associate Professor, Department of Computer Science, PUC Minas, Brazil

IRIS Research Group (four senior researchers, one MSc. Students, several undergraduate students).

VIPLAB Research Group (three senior researchers, nine MSc. Students, several undergraduate students).

2/ Highest obtained degree (with indication of place and date)

Doctor in Computer Science, UFMG, Brazil, November 2013

3/ Professional activity in the last 5 years

Academic at the Department of Computer Science, PUC Minas, Brazil.

4/ Other positions

5/ Awards, fellowships and external recognition

(2014) Microsoft Azure for Research Award Program. (40,000.00 US dolars) to develop the Microsoft RFT Brazil Project, "*Entity-related Information Extraction and Retrieval from the Web*".

 (2017) Best paper in ICEIS with the work "*Formal concept analysis applied to professional social networks analysis*".

6/ Ongoing funded research projects with dates, titles, sources of funding

(07/2016 – 06/2017) Coordinator: Project PUC/FIP 005/2016, *Stemming Baseado em Autômatos para Recuperação de Informação*. Grant PUC 2016/11086. Amount of R$ 6,300.00 for 1 year, starting in July, 2016.

(02/2016 – 01/2018) Membro: Project FAPEMIG 14/2014, *Proposta de uma Metodologia de Busca de Informações e Estimativa de Demanda e Consumo de Clientes Livres Potenciais*. Grant FAPEMIG APQ-03075-14. Amount of R$ 92,000.00 for 2 years, starting in February, 2016.

(02/2016 – 01/2018) Membro: Project FAPEMIG 01/2015, *BigDataClima - Modelos Neurais para Cenários Climáticos por meio de Big Data - Estudo de Caso para Previsão de Eventos*

Extremos na Região Metropolitana de Belo Horizonte. Grant FAPEMIG APQ-00997-15. Amount of R$ 59,220.00 for 2 years, starting in February, 2016.

(02/2016 – 01/2018) Membro: Project FAPEMIG 19/2013, *MASWeb: Modelos, Algoritmos e Sistemas para Web*. Grant FAPEMIG APQ- 01400-14. Amount of R$ 330,835.26 for 2 years, starting in February, 2015.

(11/2014 – 10/2017) Coordinator: Project MCTI/CNPQ/Universal 14/2014, *Extração e Recuperação de Informação sobre Entidades na Web*. Grant CNPq 444156/2014-3. Amount of R$ 14,200.00 for 3 years, starting in November, 2014.

7/ Projects approved in the least 5 years

(02/2015 – 01/2016) Coordinator: Project PUC/FIP 061/2014, *Recuperação de Informação sobre Entidades na Web*. Grant PUC 2015/9396. Amount of R$ 13,800.00 for 1 year, starting in February, 2015.

(07/2014 – 06/2015) Coordinator: Project Microsoft RFT Brazil, *Entity-related Information Extraction and Retrieval from the Web*. Grant Microsoft. Amount of US$ 40,000.00 for 1 year, starting in July, 2014.

(2007 – 2010) Member: Project CNPq/CT-INFO/InfoWeb, *Métodos e Ferramentas para Tratamento da Informação Disponível na Web*. Grant CNPq 55.0874/2007-0. Amount of R$ 700,000.00 for 3 years, starting in 2007.

8/ Publications

8.1 – Highlight the most important publications related to the project theme

8.2 – Publications in cooperation with the project partners

9/ Theses oriented and post-doctoral fellows supervised

9.1 – Finished/defended in the last 5 years

9.2 – Ongoing
- Leandro Lessa (MSc. student) – Learning to recommend students to graduate courses

<u>**ANNEX: CV**</u>

1/ Personal data
Name: Zenilton Kleber Gonçalves do Patrocínio Júnior
Birth date: 27/11/1967
Professional address (with telephone and e-mail):
Rua Walter Ianni, 255 – CEP: 31980-110 – Belo Horizonte – Minas Gerais - Brasil
Phone: +55 31 3439 5204, E-mail: zenilton@pucminas.br
Current job title and size of the research group:
- Associate Professor, Department of Computer Science, PUC Minas, Brazil
- VIPLAB Research Group (three senior researchers, nine MSc. Students, several undergraduate students).

2/ Highest obtained degree (with indication of place and date)
D.Sc. in Computer Science, UFMG, Brazil, 2005.

3/ Professional activity in the last 5 years
Academic at the Department of Computer Science, PUC Minas, Brazil.

4/ Other duties/ positions
Undergraduate course coordinator – Computer Engineering, PUC Minas – 2006/2009.

5/ Awards, fellowships and external recognition
- (2012) Honorable mention in WTDBD / SBBD with the master thesis "Multimodal Video Indexing and Retrieval".
- (2010) Best paper in WEBMEDIA with the work "Identification and Analysis of Video Subsequence Using Bipartite Graph Matching".

6/ Ongoing funded research projects with dates, titles, sources of funding
- (2015-2017) Main researcher in FAPEMIG Project entitled "Exploring Coherence in Multimodal Vocabulary Construction for Video Shot Retrieval".
- (2014-2017) Main researcher in CAPES/PVE (Cooperation project) entitled "Hierarchical Segmentation of Image and Video based on Graph Analysis".

7/ Projects approved in the least 5 years
- (2014-2016) Senior researcher in CAPES STIC-AMSUD (Cooperarion project) "MaxiMUM – Unsupervised Mining of Multimedia Content".
- (2014-2016) Senior researcher in FAPEMIG-INRIA-CNRS (Cooperarion project) "Unsupervised Motif Discovery in Multimedia Content".
- (2012-2014) Senior researcher in CAPES/CDEFI (Cooperation project) Project BRAFITEC BRasil France Ingénieur TEChnologie "FOR TIC! FORmação em Tecnologia da Informação e da Comunicação"
- (2011-2013) Main researcher in FAPEMIG/INRIA (Cooperation project) Project "Video summarization using multimodal Information (TVSum)" in collaboration to Dr. Patrick Gros and Dr. Guillaume Gravier
- (2011-2013) Senior researcher in FAPEMIG Project entitled "Development of reusable artifacts in cloud computing" in collaboration to CiT Group.

8/ Publications
8.1 – Highlight the most important publications related to the project theme

-   .

8.2 – Publications in cooperation with the project partners
N/A.

9/ Theses oriented and post-doctoral fellows supervised

9.1 – Finished/defended in the last 5 years
-   Aniceto Carlos de Andrade Jr. (M.Sc. student – co-advisor) – Evolutionary algorithm for parameter adjustment in color image segmentation.
-   Fabiano Pereira Bhering (M.Sc. student – co-advisor) – Image retrieval using data clustering techniques.
-   Luciana Belo (M.Sc. student – co-advisor) – Graph-based video summarization using multimodal information.
-   Maíza Cristina de Souza Dias (M.Sc. student) – The use of genetic algorithm for parameter adjustment of a metric data structure used in image indexing and retrieval.
-   Ricardo Carlini Sperandio (M.Sc. student) – Multimodal video indexing and retrieval.
-   Vinícius von Glehn De Filippo (M.Sc. student) – Hierarchical combination of global and local descriptors for image retrieval.
-   Yuri Hoffmann Ribeiro (M.Sc. student) – Evolutionary programming for kernel evolving.

9.2 – Ongoing
-   Milena Menezes Adão (M.Sc. student). Machine learning applied to hierarchies of partitions.
-   Rafael Machado Ribeiro (M.Sc. student). Object detection beased on hierarchical segmentation.
-   Franciele Cristina Parreiras Rodrigues (M.Sc. student). Hiearachical vídeo cosegmentation.

1/ Personal data

Name: Arnaldo de Albuquerque Araújo
Birth date: July 26th 1955
Professional address (with telephone and e-mail):
Federal University of Minas Gerais - UFMG
Computer Science Department - DCC
Av. Antônio Carlos, 6627 - 31270-010 Belo Horizonte – MG, Brazil
Tel.: + 55 31 3409 5873 – Fax: + 55 31 3409 5858
e-mail: arnaldo@dcc.ufmg.br – URL: http://www.npdi.dcc.ufmg
DBLP: http://www.informatik.uni-trier.de/~ley/pers/hd/a/Ara=uacute=jo:Arnaldo_de_Albuquerque
Scholar Google: http://scholar.google.com.br/citations?user=AitxvQIAAAAJ&hl=en
ResearchGate: https://www.researchgate.net/profile/Arnaldo_Araujo/publications

Current job title and size of the research group:
Full Professor and researcher at the Computer Science Department - UFMG
Digital Image Processing Group (NPDI) - One coordinator, five associated researchers, five PhD students, three MSc students, five undergraduate students.

2/ Highest obtained degree (with indication of place and date)

D.Sc. in Electrical Engineering, from the Universidade Federal da Paraiba, Brazil, 1987, with sandwich period at TH Aachen, Germany. Post-docs in Computer Science, in France, at ESIEE, 1994-1995, ENSEA, 2005, and UPMC (LIP6 – Paris 6), 2008-2009.

3/ Professional activity in the last 5 years

Teaching and research activities in the area of Image and Video Processing, CBIR, at the Computer Science Department; Enrollment at the graduate programs in Computer Science.

4/ Other duties/ positions

Head of the Image Processing Group, since 1990. Coordinator of the international cooperation Brazil-France (UFMG, UNICAMP, PUC-Minas, ESIEE, ENSEA, UPMC). Member of the UFMG Committee for Internationalization. Coordinator of the Computer Museum @ DCC/UFMG. Ad-Hoc consultant for Brazilian agencies for R&D CNPq, CAPES and FAPEMIG.

5/ Awards, fellowships and external recognition

2016 - 1st Place, Best PhD Thesis work, 2016 UFMG Computer Science Graduate Program Seminary Week.
2015 - Elevation to IEEE Senior Member Grade.
2014 - Best Paper, Workshop on Vision-based Human Activity Recognition, SIBGRAPI.
2012 - 2nd Place in the ImageCLEF Flickr Photo Annotation Task, Roma, Italy.
2012 - 2nd Place in the Workshop of Thesis and Dissertation, SIBGRAPI, Brazil.
2011 - Best Integration of Multiple Media Prize, ACM Multimedia Grand Challenge, USA.
2010 - Invited Keynote Speaker, 23rd SIBGRAPI, SBC, Brazil.

6/ Ongoing funded research projects with dates, titles, sources of funding

2014-2018 Personal Research Grant CNPq. Coord.: Arnaldo Araújo.

2014 – 2017 Intern. Coop. Project FAPEMIG-INRIA-CNRS 12/2013 (PUC-Minas, UFMG and INRIA Rennes). Coords: Silvio Guimarães and Guillaume Gravier.
2015-2018 Intern. Coop. Project CAPES/PVE 09/2014 , Invited Researcher to Brazil (PUC-Minas, UFMG and ESIEE Paris). Coord: Silvio Guimarães.

2014-2018 Research Project FAPEMIG/PRONEX 19/2013 - Program to Support Groups of Excellence (UFMG). Coord.: Nivio Ziviani.

7/ Projects approved in the least 5 years

2008-2012+2014 Intern. Coop. Project CAPES-COFECUB 005/2007 (UFMG, UNICAMP, PUC-Minas, ESIEE, ENSEA, UPMC). Coords.: Arnaldo Araújo and Matthieu Cord.

2012-2016 Research Project Edital MCT/CNPq 14/2012 (UFMG, PUC-Minas, UNICAMP, and UFOP). Coord.: Arnaldo Araújo.

2014-2016 Personal Research Grant FAPEMIG. Coord.: Arnaldo Araújo.

2014 – 2016 Intern. Coop. Project CAPES/STIC-AMSUD 001/2013 (UFMG, PUC-Minas, UNICAMP, INRIA/IRISA Rennes, UPMC, U. de Chile and ORAND Chile). Coords: Arnaldo Araújo, Guillaume Gravier and Benjamin Bustos.
2014 – 2016 Research Project CNPq/ME 91/2013 (UFMG, PUC-Minas, and UFOP). Coord: William Schwartz.

8/ Publications

8.1 – Highlight the most important publications related to the project theme
- Souza, Araújo, Patrocínio & Guimarães. Graph-based Hierarchical Video Segmentation Based on a Simple Dissimilarity Measure. Pattern Recognition Letters, DOI 10.1016/j.patrec.2014.02.016.
- Caetano, Avila, Schwartz, Guimarães & Araújo. A Mid-level Video Re-presentation Based on Binary Descriptors: A Case Study for Pornography Detection. Neurocomputing, 10.1016/j.neucom.2016.03.099.
- Avila, Thome, Cord, Valle, & Araújo. Pooling in image representation: the visual codeword point of view. CVIU journal, DOI 10.1016/j.cviu.2012.09.007, 2013.

8.2 – Publications in cooperation with the project partners
- Silva, Patrocínio, Gravier, Amsaleg, Araújo & Guimarães. Near-duplicate video detection based on an approximate similarity self-Join strategy. CBMI, DOI 10.1109/CBMI.2016.7500278.
- Pimentel, Bustos, Araújo, & Guimarães. Combining pixel domain and compressed domain index for sketch based image retrieval. MTAP, to appear 2017.

9/ Theses oriented and post-doctoral fellows supervised
9.1 – Finished/defended in the last 5 years
- Kleber Jacques F. de Souza. Hierarchical Video Segmentation Using an Observation Scale. 2016. Thesis, UFMG.
- Carlos Alberto Fraga Pimentel Filho. Sketch-Finder: Efficient and Effective Sketch-based Retrieval for Large Image Collections. 2014. Thesis, UFMG. (Sandwich at CNAM Paris – CAPES/COFECUB).

- Sandra Eliza Fontes de Avila. Extended Bag-of-Words Formalism for Image Classification. 2013. Thesis, UFMG & UPMC. (Double Diploma, CAPES/COFECUB).
- Marcelo de Miranda Coelho. Visual Information Retrieval in Historical Town Image Databases: Contributions for the Recognition and Classification of Images). 2013. Thesis, UFMG.

9.2 – Ongoing
- Henrique Batista da Silva. Identificação e Classificação Automática de Cenas de Vídeo usando Informações Multimodais. Started: 2012. Thesis, UFMG.
- Edward Jorge Cahuina. Mathematical Morphology for Video Annotation. Started: 2014. Thesis, UFMG & U. Paris-Est (Double Diploma).
- Karla Catherine Rodríguez. Mathematical Morphology for Human Action Recognition. Started: 2014. Thesis, UFMG & U. Paris-Est (Double Diploma).
- Virgínia Fernandes Mota. Reconhecimento de ações humanas usando um descritor baseado no tensor de orientação. Started: 2011. Thesis, UFMG.

1/ Personal data

Name: Laurent Amsaleg

Birth date: April, 7, 1967

Professional address (with telephone and e-mail): IRISA, Campus de Beaulieu, 35042 RENNES, FRANCE. +33299847444. Laurent.Amsaleg@irisa.fr

Current job title and size of the research group: Chargé de recherche (research scientist), CNRS, LinkMedia research group (9 permanent researchers)

2/ Highest obtained degree (with indication of place and date)

HdR, Nov 2014, Univ Rennes 1.

3/ Professional activity in the last 5 years: CNRS Research scientist, IRISA, Texmex team. High dimensional indexing, multimedia databases, content based image retrieval, parallelism, distribution

4/ Other duties/ positions

5/ Awards, fellowships and external recognition

- General chair ACM Multimedia 2019. Program committee chair SISAP 2016, MMM 2017. Steering committee member SISAP. TPC member of major conferences: ACM Multimedia, 2008, 2011, 2015; ICMR 2011—2017, MMM 2008—2017, CBMI 2012—2017, ICME 2012—2017 (non exhaustive list). Reviewer for IEEE TCSVT 2016, IEEE TMM 2016, Pattern Recognition Letters 2016, IEEE TIFS 2016, IEEE TIP 2015, Information Systems 2013, MTAP 2013—2016, IEEE TPAMI 2013 (non exhaustive list).
- 2010: grant from the Brittany Research council
- 2004-2009: Associate Teams program, INRIA. With Björn Þór Jónsson
- 2003-2009: Yearly support, International Mobility, Égide

6/ Ongoing funded research projects with dates, titles, sources of funding

- iCODA: Knowledge-mediated Content and Data Interactive Analytics--The case of data journalism. INRIA Project Lab, 2017—2021.

- European project CHISTERA :IDentification for the Internet Of Things. 2017-2020.

7/ Projects approved in the least 5 years

- 2014—2017: CNRS PICS Multimedia Analytics.
- Secure and Large image databases indexing (2012-2014), ANR Contint

8/ Publications

8.1 – Highlight the most important publications related to the project theme

1. Near-Duplicate Video Detection Based on an Approximate Similarity Self-Join Strategy. Henrique Batista Da Silva, Zenilton Patrocino Jr., Guillaume Gravier, Laurent Amsaleg, Arnaldo De A. Araújo, Silvio Jamil Guimarães. *14th International Workshop on Content-based Multimedia Indexing*, Jun 2016, bucarest, Romania.

2. Herwig Lejsek, Friðrik Heidar Ásmundsson, Björn Þór Jónsson, Laurent Amsaleg. NV-tree: An Efficient Disk-Based Index for Approximate Search in Very Large High-Dimensional Collections. IEEE Transactions on Pattern Analysis and Machine Intelligence, 31(5):869-883, May 2009.

3. Loïc Paulevé, Hervé Jégou, Laurent Amsaleg. Locality sensitive hashing: A comparison of hash function types and querying mechanisms. Pattern Recognition Letters, 31(11):1348-1358, 2010

4. Diana Moise, Denis Shestakov, Gylfi Thór Gudmundsson, Laurent Amsaleg. Indexing and Searching 100M Images with Map-Reduce. ACM International Conference on Multimedia Retrieval, Apr 2013, Dallas, United States.

5. Romain Tavenard, Laurent Amsaleg. Improving the Efficiency of Traditional DTW Accelerators. KAIS, 2013.

6. Romain Tavenard, Laurent Amsaleg, Guillaume Gravier. Model-based similarity estimation of multidimensional temporal sequences. Annales des Télécommunications, Springer, 2009, 64 (5), pp. 381-390

8.2 – Publications in cooperation with the project partners

1. Near-Duplicate Video Detection Based on an Approximate Similarity Self-Join Strategy. Henrique Batista Da Silva, Zenilton Patrocino Jr., Guillaume Gravier, Laurent Amsaleg, Arnaldo De A. Araújo, Silvio Jamil Guimarães. *14th International Workshop on Content-based Multimedia Indexing*, Jun 2016, bucarest, Romania

9/ Theses oriented and post-doctoral fellows supervised

- C. Maigrot, 2nd year, Détection de fausses informations dans les réseaux sociaux.
- R. Carlini-Sperandio, 2nd year, Unsupervised motif mining in multimedia time series.
- R. Balu, Privacy-preserving data aggregation and service personalization using highly-scalable data indexing techniques, defended 2016.
- G. Þ. Guðmundsson, Parallelism and Distribution for Very Large Scale Content-Based Image Retrieval, defended 2013.
- T.-T. Do, Security analysis of image copy detection systems based on SIFT descriptors, defended 2012.
- R. Tavenard, Indexation de séquences de descripteurs, defended 2011.
- S.-A. Berrani, Recherche approximative de plus proches voisins avec contrôle probabiliste de la precision ; application à la recherche d'images par le contenu, defended 2004, Best French Thesis in Computer Science.
- Kari Hardarson, Visiting Professor (Reykjavik University), 2*4 months
- Antoine Doucet, Post-Doc, 1 year.
- Panagiotis Adjidoukas, Post-Doc, 1 year.

- Diana Moise, Post-Doc, 1 year.
- Denis Shestakov, Post-Doc, 1 year.
- Hervé Jégou, , Post-Doc, 9 months.

9.1 – Finished/defended in the last 5 years

- R. Balu, Privacy-preserving data aggregation and service personalization using highly-scalable data indexing techniques, defended 2016.
- G. Þ. Guðmundsson, Parallelism and Distribution for Very Large Scale Content-Based Image Retrieval, defended 2013.
- T.-T. Do, Security analysis of image copy detection systems based on SIFT descriptors, defended 2012.

9.2 – Ongoing

- C. Maigrot, 2nd year, Détection de fausses informations dans les réseaux sociaux.
- R. Carlini-Sperandio, 2nd year, Unsupervised motif mining in multimedia time series.

ANNEX: CV

1/ Personal data

Name: Kleber Jacques Ferreira de Souza
Birth date: 15/10/1986
Professional address (with telephone and e-mail):
Rua Walter Ianni, 255 – São Gabriel – Belo Horizonte – Minas Gerais - Brasil
Phone: +55 31 3439 5204, E-mail: klebersouza@pucminas.br
Current job title and size of the research group:
Associate Professor, Department of Software Engineering and Information Systems, PUC Minas, Brazil VIPLAB Research Group (three senior researchers, nine MSc. Students, several undergraduate students).

2/ Highest obtained degree (with indication of place and date)

Doctor in Computer Science, UFMG, Brazil, Mars 2016.

3/ Professional activity in the last 5 years

Academic at the Department of Software Engineering and Information Systems, PUC Minas, Brazil.

4/ Other positions

5/ Awards, fellowships and external recognition

(2016) 1st Place, PhD, presented at PPGCC - Seminars Week 2016: Best Masters and Doctoral Studies, UFMG, with the Work: Hierarchical Video Segmentation using an Observation Scale.

6/ Ongoing funded research projects with dates, titles, sources of funding

7/ Projects approved in the least 5 years

8/ Publications

8.1 – Highlight the most important publications related to the project theme

SOUZA, K. J. F. ; Araújo, Arnaldo de A. ; PATROCINIO JR, Zenilton Kleber G Do ; COUSTY, Jean ; Najman, Laurent ; Yukiko Kenmochi ; GUIMARÃES, SILVIO JAMIL F. . Decreasing the Number of Features for Improving Human Action Classification. In: XXIX Conference on Graphics, Patterns and Images (SIBGRAPI), 2016, São José dos Campos. XXIX Conference on Graphics, Patterns and Images (SIBGRAPI), 2016. v. 1. p. 1-8.

SOUZA, K. J. F. ; ARAÚJO, Arnaldo de Albuquerque ; PATROCINIO JR, Zenilton Kleber G Do ; CORD, M. ; GUIMARÃES, SILVIO JAMIL F. . Streaming graph-based hierarchical video segmentation by a simple label propagation. In: Conference on Graphics, Patterns and Images (SIBGRAPI), 2015, Salvador. Conference on Graphics, Patterns and Images (SIBGRAPI), 2015.

DE SOUZA, KLEBER JACQUES FERREIRA ; DE ALBUQUERQUE ARAÚJO, ARNALDO ; DO PATROCÍNIO, ZENILTON K.G. ; GUIMARÃES, SILVIO JAMIL F. . Graph-based hierarchical video segmentation based on a simple dissimilarity measure. Pattern Recognition Letters, v. 47, p. 85-92, 2014.

SOUZA, KLEBER JACQUES DE ; ARAUJO, ARNALDO DE ALBUQUERQUE ; PATROCINIO, ZENILTON KLEBER G. DO ; COUSTY, Jean ; Najman, Laurent ; KENMOCHI, YUKIKO ; GUIMARAES, SILVIO JAMIL F. . Hierarchical Video Segmentation Using an Observation Scale. In: 2013 XXVI SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), 2013, Arequipa. 2013 XXVI Conference on Graphics, Patterns and Images. p. 320.

8.2 – Publications in cooperation with the project partners

9/ Theses oriented and post-doctoral fellows supervised

9.1 – Finished/defended in the last 5 years

9.2 – Ongoing