Activity Report 2012

Section Application Domains

ALPAGE Project-Team

4. Application Domains

4.1. Panorama

NLP tools and methods have many possible domains of application. Some of then are already mature enough to be commercialized. They can be roughly classified in three groups:

Human-computer interaction : mostly speech processing and text-to-speech, often in a dialogue context; today, commercial offers are limited to restricted domains (train tickets reservation...);

Language writing aid : spelling, grammatical and stylistic correctors for text editors, controlled-language writing aids (e.g., for technical documents), memory-based translation aid, foreign language learning tools, as well as vocal dictation;

Access to information : tools to enable a better access to information present in huge collections of texts (e.g., the Internet): automatic document classification, automatic document structuring, automatic summarizing, information acquisition and extraction, text mining, question-answering systems, as well as surface machine translation. Information access to speech archives through transcriptions is also an emerging field.

Experimental linguistics : tools to explore language in an objective way (this is related, but not limited to corpus linguistics).

Alpage focuses on some applications included in the three last points, such as information extraction and (linguistic and extra-linguistic) knowledge acquisition (4.2), text mining (4.3), spelling correction (4.5) and experimental linguistics (4.6).

4.2. Information extraction and knowledge acquisition

Participants: Éric Villemonte de La Clergerie, Rosa Stern, François-Régis Chaumartin, Benoît Sagot.

The first domain of application for Alpage parsing systems is information extraction, and in particular knowledge acquisition, be it linguistic or not, and text mining.

Knowledge acquisition for a given restricted domain is something that has already been studied by some Alpage members for several years (ACI Biotim, biographic information extraction from the Maitron corpus, Scribo project). Obviously, the progressive extension of Alpage parsing systems or even shallow processing chains to the semantic level increase the quality of the extracted information, as well as the scope of information that can be extracted. Such knowledge acquisition efforts bring solutions to current problems related to information access and take place into the emerging notion of *Semantic Web*. The transition from a web based on data (textual documents,...) to a web based on knowledge requires linguistic processing tools which are able to provide fine grained pieces of information, in particular by relying on high-quality deep parsing. For a given domain of knowledge (say, news or tourism), the extraction of a domain ontology that represents its key concepts and the relations between them is a crucial task, which has a lot in common with the extraction of linguistic information.

In the last years, such efforts have been targeted towards information extraction from news wires in collaboration with the Agence France-Presse (Rosa Stern is a CIFRE PhD student at Alpage and at AFP, and works in relation with the ANR project EDyLex) as well as in the context of the collaboration between Alpage and Proxem, a startup created by François-Régis Chaumartin, PhD student at Alpage (who has defended his PhD in 2012).

These applications in the domain of information extraction raise exciting challenges that require altogether ideas and tools coming from the domains of computational linguistics, machine learning and knowledge representation.

4.3. Processing answers to open-ended questions in surveys: vera

Participants: Benoît Sagot, Valérie Hanoka.

Verbatim Analysis is a startup co-created by Benoît Sagot from Alpage and Dimitri Tcherniak from Towers Watson, a world-wide leader in the domain of employee research (opinion mining among the employees of a company or organization). The aim of its first product, *vera*, is to provide an all-in-one environment for editing (i.e., normalizing the spelling and typography), understanding and classifying answers to open-ended questions, and relating them with closed-ended questions, so as to extract as much valuable information as possible from both types of questions. The editing part relies in part on SxPipe (see section 5.6) and Alexina morphological lexicons. Several other parts of *vera* are co-owned by Verbatim Analysis and by Inria.

4.4. Multilingual terminologies and lexical resources for companies

Participants: Éric Villemonte de La Clergerie, Mickael Morardo, Benoît Sagot.

Lingua et Machina is a small company now headed by François Brown de Colstoun, a former Inria researcher, that provides services for developing specialized multilingual terminologies for its clients. It develops the WEB framework Libellex for validating such terminologies. A formal collaboration with ALPAGE has been set up, with the recruitment of Mikael Morardo as engineer, funded by Inria's DTI. He works on the extension of the web platform *Libellex* for the visualization and validation of new types of lexical resources. In particular, he has integrated a new interface for handling monolingual terminologies, lexical networks, and bilingual wordnet-like structures.

4.5. Automatic and semi-automatic spelling correction in an industrial setting

Participants: Benoît Sagot, Éric Villemonte de La Clergerie, Laurence Danlos.

NLP tools and resources used for spelling correction, such as large n-gram collections, POS taggers and finite-state machinery are now mature and precise. In industrial setting such as post-processing after large-scale OCR, these tools and resources should enable spelling correction tools to work on a much larger scale and with a much better precision than what can be found in different contexts with different constraints (e.g., in text editors). Moreover, such industrial contexts allow for a non-costly manual intervention, in case one is able to identify the most uncertain corrections. An FUI project on this topic has been proposed in collaboration with Diadeis, a company specialized in text digitaliNzation, and two other partners. It has been rerouted to the "Investissements d'avenir" framework, and has been accepted. It started in 2012.

4.6. Experimental linguistics

Participants: Benoît Crabbé, Juliette Thuilier, Luc Boruta.

Alpage is a team that dedicates efforts in producing ressources and algorithms for processing large amounts of textual materials. These ressources can be applied not only for purely NLP purposes but also for linguistic purposes. Indeed, the specific needs of NLP applications led to the development of electronic linguistic resources (in particular lexica, annotated corpora, and treebanks) that are sufficiently large for carrying statistical analysis on linguistic issues. In the last 10 years, pioneering work has started to use these new data sources to the study of English grammar, leading to important new results in such areas as the study of syntactic preferences [60], [128], the existence of graded grammaticality judgments [83].

The reasons for getting interested for statistical modelling of language can be traced back by looking at the recent history of grammatical works in linguistics. In the 1980s and 1990s, theoretical grammarians have been mostly concerned with improving the conceptual underpinnings of their respective subfields, in particular through the construction and refinement of formal models. In syntax, the relative consensus on a generative-transformational approach [71] gave way on the one hand to more abstract characterizations of the language faculty [71], and on the other hand to the construction of detailed, formally explicit, and often implemented, alternative formulation of the generative approach [59], [94]. For French several grammars have been implemented in this trend, among which the tree adjoining grammars of [63], [73] among others. This general movement led to much improved descriptions and understanding of the conceptual underpinnings of both linguistic competence and language use. It was in large part catalyzed by a convergence of interests of logical, linguistic and computational approaches to grammatical phenomena.

However, starting in the 1990s, a growing portion of the community started being frustrated by the paucity and unreliability of the empirical evidence underlying their research. In syntax, data was generally collected impressionistically, either as ad-hoc small samples of language use, or as ill-understood and little-controlled grammaticality judgements (Schütze 1995). This shift towards quantitative methods is also a shift towards new scientific questions and new scientific fields. Using richly annotated data and statistical modelling, we address questions that could not be addressed by previous methodology in linguistics. In this line, at Alpage we have started investigating the question of choice in French syntax with a statistical modelling methodology. Currently two studies are being led on the position of attributive adjectives w.r.t. the noun and the relative position of postverbal complement. This research has contributed to establish new links with the Laboratoire de Linguistique Formelle (LLF, Paris 7) and the Laboratoire de Psychologie et Neuropsychologie Cognitives (LPNCog, Paris 5).

On the other hand we have also started a collaboration with the Laboratoire de Sciences Cognitives de Paris (LSCP/ENS) where we explore the design of algorithms towards the statistical modelling of language acquisition (phonological acquisition). This is currently supported by one PhD project.

METISS Project-Team

4. Application Domains

4.1. Introduction

This section reviews a number of applicative tasks in which the METISS project-team is particularily active :

- spoken content processing
- description of audio streams
- audio scene analysis
- advanced processing for music information retrieval

The main applicative fields targeted by these tasks are :

- multimedia indexing
- audio and audio-visual content repurposing
- description and exploitation of musical databases
- ambient intelligence
- education and leisure

4.2. Spoken content processing

speaker recognition, user authentication, voice signature, speaker adaptation, spoken document, speech modeling, speech recognition, rich transcription, beam-search, broadcast news indexing, audio-based multimodal structuring

A number of audio signals contain speech, which conveys important information concerning the document origin, content and semantics. The field of speaker characterisation and verification covers a variety of tasks that consist in using a speech signal to determine some information concerning the identity of the speaker who uttered it.

In parallel, METISS maintains some know-how and develops new research in the area of acoustic modeling of speech signals and automatic speech transcription, mainly in the framework of the semantic analysis of audio and multimedia documents.

4.2.1. Robustness issues in speaker recognition

Speaker recognition and verification has made significant progress with the systematical use of probabilistic models, in particular Hidden Markov Models (for text-dependent applications) and Gaussian Mixture Models (for text-independent applications). As presented in the fundamentals of this report, the current state-of-the-art approaches rely on bayesian decision theory.

However, robustness issues are still pending : when speaker characteristics are learned on small quantities of data, the trained model has very poor performance, because it lacks generalisation capabilities. This problem can partly be overcome by adaptation techniques (following the MAP viewpoint), using either a speaker-independent model as general knowledge, or some structural information, for instance a dependency model between local distributions.

METISS also adresses a number of topics related to speaker characterisation, in particular speaker selection (i.e. how to select a representative subset of speakers from a larger population), speaker representation (namely how to represent a new speaker in reference to a given speaker population), speaker adaptation for speech recognition, and more recently, speaker's emotion detection.

4.2.2. Speech recognition for multimedia analysis

In multimodal documents, the audio track is generally a major source of information and, when it contains speech, it conveys a high level of semantic content. In this context, speech recognition functionalities are essential for the extraction of information relevant to the taks of content indexing.

As of today, there is no perfect technology able to provide an error-free speech retranscription and operating for any type of speech input. A current challenge is to be able to exploit the imperfect output of an Automatic Speech Recognition (ASR) system, using for instance Natural Language Processing (NLP) techniques, in order to extract structural (topic segmentation) and semantic (topic detection) information from the audio track.

Along the same line, another scientific challenge is to combine the ASR output with other sources of information coming from various modalities, in order to extract robust multi-modal indexes from a multimedia content (video, audio, textual metadata, etc...).

4.3. Description and structuration of audio streams

audio stream, audio detection, audio tracking, audio segmentation, audio descriptors, multimedia indexing, audiovisual integration, multimodality, information fusion, audio-visual descriptors,

Automatic tools to locate events in audio documents, structure them and browse through them as in textual documents are key issues in order to fully exploit most of the available audio documents (radio and television programmes and broadcasts, conference recordings, etc).

In this respect, defining and extracting meaningful characteristics from an audio stream aim at obtaining a structured representation of the document, thus facilitating content-based access or search by similarity.

Activities in METISS focus on sound class and event characterisation and tracking in audio contents for a wide variety of features and documents.

4.3.1. Detecting and tracking sound classes and events

Locating various sounds or broad classes of sounds, such as silence, music or specific events like ball hits or a jingle, in an audio document is a key issue as far as automatic annotation of sound tracks is concerned. Indeed, specific audio events are crucial landmarks in a broadcast. Thus, locating automatically such events enables to answer a query by focusing on the portion of interest in the document or to structure a document for further processing. Typical sound tracks come from radio or TV broadcasts, or even movies.

In the continuity of research carried out at IRISA for many years (especially by Benveniste, Basseville, André-Obrecht, Delyon, Seck, ...) the statistical test approach can be applied to abrupt changes detection and sound class tracking, the latter provided a statistical model for each class to be detected or tracked was previously estimated. For example, detecting speech segments in the signal can be carried out by comparing the segment likelihoods using a speech and a "non-speech" statistical model respectively. The statistical models commonly used typically represent the distribution of the power spectral density, possibly including some temporal constraints if the audio events to look for show a specific time structure, as is the case with jingles or words. As an alternative to statistical tests, hidden Markov models can be used to simultaneously segment and classify an audio stream. In this case, each state (or group of states) of the automaton represent one of the audio event to be detected. As for the statistical test approach, the hidden Markov model approach requires that models, typically Gaussian mixture models, are estimated for each type of event to be tracked.

In the area of automatic detection and tracking of audio events, there are three main bottlenecks. The first one is the detection of simultaneous events, typically speech with music in a speech/music/noise segmentation problem since it is nearly impossible to estimate a model for each event combination. The second one is the not so uncommon problem of detecting very short events for which only a small amount of training data is available. In this case, the traditional 100 Hz frame analysis of the waveform and Gaussian mixture modeling suffer serious limitations. Finally, typical approaches require a preliminary step of manual annotation of a training corpus in order to estimate some model parameters. There is therefore a need for efficient machine learning and statistical parameter estimation techniques to avoid this tedious and costly annotation step.

4.3.2. Describing multi-modal information

Applied to the sound track of a video, detecting and tracking audio events can provide useful information about the video structure. Such information is by definition only partial and can seldom be exploited by itself for multimedia document structuring or abstracting. To achieve these goals, partial information from the various media must be combined. By nature, pieces of information extracted from different media or modalities are heterogeneous (text, topic, symbolic audio events, shot change, dominant color, etc.) thus making their integration difficult. Only recently approaches to combine audio and visual information in a generic framework for video structuring have appeared, most of them using very basic audio information.

Combining multimedia information can be performed at various level of abstraction. Currently, most approaches in video structuring rely on the combination of structuring events detected independently in each media. A popular way to combine information is the hierarchical approach which consists in using the results of the event detection of one media to provide cues for event detection in the other media. Application specific heuristics for decision fusions are also widely employed. The Bayes detection theory provides a powerful theoretical framework for a more integrated processing of heterogeneous information, in particular because this framework is already extensively exploited to detect structuring events in each media. Hidden Markov models with multiple observation streams have been used in various studies on video analysis over the last three years.

The main research topics in this field are the definition of structuring events that should be detected on the one hand and the definition of statistical models to combine or to jointly model low-level heterogeneous information on the other hand. In particular, defining statistical models on low-level features is a promising idea as it avoids defining and detecting structuring elements independently for each media and enables an early integration of all the possible sources of information in the structuring process.

4.3.3. Recurrent audio pattern detection

A new emerging topic is that of motif discovery in large volumes of audio data, i.e. discovering similar units in an audio stream in an unsupervised fashion. These motifs can constitue some form of audio "miniatures" which characterize some potentially salient part of the audio content : key-word, jingle, etc...

This problem naturally requires the definition of a robuste metric between audio segments, but a key issue relies in an efficient search strategy able to handle the combinatorial complexity stemming from the competition between all possible motif hypotheses. An additional issue is that of being able to model adequately the collection of instances corresponding to a same motif (in this respect, the HMM framework certainly offers a reasonable paradigm).

4.4. Advanced processing for music information retrieval

audio object, music description, music language modeling, multi-level representations

4.4.1. Music content modeling

Music pieces constitue a large part of the vast family of audio data for which the design of description and search techniques remain a challenge. But while there exist some well-established formats for synthetic music (such as MIDI), there is still no efficient approach that provide a compact, searchable representation of music recordings.

In this context, the METISS research group dedicates some investigative efforts in high level modeling of music content along several tracks. The first one is the acoustic modeling of music recordings by deformable probabilistic sound objects so as to represent variants of a same note as several realisation of a common underlying process. The second track is music language modeling, i.e. the symbolic modeling of combinations and sequences of notes by statistical models, such as n-grams.

4.4.2. Multi-level representations for music information retrieval

New search and retrieval technologies focused on music recordings are of great interest to amateur and professional applications in different kinds of audio data repositories, like on-line music stores or personal music collections.

The METISS research group is devoting increasing effort on the fine modeling of multi-instrument/multi-track music recordings. In this context we are developing new methods of automatic metadata generation from music recordings, based on Bayesian modeling of the signal for multilevel representations of its content. We also investigate uncertainty representation and multiple alternative hypotheses inference.

4.5. Audio scene analysis

source separation, multi-channel audio, source characterization, source localization, compressive sensing

Audio signals are commonly the result of the superimposition of various sources mixed together : speech and surrounding noise, multiple speakers, instruments playing simultaneously, etc...

Source separation aims at recovering (approximations of) the various sources participating to the audio mixture, using spatial and spectral criteria, which can be based either on a priori knowledge or on property learned from the mixture itself.

4.5.1. Audio source separation

The general problem of "source separation" consists in recovering a set of unknown sources from the observation of one or several of their mixtures, which may correspond to as many microphones. In the special case of *speaker separation*, the problem is to recover two speech signals contributed by two separate speakers that are recorded on the same media. The former issue can be extended to *channel separation*, which deals with the problem of isolating various simultaneous components in an audio recording (speech, music, singing voice, individual instruments, etc.). In the case of *noise removal*, one tries to isolate the "meaningful" signal, holding relevant information, from parasite noise.

It can even be appropriate to view audio compression as a special case of source separation, one source being the compressed signal, the other being the residue of the compression process. The former examples illustrate how the general source separation problem spans many different problems and implies many foreseeable applications.

While in some cases –such as multichannel audio recording and processing– the source separation problem arises with a number of mixtures which is at least the number of unknown sources, the research on audio source separation within the METISS project-team rather focusses on the so-called under-determined case. More precisely, we consider the cases of one sensor (mono recording) for two or more sources, or two sensors (stereo recording) for $n > 2$ sources.

We address the problem of source separation by combining spatial information and spectral properties of the sources. However, as we want to resort to as little prior information as possible we have designed self-learning schemes which adapt their behaviour to the properties of the mixture itself [1].

4.5.2. Compressive sensing of acoustic fields

Complex audio scene may also be dealt with at the acquisition stage, by using "intelligent" sampling schemes. This is the concept behind a new field of scientific investigation : compressive sensing of acoustic fields.

The challenge of this research is to design, implement and evaluate sensing architectures and signal processing algorithms which would enable to acquire a reasonably accurate map of an acoustic field, so as to be able to locate, characterize and manipulate the various sources in the audio scene.

PAROLE Project-Team

4. Application Domains

4.1. Application Domains

Our research is applied in a variety of fields from ASR to paramedical domains. Speech analysis methods will contribute to the development of new technologies for language learning (for hearing-impaired persons and for the teaching of foreign languages) as well as for hearing aids. In the past, we developed a set of teaching tools based on speech analysis and recognition algorithms of the group (cf. the ISAEUS [43] project of the EU that ended in 2000). We are continuing this effort towards the diffusion of a course on Internet.

Speech is likely to play an increasing role in man-machine communication. Actually, speech is a natural mean of communication, particularly for non-specialist persons. In a multimodal environment, the association of speech and designation gestures on touch screens can, for instance, simplify the interpretation of spatial reference expressions. Besides, the use of speech is mandatory in many situations where a keyboard is not available: mobile and on-board applications (for instance in the framework of the HIWIRE European project for the use of speech recognition in a cockpit plane), interactive vocal servers, telephone and domestic applications, etc. Most of these applications will necessitate to integrate the type of speech understanding process that our group is presently studying. Furthermore, speech to speech translation concerns all multilingual applications (vocal services, audio indexing of international documents). The automatic indexing of audio and video documents is a very active field that will have an increasing importance in our group in the forthcoming years, with applications such as economic intelligence, keyword spotting and automatic categorization of mails.

SEMAGRAMME Team

4. Application Domains

4.1. Introduction

Our applicative domains concern natural language processing applications that rely on a deep semantic analysis. For instance, one may cite the following ones:

- textual entailment and inference;
- dialogue systems;
- semantic-oriented query systems;
- content analysis of unstructured documents;
- text transformation and automatic summarization;
- (semi) automatic knowledge acquisition.

However, if the need for semantics seems to be ubiquitous. There is a challenge in finding applications for which a deep semantic analysis results in a real improvement over non semantic-based techniques.

4.2. Text Transformation

Text transformation is an application domain featuring two important sub-fields of computational linguistics:

- parsing: from surface form to abstract representation;
- generation: from abstract representation to surface form.

Text simplification or automatic summarization belong to that domain.

We aim at using the framework of Abstract Categorial Grammars we develop to this end. It is indeed a *reversible* framework that allows both parsing and generation. Its underlying mathematical structure of λ-calculus makes it fit with our type-theoretic approach to discourse dynamics modeling. The ANR project POLYMNIE(see section 7.2.1.1) is especially dedicated to this aim.

ALICE Project-Team (section vide)

<p style="text-align:center;color:red;font-weight:bold;">AVIZ Project-Team</p>

4. Application Domains

4.1. Application Domains

AVIZ develops active collaboration with users from various application domains, making sure it can support their specific needs. By studying similar problems in different domains, we can begin to generalize our results and have confidence that our solutions will work for a variety of applications.

Our current application domains include:

- *Genealogy*, in cooperation with North Carolina State University;
- *Biological research*, in cooperation with Institut Pasteur;
- *Digital Libraries*, in cooperation with the French National Archives and the Wikipedia community;
- *Open Data*, in cooperation with Google Open Data and Data Publica;
- *Agrifood Process Modeling*, in cooperation with the DREAM project (see section 8.2.1.1);

4. Application Domains

4.1. Application Domains

This research can be applied to any situation where users need to create new, imaginary, 3D content. Our work should be instrumental, in the long term, for the visual arts, from the creation of 3D films and games to the development of new digital planning tools for theatre or cinema directors. Our models can also be used in interactive prototyping environments for engineering. They can help promoting interactive digital design to scientists, as a tool to quickly express, test and refine models, as well as an efficient way for conveying them to other people. Lastly, we expect our new methodology to put digital modelling within the reach of the general public, enabling educators, media and other practitioners to author their own 3D content.

In practice, fully developing a few specialized interactive systems will be instrumental for testing our models. The multi-disciplinary expertise and professional background of our team members will ease the set up of projects in the domains listed below. The diversity of users these domains bring, from digital experts to other professionals and novices, will be excellent for validating our general methodology. Our ongoing projects in these various application domains are listed in Section 6.

- Visual arts
    - Modeling and animation for 3D films and games (François Faure, Marie-Paule Cani,)
    - Virtual cinematography and tools for theatre directors (Rémi Ronfard)
- Engineering
    - Industrial design (Stéfanie Hahmann, Jean-Claude Léon)
    - Mechanical & civil engineering (Jean-Claude Léon, François Faure)
- Natural Sciences
    - Virtual functional anatomy (Olivier Palombi, François Faure)
    - Virtual plants (Marie-Paule Cani, François Faure)
- Education and Creative tools
    - Sketch-based teaching (Olivier Palombi, Marie-Paule Cani)
    - Creative environments for novice users (Marie-Paule Cani, Jean-Claude Léon)

<p style="text-align: center; color: red;">**IN-SITU Project-Team**</p>

4. Application Domains

4.1. Research Domains

INSITU works on general problems of interaction in multi-surface environments as well as on challenges associated with specific research groups. The former requires a combination of controlled experiments and field studies; the latter involves participatory design with users. We are currently working with highly creative people, particularly designers and music composers, to explore interaction techniques and technologies that support the earliest phases of the design process. We are also working with research scientists, particularly neuroscientists and astrophysicists, in our explorations of interaction in multisurface environments, and with doctors and nurses to support crisis management situations.

MANAO Team (section vide)

MAVERICK Team

4. Application Domains

4.1. Introduction

Maverick is part of the research theme "Interaction and Visualization" at Inria. This research theme has historically been very successful inside Inria. It nicely connects industrial applications with fundamental research using advanced mathematics, algorithmic and computer science, and it connects computer science with other sciences such as physics, biology, medicine, environment, psychophysiology.

We envision Maverick at this crossroad. We have several industrial partnerships, with companies making video games (Eden Games), special effects for motion pictures (WetaFX), planetarium (RSA Cosmos), graphical edition software (Adobe), tomography (Digisens) or visualizing simulated data (EDF). The constraints and needs of our partners motivate new problems for us to solve. At the same time, we are looking into fundamental research problems, such as analysis of light transport, human perception, filtering and sampling.

The fundamental research problems we target are not necessarily "long term research": the computer graphics industry is very dynamic and can adopt (and adapt) a research paper in a matter of months if it sees benefits in it. The research problems we describe as "fundamental" correspond to high-risk, high-benefit research problems. Solving these problems would result in a significant breakthrough for the whole domain of Computer Graphics, both in research and in industry.

4.2. Illustration

Although it has long been recognized that the visual channel is one of the most effective means for communicating information, the use of computer processing to generate effective visual content has been mostly limited to very specific image types: realistic rendering, computer-aided cell animation, etc.

The ever-increasing complexity of available 3d models is creating a demand for improved image creation techniques for general illustration purposes. Recent examples in the literature include computer systems to generate road maps, or assembly instructions, where a simplified visual representation is a necessity.

Our work in expressive rendering and in relevance-guided rendering aims at providing effective tools for all illustration needs that work from complex 3d models. We also plan to apply our knowledge of lighting simulation, together with expressive rendering techniques, to the difficult problem of sketching illustrations for architectural applications.

4.3. Video-games and visualization

Video games represent a particularly challenging domain of application since they require both real-time interaction and high levels of visual quality. Moreover, video games are developed on a variety of platforms with completely different capacities. Automatic generation of appropriate data structures and runtime selection of optimal rendering algorithms can save companies a huge amount of development.

More generally, interactive visualization of complex data (e.g. in scientific engineering) can be achieved only by combining various rendering accelerations (e.g. visibility culling, levels of details, etc.), an optimization task that is hard to perform "by hand" and highly data dependent. One of Maverick' goals is to understand this dependence and automate the optimization.

4.4. Virtual heritage

Virtual heritage is a recent area which has seen spectacular growth over the past few years. Archeology and heritage exhibits are natural application areas for virtual environments and computer graphics, since they provide the ability to navigate 3D models of environments that no longer exist and can not be recorded on a videotape. Moreover, digital models and 3D renderings give the ability to enrich the navigation with annotations.

Our work on style has proved very interesting to architects who have a long habit of using hand-drawn schemas and wooden models to work and communicate. Wooden models can advantageously be replaced by 3D models inside a computer. Drawing, on the other hand, offers a higher level of interpretation and a richness of expression that are really needed by architects, for example to emphasize that such model is an hypothesis.

By investigating style analysis and expressive rendering, we could "sample" drawing styles used by architects and "apply" them to the rendering of 3D models. The computational power made available by computer assisted drawing can also lead to the development of new styles with a desired expressiveness, which would be harder to produce by hand. In particular, this approach offers the ability to navigate a 3D model while offering an expressive rendering style, raising fundamental questions on how to "animate" a style.

MIMETIC Team

4. Application Domains

4.1. Motion Sensing

Recording human activity is a key point of many applications and fundamental works. Numerous sensors and systems have been proposed to measure positions, angles or accelerations of the user's body parts. Whatever the system is, one of the main is to be able to automatically recognize and analyze the user's performance according to poor and noisy signals. Human activity and motion are subject to variability: intra-variability due to space and time variations of a given motion, but also inter-variability due to different styles and anthropometric dimensions. MimeTIC has addressed the above problems in two main directions.

Firstly, we have studied how to recognize and quantify motions performed by a user when using accurate systems such as Vicon (product of Oxford Metrics) or Optitrack (product of Natural Point) motion capture systems. These systems provide large vectors of accurate information. Due to the size of the state vector (all the degrees of freedom) the challenge is to find the compact information (named features) that enables the automatic system to recognize the performance of the user. Whatever the method is used, finding these relevant features that are not sensitive to intra-individual and inter-individual variability is a challenge. Some researchers have proposed to manually edit these features (such as a Boolean value stating if the arm is moving forward or backward) so that the expertise of the designer is directly linked with the success ratio. Many proposals for generic features have been proposed, such as using Laban notation which was introduced to encode dancing motions. Other approaches tend to use machine learning to automatically extract these features. However most of the proposed approaches were used to seek a database for motions which properties correspond to the features of the user's performance (named motion retrieval approaches). This does not ensure the retrieval of the exact performance of the user but a set of motions with similar properties.

Secondly, we wish to find alternatives to the above approach which is based on analyzing accurate and complete knowledge on joint angles and positions. Hence new sensors, such as depth-cameras (Kinect, product of Microsoft) provide us with very noisy joint information but also with the surface of the user. Classical approaches would try to fit a skeleton into the surface in order to compute joint angles which, again, lead to large state vectors. An alternative would be to extract relevant information directly from the raw data, such as the surface provided by depth cameras. The key problem is that the nature of these data may be very different from classical representation of human performance. In MimeTIC, we try to address this problem in specific application domains that require picking specific information, such as gait asymmetry or regularity for clinical analysis of human walking.

4.2. VR and Sports

Sport is characterized by complex displacements and motions. These motions are dependent on visual information that the athlete can pick up in his environment, including the opponent's actions. The perception is thus fundamental to the performance. Indeed, a sportive action, as unique, complex and often limited in time, requires a selective gathering of information. This perception is often seen as a prerogative for action, it then takes the role of a passive collector of information. However, as mentioned by Gibson in 1979, the perception-action relationship should not be considered sequential but rather as a coupling: we perceive to act but we must act to perceive. There would thus be laws of coupling between the informational variables available in the environment and the motor responses of a subject. In other words, athletes have the ability to directly perceive the opportunities of action directly from the environment. Whichever school of thought considered, VR offers new perspectives to address these concepts by complementary using real time motion capture of the immersed athlete.

In addition to better understanding sports and interaction between athletes, VR can also be used as a training environment as it can provide complementary tools to coaches. It is indeed possible to add visual or auditory information to better train an athlete. The knowledge found in perceptual experiments can be for example used to highlight the body parts that are important to look at to correctly anticipate the opponent's action.

4.3. Biomechanics and Motion Analysis

Biomechanics is obviously a very large domain. This large set can be divided regarding to the scale at which the analysis is performed going from microscopic evaluation of biological tissues' mechanical properties to macroscopic analysis and modeling of whole body motion. Our topics in the domain of biomechanics mainly lie within this last scope.

The first goal of such kind of research projects is a better understanding of human motion. The MimeTic team addresses three different situations: everyday motions of a lambda subject, locomotion of pathological subjects and sports gesture.

In the first set, Mimetic is interested in studying how subjects maintain their balance in highly dynamic conditions. Until now, balance havec nearly always been considered in static or quasi-static conditions. The knowledge of much more dynamic cases still has to be improved. Our approach has demonstrated that first of all, the question of the parameter that will allow to do this is still open. We have also taken interest into collision avoidance between two pedestrian. This topic includes the research of the parameters that are interactively controlled and the study of each one's role within this interaction.

When patients, in particular those suffering from central nervous system affection, cannot have an efficient walking it becomes very useful for practicians to benefit from an objective evaluation of their capacities. To propose such help to patients following, we have developed two complementary indices, one based on kinematics and the other one on muscles activations. One major point of our research is that such indices are usually only developed for children whereas adults with these affections are much more numerous.

Finally, in sports, where gesture can be considered, in some way, as abnormal, the goal is more precisely to understand the determinants of performance. This could then be used to improve training programs or devices. Two different sports have been studied: the tennis serve, where the goal was to understand the contribution of each segments of the body in ball's speed and the influence of the mechanical characteristics of the fin in fin swimming.

After having improved the knowledge of these different gestures a second goal is then to propose modeling solutions that can be used in VR environments for other research topics within MimeTic. This has been the case, for exemple, for the colision avoidance.

4.4. Crowds

Crowd simulation is a very active and concurrent domain. Various disciplines are interested in crowds modeling and simulation: Mathematics, Cognitive Sciences, Physics, Computer Graphics, etc. The reason for this large interest is that crowd simulation raise fascinating challenges.

At first, crowd can be first seen as a complex system: numerous local interactions occur between its elements and results into macroscopic emergent phenomena. Interactions are of various nature and are undergoing various factors as well. Physical factors are crucial as a crowd gathers by definition numerous moving people with a certain level of density. But sociological, cultural and psychological factors are important as well, since crowd behavior is deeply changed from country to country, or depending on the considered situations.

On the computational point of view, crowd push traditional simulation algorithms to their limit. An element of a crowd is subject to interact with any other element belonging the same crowd, a naive simulation algorithm has a quadratic complexity. Specific strategies are set to face such a difficulty: level-of-detail techniques enable scaling large crowd simulation and reach real-time solutions.

MimeTIC is an international key contributor in the domain of crowd simulation. Our approach is specific and based on three axis. First, our modeling approach is founded on human movement science: we conducted challenging experiment on the motion of groups. Second: we developed high-performance solutions for crowd simulation. Third, we develop solutions for realistic navigation in virtual world to enable interaction with crowds in Virtual Reality.

4.5. Interactive Digital Storytelling

Interactive digital storytelling, including novel forms of edutainment and serious games, provides access to social and human themes through stories which can take various forms and contains opportunities for massively enhancing the possibilities of interactive entertainment, computer games and digital applications. It provides chances for redefining the experience of narrative through interactive simulations of computer-generated story worlds and opens many challenging questions at the overlap between computational narratives, autonomous behaviours, interactive control, content generation and authoring tools.

Of particular interest for the Mimetic research team, virtual storytelling triggers challenging opportunities in providing effective models for enforcing autonomous behaviours for characters in complex 3D environments. Offering both low-level capacities to characters such as perceiving the environments, interacting with the environment and reacting to changes in the topology, on which to build higher-levels such as modelling abstract representations for efficient reasonning, planning paths and activities, modelling cognitive states and behaviours requires the provision of expressive, multi-level and efficient computational models. Furthermore virtual storytelling requires the seamless control of the balance between the autonomy of characters and the unfolding of the story story through the narrative discourse. Virtual storytelling also raises challenging questions on the conveyance of a narrative through interactive or automated control of the cinematography (how to stage the characters, the lights and the cameras). For example, estimating visibility of key subjects, or performing motion planning for cameras and lights are central issues for which have not received satisfactory answers in the litterature.

4.6. Autonomous characters

Autonomous characters are becoming more and more popular has they are used in an increasing number of application domains. In the field of special effects, virtual characters are used to replace secondary actors and generate highly populated scenes that would be hard and costly to produce with real actors. In video games and virtual storytelling, autonomous characters play the role of actors that are driven by a scenario. Their autonomy allows them to react to unpredictable user interactions and adapt their behavior accordingly. In the field of simulation, autonomous characters are used to simulate the behavior of humans in different kind of situations. They enable to study new situations and their possible outcomes.

One of the main challenges in the field of autonomous characters is to provide a unified architecture for the modeling of their behavior. This architecture includes perception, action and decisional parts. This decisional part needs to mix different kinds of models, acting at different time scale and working with different nature of data, ranging from numerical (motion control, reactive behaviors) to symbolic (goal oriented behaviors, reasoning about actions and changes).

In the MimeTIC team, we focus on autonomous virtual humans. Our problem is not to reproduce the human intelligence but to propose an architecture making it possible to model credible behaviors of anthropomorphic virtual actors evolving/moving in real time in virtual worlds. The latter can represent particular situations studied by psychologists of the behavior or to correspond to an imaginary universe described by a scenario writer. The proposed architecture should mimic all the human intellectual and physical functions.

MINT Project-Team

4. Application Domains

4.1. Next-generation desktop systems

The term *desktop system* refers here to the combination of a window system handling low-level graphics and input with a window manager and a set of applications that share a distinctive look and feel. It applies not only to desktop PCs but also to any other device or combination of devices supporting graphical interaction with multiple applications. Interaction with these systems currently rely on a small number of interaction primitives such as text input, pointing and activation as well as a few other basic gestures. This limited set of primitives is one reason the systems are simple to use. There is, however, a cost. Most simple combinations being already used, few remain to trigger and control innovative techniques that could facilitate task switching or data management, for example. Desktop systems are in dire need of additional interaction primitives, including gestural ones.

4.2. Ambient Intelligence

Ambient intelligence (AmI) refers to the concept of being surrounded by intelligent systems embedded in everyday objects [33]. Envisioned AmI environments are aware of human presence, adapt to users' needs and are capable of responding to indications of desire and possibly engaging in intelligent dialogue. Ambient Intelligence should be unobtrusive: interaction should be relaxing and enjoyable and should not involve a steep learning curve. Gestural interaction is definitely relevant in this context.

4.3. Serious Games

Serious game refers to techniques extensively used in computer games, that are being used for other purposes than gaming. Fields such as learning, use of Virtual Reality for rehabilitation, 3D interactive worlds for retail, art-therapy, are specific context with which the MINT group has scientific connection, and industrial contacts. This field of application is a good opportunity for us to test and transfer our scientific knowledge and results.

4.4. Interactive Art

The heart of Mint project is about interaction gesture, and aims at making relation between application and user more intimate through the production of tools and methods for application to use more information from user gesture. There seems to be, at first sight, very strong difference of fields, tools, vocabulary, between Science and Art. Up to basic intellectual schemes are classically thought to be different. Yet, a closer look needs to be taken on things. Through time, Art is more and more involved in relation between people and content. For example, *relational art*[1] is centered on inter-human relations and social context. Because of this similar analysis about relation between person and content, research on interactive systems probably has a lot to develop relations with Art, this is also true for research on gestural interaction.

---

[1] http://en.wikipedia.org/wiki/Relational_Art

POTIOC Team

4. Application Domains

4.1. Application domains

Since our project aims at providing 3D digital worlds to all, including the general public, in order to stimulate understanding, learning, communication and creation, our scope of applications will naturally be the following one:

- **Culture and education:** We are convinced that a 3D digital world is a powerful media that may contribute to enhance understanding processes. For example, a museum would benefit from new 3D user interfaces allowing visitors to better understand complex content. Similarly, at school, this media has an extraordinary potential for enhancing learning. For example, a child being able to navigate in archaeological sites, or being able to manipulate by himself 3D molecules as described previously, will probably understand and learn things while having pleasure in interacting with the content.

- **Art:** We believe that 3D digital worlds may stimulate creativity, too. Our first investigations with music and drawings have shown that this media opens new possibilities for creation. The challenge here will be to design good interfaces that will allow artists to explore new dimensions. The user may be an experienced artist, or a three years old child who would express his creativity thought tools that go beyond papers and pens.

- **Healthcare:** People suffering from motor or cognitive impairments are one of the target populations of the Potioc project. Indeed, we believe that new interfaces that exploit 3D digital worlds may help people to overcome their disabilities. For example, someone with very reduced motor capabilities could benefit from BCI to explore a virtual museum, or a children having difficulties for concentrating may benefit from new 3D interactive systems.

- **Entertainment:** The objective of Potioc is to open 3D digital worlds to everyone by designing innovative interfaces driven by enjoyment. Consequently, the entertainment industry will be an obvious application domain where Potioc can contribute. This can be in the scope of videogames, entertainment parks, Web and TV of the future, applications for mobile devices, and so on.

Naturally, we will not necessarily address all these applications, and certainly not all at the same time. These areas just define the applicative scope of our work. As an example, much of our current work is targeted at artistic and entertainment applications, with VR-based musical performances, augmented paper-based drawing or BCI-based video games. We are also currently starting to conduct research on digital cities, in order to provide ordinary citizens suitable tools and UI to explore 3D content related to their city, such as 3D maps, information about population density or sound nuisance, among other. It should also be noted that our work might find applicative connexions outside these main application domains and benefit to a large range of academic and industrial areas, with which we could build relationships. For example, in the scope of medicine, new and easy to adopt user interfaces designed in Potioc could prove valuable for medical professionals as well, to better access and interact with biological 3D content (e.g., X-rays or MRI scans).

REVES Project-Team (section vide)

<p align="center">**VR4I Team**</p>

4. Application Domains

4.1. Panorama

The research topics of the VR4i team are related to applications of the industrial, training and education domains.

The applications to the industrial domain are very promising. For instance, the PSA Automotive Design Network, which is a new design center, groups all the tools used for automotive design, from classical CAD systems to Virtual Reality applications. The coupling of virtual reality and simulation algorithms is a key point and is the core of VR4i simulation activities. Major issues in which industrials are strongly involved are focussing on collaborative tasks between multiple users in digital mockups and for scientific visualization (ANR Part@ge and ANR Collaviz 7.1.2), tackling the challenging problem of training in Virtual Reality by providing interactive scenario languages with realists actions and reactions within the environment (GVT Project, ANR Corvette 7.1.3 and FUI SIFORAS 7.1.1). In this context, we are tackling the problem of using Virtual Reality environments for improving the ergonomics of workstations. Collaborative work is now a hot issue for facing the question of sharing expertise of distant experts for project review, for collaborative design or for analysis of data resulting from scientific computations (FP7-Infra VISIONAIR project 7.2.1) where we propose new software architectures ensuring the data distribution and the synchronization of the users (Figure 1).
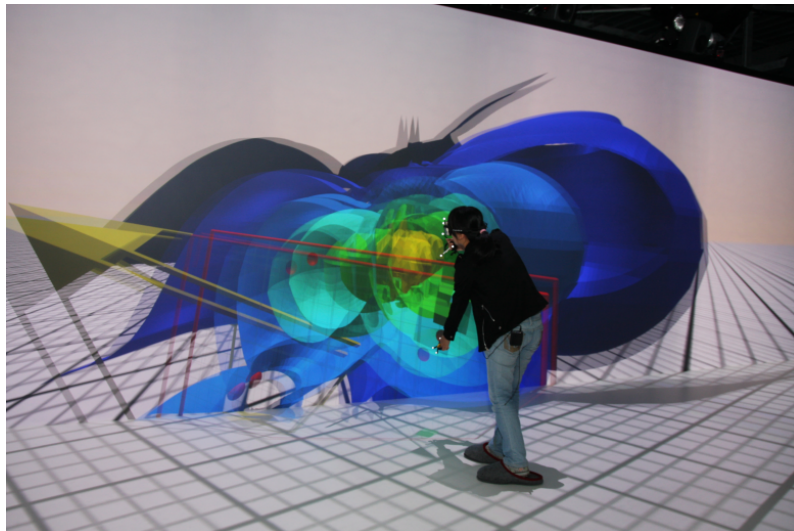

Figure 1. Collaboration between VR4i team in Immersia Room 6.4 and UCL on shared analysis of earthquake simulation within VISIONAIR project 7.2.1

AXIS Project-Team

3. Application Domains

3.1. Panorama: Living Labs, Smart Cities

AxIS addresses applicative field which has the following features:

a) requiring usage/data storage, preprocessing and analysis tools

- for designing, evaluating and improving huge evolving hypermedia information systems (mainly Web-based ISs), for which end-users are of primary concern,
- for a better understanding of service/product used with data mining techniques and knowledge management
- for social network analysis (for example in Web 2.0 applications, Business Intelligence, Sustainable Development, etc.)

b) requiring user-driven innovation methods.

Even if our know-how, methods and algorithms have a cross domain applicability, our team chooses to focus on **Living Lab projects** (and mainly related to **Sustainable Development for Smart Cities**) (cf. section 5.5.5 which imply user involvement for the generation of future services/products. Indeed, following the Rio Conference (1992) and the Agenda for the 21st Century, local territories are now directly concerned with the set up of actions for a sustainable development. In this frame, ICT tools are supposed to be very efficient to re-engage people in the democratic process and to make decision-making more transparent, inclusive and accessible. So, sustainable development is closely associated with citizen participation. The emerging research field of e-democracy (so called Digital Democracy or eParticipation), concerned with the use of communications technologies such as the Internet to enhance the democratic processes is now a very active field. Though still in its infancy, a lot of literature is already available (see for instance: http://itc.napier.ac.uk/ITC/publications.asp or http://www.demo-net.org/ for a global view of work in Europe) and numerous different topics are addressed in the field.

Our experience particularly stressed on the following applicative domains:

- Transportation systems & Mobility (cf. section 3.2),
- Tourism (cf. section 3.3),
- User Involvement in Energy, Environment, Well Being & Health and e-governement (cf. section 3.4).

3.2. Transportation Systems & Mobility

Major recent evolutions in Intelligent Transportation Systems (ITS) are linked to rapid changes in communication technologies, such as ubiquitous computing, semantic web, contextual design. A strong emphasis is now put on mobility improvements. In addition to development of sustainable transportation systems (better ecological vehicles' performance, reduction of impacts on town planning ...) these improvements concern also mobility management, that is specific measures to encourage people to adopt new mobility behaviour such as public transportation services rather than their personal car. These prompting measures concern for instance the quality of traveller's information systems for trip planning, the ability to provide real time recommendations for changing transportation means according to traffic information, and the quality of embedded services in vehicles to provide enhanced navigation aids with contextualised and personalised information.

Since 2004, AxIS has been concerned with mobility projects :

- PREDIT (2004-2007): The MobiVIP project has been an opportunity to collaborate with local Institutions (Communauté d'Agglomération de Sophia Antipolis - CASA) and SMEs (VU Log) and to apply AxIS' know-how in data and web mining to the field of transportation systems.

- Traveller's information systems and recommender systems have been studied with the evaluation of two CASA web sites : the "Envibus" web site which provides information about a bus network and the "Otto&co" web site support car-sharing.

- Advanced transportation systems has been studied in PREDIT TIC TAC (2010-2012): this project (cf. section 6.1.1) aimed at optimizing travel time by providing in an aera with weak transportation services, a just in time on demand shuttle based on real time information. It was for AxIS the opportunity to experiment user implication in the design of a new travel information system called MOBILTIC

- User Experience: in the ELLIOT project (cf. section 6.3.1.1), the mobility scenario is addressed in relation to information on air quality and noise and the use of internet of things.

3.3. Tourism

As tourism is a highly competitive domain, local tourism authorities have developed Web sites in order to offer of services to tourists. Unfortunately, the way information is organised does not necessarily meet Internet users expectations and numerous improvements are necessary to enhance their understanding of visited sites. Thus, even if only for economical reasons, the quality and the diversity of tourism packages have to be improved, for example by highlighting cultural heritage.

Again to illustrate our role in such a domain, let us cite some past projects where AxIS is involved related mainly to **Semantic Web Mining**[3]. In our case, a) we exploit ontologies and semantic data for improving usage analysis, personnalised services, the quality of results of search engines and for checking the content of an IS and also b) we exploit usage data for updating ontologies.) and Information Retrieval.

- Research has been carried out using log files from the city of Metz. This city was chosen because its Web site is in constant development and has been awarded several times, notably in 2003, 2004 and 2005 in the context of the Internet City label. The objective was to extract information about tourists behaviours from this site log files and to identify possible benefits in designing or updating a tourism ontology.

- Providing Tourism Information linked to Transportation information: AxIS has already studied recommender systems in order to provide users with personalised transportation information while looking for tourism information such as cultural information, leisure etc (cf. our recommender Be-TRIP (2006) based on CBR*tools).

- In the context of HOTEL-REF-PACA project (cf. section 6.1.2 , we aimed to better refer the web sites of hotels/campings from the region of TOURVAL in PACA (mainly Vésubie territory), with an approach based on a better understanding of usage from the internauts. To address this, we proposed and adopted a multidisicplinary approach combining various AxIs know-how: knowledge engineering (ontology in tourism), data mining (analysis of Google logs, hotel web site logs and user queries, visual behaviours from eye tracker), Ergonomics (clustering of hotel web sites based on their ergonomical quality).

- Several contacts (PACA, France Living Labs, Island of the Reunion) have been done related to projects in tourism and eco-tourism.

---

[3]By Semantic Web Mining, we mean the mutual benefits between two communities Semantic Web and Web Mining

3.4. User Involvement in Energy, Environment, Health and E-governement

Below are some topics where AxIS was or is involved in:

- **Preprocessing and analysing collective usage data and social networks** from group discussions related to design process: cf. ANR Intermed (2009) and FP7 Elliot (cf. section 6.3.1.1) where citizen generate ideas in terms of specific environmental sensors based services according to their needs.

- **Methods and tools for supporting open innovation based on public data**: a first work was made in 2010 with the CDISOD Color action related Public Data in collaboration with Fing (Marseille) and ADEME (Sophia Antipolis). We pursue such a study in the context of FP7 Elliot by providing to citizen environnemental data (air quality and noise) issued from citizen and/or territories sensors.

All AxIS topics are relevant for these domains. let us cite: social network analysis, personalization and information retrieval, recommender systems, expert search, design and evaluation of methods and tools for open innovation and user co-creation in the context of Living Labs, usage mining, mining data streams.

We have addressed specific works:

- Energy (cf. section 6.1.3): the main AxIS topic here was usage analysis in the context of an energy challenge in an enterprise (ECOFFICES) taking into account the complex and real situation (installation fo more than 400 sensors, differences between the three concerned teams, differences between the offices). Such an analysis aims to correlate team/office energy consuming, team/office eco-responsible behaviours and team/office profile.

- Health (cf. section 3.4): Axis contributed in 2011 to a Living Lab characterisation in Health domain through the visit of several Living Labs, which operate in the domain of Health and Autonomy, and conducted interviews. This work was done in relation with the CGIET [4].

- E-gov: The future Internet will bring a growing number of networked applications (services), devices and individual data (including private ones) to end-users. The important challenges are the organization of their access, and the guarantee of trust and privacy. The objectives of the PIMI [5] project (cf. section 6.2.1) are the definition of a design environment and a deployment platform for Personal Information Management system (PIM). The future PIM must provide the end-user personal data access with services that are relevant to his needs. In order to take mobility into account, the PIM will be accessed both by mobile devices (smartphone) and Personal Computers. With the increasing number of services and associated data being accessible through Internet, the number and complexity of PIM will augment dramatically in the near future. This will require strong research investment in a number of topics, all contributing to the expected usability and accessibility of Individual Information Spaces for the end-user.

---

[4]CGIET: "Conseil Général de l'Economie, de l'Industrie, de l'Energie et des technologies" http://www.cgeiet.economie.gouv.fr
[5]Personal Information Management through Internet

DAHU Project-Team

4. Application Domains

4.1. Application Domains

Databases are pervasive across many application fields. Indeed, most human activities today require some form of data management. In particular, all applications involving the processing of large amounts of data require the use of a database. Increasingly complex Web applications and services also rely on DBMS, and their correctness and robustness is crucial.

We believe that the automated solutions that Dahu aims to develop for verifying such systems will be useful in this context.

<p style="text-align:center; color:red">DREAM Project-Team</p>

4. Application Domains

4.1. Introduction

The DREAM research applications have been oriented towards surveillance of large networks as telecommunication networks and more recently of web services. During the past few years, we have focussed more and more on agricultural and environmental applications by means of research collaborations with INRA and Agrocampus Ouest.

4.2. Software components monitoring

software components, web services, distributed diagnosis

Web-services, i.e., services that are provided, controlled and managed through Internet, cover nowadays more and more application areas, from travel booking to goods supplying in supermarkets or the management of an e-learning platform. Such applications need to process requests from users and other services on line, and respond accurately in real time. Anyway, errors may occur, which need to be addressed in order to still be able to provide the correct response with a satisfactory quality of service (QoS): on-line monitoring, especially diagnosis and repair capabilities, become then a crucial concern.

We have been working on this problem within the WS-DIAMOND project [68], a large European funded project involving eight partners in Italy, France, Austria and Netherlands http://wsdiamond.di.unito.it/. Our own work consisted in two distinct contributions.

The first issue has been to extend the decentralized component-oriented approach, initially developed for monitoring telecommunication networks [4] to this new domain. To this end we have proposed the concept of distributed chronicles, with synchronization events, and the design of an architecture consisting of distributed CRSs (Chronicle Recognition Systems) communicating their local diagnoses to a broker agent which is in charge of merging them to compute a global diagnosis.

Our current work aims at coupling diagnosing and repair, in order to implement *adaptive web services*. We started this study by proposing an architecture inspired from the one developed during the WS-DIAMOND project and dedicated to the adaptive process of a request event when faults occur and propagate through the orchestration.

4.3. Environmental decision making

environment, decision methods

The need of decision support systems in the environmental domain is now well-recognized. It is especially true in the domain of water quality. For instance the program, named "Bretagne Eau Pure". was launched a few years ago in order to help regional managers to protect this important resource in Brittany. The challenge is to preserve the water quality from pollutants as nitrates and herbicides, when these pollutants are massively used by farmers to weed their agricultural plots and improve the quality and increase the quantity of their crops. The difficulty is then to find solutions which satisfy contradictory interests and to get a better knowledge on pollutant transfer.

In this context, we are cooperating with INRA (Institut National de Recherche Agronomique) and developing decision support systems to help regional managers in preserving the river water quality. The approach we advocate is to rely on a qualitative modeling, in order to model biophysical processes in an explicative and understandable way. The SACADEAU model associates a qualitative biophysical model, able to simulate the biophysical process, and a management model, able to simulate the farmer decisions. One of our main contribution is the use of qualitative spatial modeling, based on runoff trees, to simulate the pollutant transfer through agricultural catchments.

The second issue is the use of learning/data mining techniques to discover, from model simulation results, the discriminant variables and automatically acquire rules relating these variables. One of the main challenges is that we are faced with spatiotemporal data. The learned rules are then analyzed in order to recommend actions to improve a current situation.

This work has been done in the framework of the APPEAU project, funded by ANR and of the ACASSYA project, funded by ANR, having started at the beginning of 2009 and ended at the end of 2012. We were also involved in the PSDR GO CLIMASTER project, that started in september 2008 and end in 2011. CLIMASTER stands for "Changement climatique, systèmes agricoles, ressources naturelles et développement territorial" and is dedicated to the impact of climate changes on the agronomical behaviors in west of France ("Grand Ouest"). PSDR GO stands for "Programme Pour et Sur le Développpement Régional Grand Ouest".

Our main partners are the SAS INRA research group, located in Rennes and the BIA INRA and AGIR INRA research groups in Toulouse.

<p style="text-align:center;color:red;font-weight:bold;">EXMO Project-Team</p>

4. Application Domains

4.1. Application Domains

The main application context motivating our work is the "semantic web" infrastructure (§4.1.1), but it can be applied in any context where semantic technologies are used: semantic social networks, ambient intelligence, linked data, etc. [2].

4.1.1. Semantic web technologies

Internet technologies support organisations and people in accessing and sharing knowledge, often difficult to access in a documentary form. However, these technologies quickly reach their limits: web site organisation is expensive and full-text search inefficient. Content-based information search is becoming a necessity. Content representation will enable computers to manipulate knowledge on a more formal ground and to carry out similarity or generality search. Knowledge representation formalisms are good candidates for expressing content.

The vision of a "semantic web" [29] supplies the web, as we know it (informal) with annotations expressed in a machine-processible form and linked together. In the context where web documents are formally annotated, it becomes necessary to import and manipulate annotations according to their semantics and their use. Taking advantage of this semantic web will require the manipulation of various knowledge representation formats. Exmo concerns are thus central to the semantic web implementation. Our work aims at enhancing content understanding, including the intelligibility of communicated knowledge and formal knowledge transformations.

In addition, Exmo also considers a more specific use of semantic web technologies in semantic peer-to-peer systems, social semantic networks and ambient intelligence (see §6.3). In short, we would like to bring the semantic web to everyone's pocket. Semantic peer-to-peer systems are made of a distributed network of independent peers which share local resources annotated semantically and locally. This means that each peer can use its own ontology for annotating resources and these ontologies have to be confronted before peers can communicate. In social semantic networks, relationships between people are infered from relationships between knowledge they use. In ambient intelligence, applications have to reconcile device and sensor descriptions provided by independent sources.

Exmo's work can be implemented in software: in particular, we have developed an API for expressing ontology alignment (§5.1) and a library of ontology distances and similarities OntoSim (§5.2).

<p style="text-align:center; color:red; font-weight:bold;">GRAPHIK Project-Team</p>

4. Application Domains

4.1. Introduction

We currently focus on two application domains: knowledge representation in agronomy, more precisely applied to the quality in agri-food chains, and metadata management, in particular for bibliographic metadata.

The choice of the agronomy domain is motivated both by the local context of GraphIK (UMR IATE) and by its adequation to our research themes. Indeed, the agri-food domain seems to be particularly well-adapted to artificial intelligence techniques: there are no mathematical models available to solve the problems related to the quality of agrifood chains, which need to be stated at a more conceptual level; solving these problems requires an integrated approach that takes into account expert knowledge, which is typically symbolic, as well as numeric data, vague or uncertain information, multi-granularity knowledge, multiple and potentially conflicting viewpoints and actors.

The second area, metadata management, is not strictly speaking an application domain, but rather a cross-cutting axis. Indeed, metadata can be used to describe data in various areas (including for instance scientific publications in agronomy). We have a long experience in this domain, and we currently focus on document metadata.

4.2. Agronomy

Quality control within agri-food chains, but also non-food chains relies on numerous criteria (environmental, economical, functional, sanitary quality, etc.). The objectives of quality are based on several actors. The current structure of chains is questioned as for system perenniality, protection of the environment, cost and energy. In all cases, the following questions have to be taken into account:

1. the actors' viewpoints are divergent, hence it is necessary to define reasoning mechanisms able to model and take into account the balance between viewpoints, and the risks and benefits they imply;
2. the successive steps involved in a chain, impacting the quality of end products, have limiting factors. Their improvement is a complex objective that has no simple solution;
3. data from literature are dispersed and scattered, which makes their use difficult.

These questions highlight the need for an integrated approach of agri-food chains, respectively with symbolic reasoning mechanisms, reverse engineering methods, and knowledge organization and modelling.

Our general objective is the conception of a decision support tool for the actors of an agri-food chain, in presence of contradictory viewpoints and priorities, including the concepts of gravity and certainty of a risk or a benefit. The first step is to build a knowledge-based system able to represent the different kinds of knowledge needed, and provided with consistency checking, querying and symbolic simulation mechanisms, which will allow to refine and validate the modelling.

Our results in Sect. 6.1 and Sect. 6.2 can be seen as theoretical requirements towards this objective.

4.3. Document Metadata

Semantic metadata, in particular semantic annotations for multimedia documents, are at the core of the applications we are working on for several years. In the applications we developed in the previous years, mainly with INA (National Institute of Audiovisual) and FMSH (Fondation Maison des Sciences de l'Homme), we have built tools aimed at helping the manual construction of semantic annotations. In these projects, manual construction was unavoidable because semantically rich annotations, not obtainable by automatic processes, had to be built. In our current project Qualinca (see Sect. 8.1), the semantic metadata considered consists of information present in bibliographic databases and authority notices (which respectively describe documents and so-called authorities, such as authors typically). The challenge is not to build these metadata, which have been built by human specialists and already exist, but, for instance, to check their validity, to link or to merge different metadata bases.

MAIA Project-Team

4. Application Domains

4.1. Decision Making

Our group is involved in several applications of its more fondamental work on autonomous decision making and complex systems. Applications addressed include:

- Robotics, where the decision maker or agent is supported by a physical entity moving in the real world;
- Medicine or Personal Assisting Living, where the agent can be an analytic device recommending tests and/or treatments, or able to gather different sources of information (sensors for example) in order to help a final user, detecting for example anormal situation needing the rescue of a person (fall detection of elderly people, risk of hospitalization of a person suffering from chronic desease;
- Computer Security, where the agent can be a virtual attacker trying to identify security holes in a given network;
- and Business Process Management, where the agent can provide an auto-completion facility helping to decide which steps to include into a new or revised process.

4.2. Ambient intelligence

Taking into consideration some scientific strategic choices made by the Nancy – Grand Est Research Center such as developing some platforms around Robotics and Smart Living Apartments, some members of the team have recentered their research toward "ambient intelligence and AI" . This choice has been also comforted by the launch by Inria of a Large-scale initiative project termed PAL (Personal assistant Living) in which we are strongly involved. The regional council of Lorraine also supports this new research line through the CPER, (project "situated computing" or "INFOSITU" (http://infositu.loria.fr) whose the coordinator is a member of MAIA Team. Within this new domain of research in MAIA, we are thinking about the design of intelligent environments dedicated to eldely people with loss of autonomy. This domain of research is currently very active, taking up a societal challenge that developped countries have to address. In order to contribute to this resarch domain, we are currently developping several action lines both at the technological, scientific and service levels:

- Evaluation of the degree of frailty of the elderly : We are currently designing a system whose the objective is twofold : assessing the risk of fall and evaluating the degree of frailty of the elderly. This issue is considered with more or less sophisticated sensors : one kinect, a network of kinects, an heterogeneous sensor network made up of an intelligent floor and a network of kinects. One simple idea which is currently developped is to determine either the center of mass of a person using one or several kinect, or the center of pressure and footsteps localization using an intelligent floor. The idea is to induce from these simple measures, the walking speed, the length of the steps and the position of the monitored persons.
- People activity analysis: Sensitive or intelligent floors have attracted a lot of attention during the last two decades for different applications going from interaction capture in immersive virtual environments to robotics or human tracking, fall detection or activity recognition. Different technologies have been proposed so far either based on optical fiber sensing, pressure sensing or electrical near field. In PAL we envision a more more sophisticate approach in which both computation and sensing is distributed within the floor. This floor is made up of interconnected intelligent titles with can:
  - communicate with each other,
  - provide some computation,
  - sense the environment activity through four weight sensors, an accelerometer and a magnetometer,
  - interact with users, robots or other sensor networks either by wireless/wire communication or through visual communication (each tiles being equipped be 16 leds).

Several scientific challenges are open to us

– in decentralized spatial computing

– in designing real application for assisting people suffering from loss of autonomy

Concerning the second point, we envision several applications :

– evaluation of the degree of frailty of the elderly, especially evaluating the risk of fall

– activity recognition

– monitoring assistant robots

<p align="center" style="color:red">MOSTRARE Project-Team</p>

4. Application Domains

4.1. Context

XML transformations are basic to data integration: HTML to XML transformations are useful for information extraction from the Web; XML to XML transformations are useful for data exchange between Web services or between peers or between databases. Doan and Halevy [40] survey novel integration tasks that appear with the Semantic Web and the usage of ontologies. Therefore, the semi-automatic generation of XML transformations is a challenge in the database community and in the semantic Web community.

Also, XML transformations are useful for document processing. For instance, there is need of designing transformations from documents organized w.r.t visual format (HTML, DOC, PDF) into documents organized w.r.t. semantic format (XML according to a DTD or a schema). The semi-automatic design of such transformations is obviously a very challenging objective.

Furthermore, quite some activities of Mostrare concern efficient evaluation of XPath queries on XML documents and XML streams. XPath is fundamental to all XML standards, in particular to XQuery, XSLT, and XProc.

OAK Team

4. Application Domains

4.1. Online content management

This concerns archiving filtered content from online information sources (journals, blogs, ...) with the purpose of recording their perspective on facts involving specific countries, regions or enterprises, key actors etc. We have recently explored such an application in a demonstration proposal, built using our XR model for XML documents with semantic annotations (currently under evaluation). The benefit of XR was to simplify and streamline the specification of document- and semantics-rich data management applications. Support for documents allows to keep an evidence (source) for facts and statements that may be extracted from them. At the same time, support for semantics allows to connect documents to people, ideas, trends etc. and to reason across data sources [9].

4.2. Open data intelligence

Open data intelligence: the goal is to build and efficiently exploit warehouses of Open Data, integrated from several data sources on the Web, in order to produce consolidated rich information to be given to decision makers and to the citizens. Such projects have been started in France notably by the city areas of Rennes (pioneer in Open Data usage), Paris, and more recently by Grenoble, in a project to which we participate; the ICT Labs DataBridges also focused on such topics. Oak competencies required for such projects are related to large-scale RDF data management as well as to the design of innovative data models for semantic-rich content.

4.3. Efficient complex data management in the cloud

In a cloud environment, data catalogs and indices need to be efficiently built and maintained, typically in parallel; queries need to be routed to only those data subsets which are likely to lead to results, and efficiently executed, using parallelism and the available indexes. We work on such topics within the Europa ICT Labs activity. Algorithms for efficiently handling indexes and views in the cloud for heterogeneous, complex-structure data are also at the core of our work proposal in the Datalyse project (see below). Our 2012 work in this context has lead to [10], [12], [13].

4.4. 360 degree customer view

Companies seek to gather as much information as possible about their customers, for instance for more targeted publicity campaigns, market analysis, offer personalization, etc. That is, they want to consider data beyond the data they collected about their customers in their proprietary databases. Complementary data include for instance social data extracted from customers' activities in social networks, public data related to their place of residence (e.g., crime rate or housing price evolution). To achieve this goal, data integration on highly heterogeneous and massive data is necessary. Furthermore, as one means to assess both the correctness of the integration result and the quality (in this application most notably trustworthiness), we can resort to data provenance. These topics are explored in the project Datalyse which we submitted in 2012 to the French national "AAP Cloud 3: Big Data" call, a project headed by the "Business & Decision" company and whose ongoing evaluation will continue in early 2013.

<div align="center">**ORPAILLEUR Project-Team**</div>

4. Application Domains

4.1. Life Sciences

Participants: Yasmine Assess, Thomas Bourquard, Emmanuel Bresso, Marie-Dominique Devignes, Elias Egho, Anisah Ghoorah, Renaud Grisoni, Nicolas Jay, Bernard Maigret, Amedeo Napoli, Violeta Pérez-Nueno, Dave Ritchie, Malika Smaïl-Tabbone, Yannick Toussaint.

knowledge discovery in life sciences, bioinformatics, biology, chemistry, gene

> **Knowledge discovery in life sciences**   is a process for extracting knowledge units from large biological databases, e.g. collection of genes.

One major application domain which is currently investigated by Orpailleur team is related to life sciences, with particular emphasis on biology, medicine, and chemistry. The understanding of biological systems provides complex problems for computer scientists, and, when they exist, solutions bring new research ideas for biologists and for computer scientists as well. Accordingly, the Orpailleur team includes biologists, chemists, and a physician, making Orpailleur a very original EPI at Inria.

Knowledge discovery is gaining more and more interest and importance in life sciences for mining either homogeneous databases such as protein sequences and structures, or heterogeneous databases for discovering interactions between genes and environment, or between genetic and phenotypic data, especially for public health and pharmacogenomics domains. The latter case appears to be one main challenge in knowledge discovery in biology and involves knowledge discovery from complex data and thus KDDK. The interactions between researchers in biology and researchers in computer science improve not only knowledge about systems in biology, chemistry, and medicine, but knowledge about computer science as well. Solving problems for biologists using KDDK methods involves the design of specific modules that, in turn, leads to adaptations of the KDDK process, especially in the preparation of data and in the interpretation of the extracted units.

4.2. Knowledge Management in Medicine

Participants: Nicolas Jay, Jean Lieber, Thomas Meilender, Amedeo Napoli.

knowledge representation, description logics, classification-based reasoning, case-based reasoning, formal concept analysis, semantic web

The Kasimir research project holds on decision support and knowledge management for the treatment of cancer [103]. This is a multidisciplinary research project in which participate researchers in computer science (Orpailleur), experts in oncology ("Centre Alexis Vautrin" in Vandœuvre-lès-Nancy), Oncolor (a healthcare network in Lorraine involved in oncology), and A2Zi (a company working in Web technologies and involved in several projects in the medical informatics domain, http://www.a2zi.fr/). For a given cancer localization, a treatment is based on a protocol similar to a medical guideline, and is built according to evidence-based medicine principles. For most of the cases (about 70%), a straightforward application of the protocol is sufficient and provides a solution, i.e. a treatment, that can be directly reused. A case out of the 30% remaining cases is "out of the protocol", meaning that either the protocol does not provide a treatment for this case, or the proposed solution raises difficulties, e.g. contraindication, treatment impossibility, etc. For a case "out of the protocol", oncologists try to *adapt* the protocol. Actually, considering the complex case of breast cancer, oncologists discuss such a case during the so-called "breast cancer therapeutic decision meetings", including experts of all specialties in breast oncology, e.g. chemotherapy, radiotherapy, and surgery.

The semantic Web technologies have been used and adapted in the Kasimir project for several years. Recently, a semantic wiki has been deployed in its production version (http://www.oncologik.fr) It allows the management of decision protocols [48] More precisely, the migration from the static HTML site of Oncolor to a semantic wiki (with limited editing rights and unlimited reading rights) has been done [49]. This has consequences on the editorial chain of the published protocols which is more collaborative. A decision tree editor that has been integrated into the wiki and that has an export facility to formalized protocols in OWL DL has also been developed [61].

4.3. Cooking

Participants: Valmi Dufour-Lussier, Emmanuelle Gaillard, Laura Infante Blanco, Florence Le Ber, Jean Lieber, Amedeo Napoli, Emmanuel Nauer.

cooking, knowledge representation, knowledge discovery, case-based reasoning, semantic wiki

The origin of the Taaable project is the Computer Cooking Contest (CCC). A contestant of the CCC is a system that answers queries of recipes, using a recipe base; if no recipe exactly matches the query, then the system adapts another recipe. Taaable is a case-based reasoning system that uses various technologies used and developed in the Orpailleur team, such as technologies of the semantic web, knowledge discovery techniques, knowledge representation and reasoning techniques, etc. From a research viewpoint it enables to test the scientific results on an application domain that is at the same time simple to understand and raising complex issues, and to study the complementarity of various research domains. Taaable has been at the origin of the project Kolflow of the ANR CONTINT program, whose application domain is WikiTaaable, the semantic wiki of Taaable. It is also used for other projects under submission.

<p style="text-align:center; color:red">SMIS Project-Team</p>

4. Application Domains

4.1. Application Domains

data privacy, personal information management, healthcare, ambient intelligence

Our work addresses varied application domains. Typically, data management techniques on chip are required each time data-driven applications have to be embedded in ultra-light computing devices. This situation occurs for example in healthcare applications where medical folders are embedded into smart tokens (e.g., smart cards, secured USB keys), in telephony applications where personal data (address book, agenda, etc.) is embedded into cellular phones, in sensor networks where sensors log row measurements and perform local computation on them, in smart-home applications where a collection of smart appliances gather information about the occupants to provide them a personalized service, and more generally in most applications related to ambient intelligence.

Safeguarding data confidentiality has become a primary concern for citizens, administrations and companies, broadening the application domains of our work on access control policies definition and enforcement. The threat on data confidentiality is manifold: external and internal attacks on the data at rest, on the data on transit, on the data hosted in untrusted environments (e.g., Database Service Providers, Web-hosting companies) and subject to illegal usage, insidious gathering of personal data in an ambient intelligence surrounding. Hence, new access control models and security mechanisms are required to accurately declare and safely control who is granted access to which data and for which purpose.

While the application domain mentioned above is rather large, one application is today more specifically targeted by the SMIS project. This application deals with privacy preservation in EHR (Electronic Health Record) systems. Several countries (including France) launched recently ambitious EHR programs where medical folders will be centralized and potentially hosted by private Database Service Providers. Centralization and hosting increase the risk of privacy violation. In 2007, we launched two projects (PlugDB and DMSP) tackling precisely this issue, with the final objective to experiment our technologies in the field. In 2011, we launched a new project (KISS) capitalizing on the previous ones and extending their scope towards the protection of any personal data delivered to individuals in an electronic form.

WAM Project-Team

4. Application Domains

4.1. Introduction

Broadly speaking, the main application domain of our research is the web and its numerous applications. This includes the recent evolutions of the web, with a special attention paid to the mobile web, the multimedia web, and the web as a platform for applications. The goal of our research is to enable new multimedia and mobile applications that can be deployed easily on the web, taking advantage of the existing infrastructure and the latest advances in web technology.

More specifically, our work this year has focused on two main application domains: web development and pedestrian navigation.

4.2. Web Development

Current content representation practices and programming methods for the web remain severely limited. Designing web applications is becoming increasingly complex as it relies more and more on a jungle of programming languages, tools and data formats, each targeted toward the different application layers (presentation, application and storage). This often yields complex and opaque applications organized in silos, which are costly, inefficient, hard to maintain and evolve, and vulnerable to errors. In addition, the communication aspects are often handled independently via remote service invocations and represent another source of complexity and vulnerability.

Most research activities in WAM address these issues and try to cope with the fundamentals aspects of web applications (advanced content, data and communication) by studying rich document formats, data models and communication patterns, to offer correction guarantees and flexibility in the application architecture. For instance, applications need to be checked, optimized and managed as a whole while leveraging on the consistency of their individual components and data fragments.

4.3. Pedestrian Navigation

A number of factors are changing our thinking of an accessible town, namely the open data movement exemplified by OpenStreetMap, MEMS sensors embedded in mobile phones (accelerometers, magnetometers, gyroscopes), web and Augmented Reality technologies, increase in processing power of mobile phones. All these changes are allowing us to build energy efficient urban pedestrian navigation systems. These systems can now be based on Inertial Measurement Unit (IMU) and will run on mobile phones with customized embedded geographical data, a routing engine, and real time queries to urban information systems.

Richer and more precise AR mobile applications in such fields as cultural heritage visits, outdoor games, or guidance of people with disabilities will be enabled by MMG navigation, i.e. the join use of micro, macro and global navigation.

Micro-navigation builds upon embedded software ability to create a greater awareness of the immediate environment, using texture-based tracking or vision algorithms and relating this information to map and IMU data. Micro-navigation includes avoiding obstacles, locating a clear path in the proximate surroundings or at a complex crossing, finding objects and providing absolute positioning using known landmarks or beacons. Micro-navigation works at a precision level of a few centimeters by using predefined landmarks.

Macro-navigation refers to the actions required to find a route in a larger, not immediately perceptible environment, and builds upon carefully designed pedestrian ways incorporating speech instructions, audio guidance, environmental queries and IMU instructions among other things. Macro-navigation works at a precision level of one step using carefully designed routes with map-matching instructions.

Global navigation is based on an absolute global localization system like the GPS. Its precision is that of a few meters if used in a adequate geographical environment where data from external sensors is accessible. It can be used to bootstrap macro-navigation.

There is a duality relation between micro-navigation and macro-navigation. Micro-navigation is based on a localization system giving an absolute position which makes it possible to compute a relative position with respect to the planned route. Macro-navigation is based on a localization system giving a relative position which allows to compute an absolute position on the route through a process called map-matching. As a consequence, this two kinds of navigation complement and enhance each other.

WIMMICS Team

4. Application Domains

4.1. Introduction

A number of evolutions have changed the face of information systems in the past decade but the advent of the Web is unquestionably a major one and it is here to stay. From an initial wide-spread perception of a public documentary system, the Web as an object turned into a social virtual space and, as a technology, grew as an application design paradigm (services, data formats, query languages, scripting, interfaces, reasoning, etc.). The universal deployment and support of its standards led the Web to take over nearly all of our information systems. As the Web continues to evolve, our information systems are evolving with it.

Today in organizations, not only almost every internal information system is a Web application, but these applications also more and more often interact with external Web applications. The complexity and coupling of these Web-based information systems call for specification methods and engineering tools. From capturing the needs of users to deploying a usable solution, there are many steps involving computer science specialists and non-specialists.

We defend the idea of relying on Semantic Web formalisms to capture and reason on the models of these information systems supporting the design, evolution, interoperability and reuse of the models and their data as well as the workflows and the processing.

4.2. Linked Data on the Web and on Intranets

With billions of triples online (see Linked Open Data initiative), the Semantic Web is providing and linking open data at a growing pace and publishing and interlinking the semantics of their schemas. Information systems can now tap into and contribute to this Web of data, pulling and integrating data on demand. Many organisations also started to use this approach on their intranets leading to what is called linked enterprise data.

A first application domain for us is the publication and linking of data and their schemas through Web architectures. Our results provide software platforms to publish and query data and their schemas, to enrich these data in particular by reasoning on their schemas, to control their access and licences, to assist the workflows that exploit them, to support the use of distributed datasets, to assist the browsing and visualization of data, etc.

Examples of collaboration and applied projects include: Corese/KGRAM, Datalift, DBpedia, ALU/BLF Convention, ADT SeGViz.

4.3. Assisting Web-based Epistemic Communities

In parallel to linked open data on the Web, social Web applications also spread virally (e.g. Facebook growing toward 800 million users) first giving the Web back its status of a social read-write media and then leading it to its full potential of a virtual place where to act, react and interact. In addition, many organizations are now considering deploying social Web applications internally to foster community building, expert cartography, business intelligence, technological watch and knowledge sharing in general.

Reasoning on the linked data and the semantics of the schemas used to represent social structures and Web resources, we intend to provide applications supporting communities of practice and interest and fostering their interactions.

We use typed graphs to capture and mix: social networks with the kinds of relationships and the descriptions of the persons; compositions of Web services with types of inputs and outputs; links between documents with their genre and topics; hierarchies of classes, thesauri, ontologies and folksonomies; recorded traces and suggested navigation courses; submitted queries and detected frequent patterns; timelines and workflows; etc.

Our results assist epistemic communities in their daily activities such as biologists exchanging results, business intelligence and technological watch networks informing companies, engineers interacting on a project, conference attendees, students following the same course, tourists visiting a region, mobile experts on the field, etc. Examples of collaboration and applied projects include: Kolflow, OCKTOPUS, ISICIL, SAP Convention.

<p style="text-align:center">**ZENITH Project-Team**</p>

4. Application Domains

4.1. Data-intensive Scientific Applications

The application domains covered by Zenith are very wide and diverse, as they concern data-intensive scientific applications, i.e. most scientific applications. Since the interaction with scientists is crucial to identify and tackle data management problems, we are dealing primarily with application domains for which Montpellier has an excellent track record, i.e. agronomy, environmental science, life science, with scientific partners like INRA, IRD, CIRAD and IRSTEA (prev. CEMAGREF). However, we are also addressing other scientific domains (e.g. astronomy, oil extraction) through our international collaborations (e.g. in Brazil).

Let us briefly illustrate some representative examples of scientific applications on which we have been working on.

- **Management of astronomical catalogs**. An example of data-intensive scientific applications is the management of astronomical catalogs generated by the Dark Energy Survey (DES) project on which we are collaborating with researchers from Brazil. In this project, huge tables with billions of tuples and hundreds of attributes (corresponding to dimensions, mainly double precision real numbers) store the collected sky data. Data are appended to the catalog database as new observations are performed and the resulting database size is estimated to reach 100TB very soon. Scientists around the globe can query the database with queries that may contain a considerable number of attributes. The volume of data that this application holds poses important challenges for data management. In particular, efficient solutions are needed to partition and distribute the data in several servers. An efficient partitioning scheme should try to minimize the number of fragments accessed in the execution of a query, thus reducing the overhead associated to handle the distributed execution.

- **Pesticide reduction**. In a pesticide reduction application, with CEMAGREF, we plan to work on sensor data for plant monitoring. Sensors are used to observe the development of diseases and insect attacks in the agricultural farms, aiming at using pesticides only when necessary. The sensors periodically send to a central system their data about different measures such as plants contamination, temperature or moisture level. A decision support system analyzes the sent data, and triggers a pesticide treatment only when needed. However, the data sent by sensors are not entirely certain. The main reasons for uncertainty are the effect of climate events on sensors, e.g. rain, unreliability of the data transmission media, fault in sensors, etc. This requires to deal with uncertain data in modeling and querying to be used for data analysis and data mining.

- **Botanical data sharing**. Botanical data is highly decentralized and heterogeneous. Each actor has its own expertise domain, hosts its own data, and describes them in a specific format. Furthermore, botanical data is complex. A single plant's observation might include many structured and unstructured tags, several images of different organs, some empirical measurements and a few other contextual data (time, location, author, etc.). A noticeable consequence is that simply identifying plant species is often a very difficult task; even for the botanists themselves (the so-called taxonomic gap). Botanical data sharing should thus speed up the integration of raw observation data, while providing users an easy and efficient access to integrated data. This requires to deal with social-based data integration and sharing, massive data analysis and scalable content-based information retrieval. We address this application in the context of the French initiative Pl@ntNet, with CIRAD and IRD.

- **Deepwater oil exploitation**. An important step in oil exploitation is pumping oil from ultra-deepwater from thousand meters up to the surface through long tubular structures, called risers. Maintaining and repairing risers under deep water is difficult, costly and critical for the environment. Thus, scientists must predict risers fatigue based on complex scientific models and observed data for the risers. Risers fatigue analysis requires a complex workflow of data-intensive activities which

may take a very long time to compute. A typical workflow takes as input files containing riser information, such as finite element meshes, winds, waves and sea currents, and produces result analysis files to be further studied by the scientists. It can have thousands of input and output files and tens of activities (e.g. dynamic analysis of risers movements, tension analysis, etc.). Some activities, e.g. dynamic analysis, are repeated for many different input files, and depending on the mesh refinements, each single execution may take hours to complete. To speed up risers fatigue analysis requires parallelizing workflow execution, which is hard to do with existing SWfMS. We address this application in collaboration with UFRJ, and Petrobras.

These application examples illustrate the diversity of requirements and issues which we are addressing with our scientific application partners (CIRAD, INRA, CEMAGREF, etc.). To further validate our solutions and extend the scope of our results, we also want to foster industrial collaborations, even in non scientific applications, provided that they exhibit similar challenges.

<div align="center" style="color:red">**COPRIN Project-Team**</div>

4. Application Domains

4.1. Application Domains

While the methods developed in the project can be used for a very broad set of application domains (for example we have an activity in CO2 emission allowances [32]), it is clear that the size of the project does not allow us to address all of them. Hence we have decided to focus our applicative activities on *mechanism theory*, where we focus on *optimal design* and geometrical modeling of mechanisms. Along the same line our focus is *robotics* and especially *service robotics* which includes rescue robotics, rehabilitation and assistive robots for elderly and handicapped people (section 6.2.1.3). Although these topics were new for us in 2008 we have spent two years determining priorities and guidelines by conducting about 200 interviews with field experts (end-users, praticians, family and caregivers, institutes), establishing strong collaboration with them (e.g. with the CHU of Nice-Cimiez) and putting together an appropriate experimental setup for testing our solutions (see 6.2.1.4). A direct consequence of setting up this research framework is a reduction in our publication and contract activities. But this may be considered as an investment as assistance robotics will constitute the major research axis of the project on the long term (COPRIN will reach its twelve years of existence in 2014 but we intend to propose in 2013 a new project on this topic, code-named HEPHAISTOS).

E-MOTION Project-Team

3. Application Domains

3.1. Introduction

The main applications of our research are those aiming at introducing advanced and secured robotized systems into human environments. In this context, we are focusing onto the following application domains: Future cars and transportation systems, Service and Human assistance robotics, and Potential spin-offs in some other application domains.

3.2. Future cars and transportation systems

Thanks to the introduction of new sensor and ICT technologies in cars and in mass transportation systems, and also to the pressure of economical and security requirements of our modern society, this application domain is quickly changing. Various technologies are currently developed by both research and industrial laboratories. These technologies are progressively arriving at maturity, as it is witnessed by the results of large scale experiments and challenges (e.g., Darpa Urban Challenge 2007) and by the fast development of ambitious projects such as the Google's car project. Moreover, the legal issue starts to be addressed (see for instance the recent laws in Nevada and in California authorizing autonomous vehicles on roads).

In this context, we are interested in the development of *ADAS*[1] systems aimed at improving comfort and safety of the cars users (e.g., ACC, emergency braking, danger warnings), and of *Fully Autonomous Driving* functions for controlling the displacements of private or public vehicles in some particular driving situations and/or in some equipped areas (e.g., automated car parks or captive fleets in downtown centers or private sites).

3.3. Service, intervention, and human assistance robotics

This application domain is currently quickly emerging, and more and more industrials companies (e.g., IS-Robotics, Samsung, LG) are now commercializing service and intervention robotics products such as vacuum cleaner robots, drones for civil or military applications, entertainment robots . . .). One of the main challenges is to propose robots which are sufficiently robust and autonomous, easily usable by non-specialists, and marked at a reasonable cost. A more recent challenge for the coming decade is to develop robotized systems for assisting elderly and/or disabled people. We are strongly involved in the development of such technologies, which are clearly tightly connected to our research work on robots in human environments.

3.4. Potential spin-offs in some other application domains

Our *Bayesian Programming* tools (including the functions for decision making under uncertainty) are also impacting a large spectrum of application domains such as autonomous systems, surveillance systems, preventive maintenance for large industrial plants, fraud detection, video games, etc. These application domains are covered by our start-up *Probayes*.

---

[1]Advanced Driver Assistance Systems

<center>**FLOWERS Project-Team**</center>

4. Application Domains

4.1. Application Domains

- **Personal robotics.** Many indicators show that the arrival of personal robots in homes and everyday life will be a major fact of the 21st century. These robots will range from purely entertainment or educative applications to social companions that many argue will be of crucial help in our aging society. For example, UNECE evaluates that the industry of entertainment, personal and service robotics will grow from \$5.4Bn to \$17.1Bn over 2008-2010. Yet, to realize this vision, important obstacles need to be overcome: these robots will have to evolve in unpredictable homes and learn new skills while interacting with non-engineer humans after they left factories, which is out of reach of current technology. In this context, the refoundation of intelligent systems that developmental robotics is exploring opens potentially novel horizons to solve these problems.

- **Human-Robot Collaboration.** Robots play a vital role for industry and ensure the efficient and competitive production of a wide range of goods. They replace humans in many tasks which otherwise would be too difficult, too dangerous, or too expensive to perform. However, the new needs and desires of the society call for manufacturing system centered around personalized products and small series productions. Human-robot collaboration could widen the use of robot in this new situations if robots become cheaper, easier to program and safe to interact with. The most relevant systems for such applications would follow an expert worker and works with (some) autonomy, but being always under supervision of the human and acts based on its task models.

- **Video games.** In conjunction with entertainment robotics, a new kind of video games are developing in which the player must either take care of a digital creature (e.g. Neopets), or tame it (e.g. Nintendogs), or raise/accompany them (e.g. Sims). The challenges entailed by programming these creatures share many features with programming personal/entertainment robots. Hence, the video game industry is also a natural field of application for FLOWERS.

- **Environment perception in intelligent vehicles.** When working in simulated traffic environments, elements of FLOWERS research can be applied to the autonomous acquisition of increasingly abstract representations of both traffic objects and traffic scenes. In particular, the object classes of vehicles and pedestrians are if interest when considering detection tasks in safety systems, as well as scene categories ("scene context") that have a strong impact on the occurrence of these object classes. As already indicated by several investigations in the field, results from present-day simulation technology can be transferred to the real world with little impact on performance. Therefore, applications of FLOWERS research that is suitably verified by real-world benchmarks has direct applicability in safety-system products for intelligent vehicles.

- **Automated Tutoring Systems.** Optimal teaching and efficient teaching/learning environments can be applied to aid teaching in schools aiming both at increase the achievement levels and the reduce time needed. From a practical perspective, improved models could be saving millions of hours of students' time (and effort) in learning. These models should also predict the achievement levels of students in order to influence teaching practices.

IMARA Project-Team

4. Application Domains

4.1. Introduction

While the preceding section focused on methodology, in connection with automated guided vehicles, it should be stressed that the evolution of the problems which we deal with remains often guided by the technological developments. We enumerate three fields of application, whose relative importance varies with time and who have strong mutual dependencies: driving assistance, cars available in self-service mode and fully automated vehicles (cybercars).

4.2. Driving assistance

Several techniques will soon help drivers. One of the first immediate goal is to improve security by alerting the driver when some potentially dangerous or dangerous situations arise, i.e. collision warning systems or lane tracking could help a bus driver and surrounding vehicle drivers to more efficiently operate their vehicles. Human factors issues could be addressed to control the driver workload based on additional information processing requirements.

Another issue is to optimize individual journeys. This means developing software for calculating optimal (for the user or for the community) path. Nowadays, path planning software is based on a static view of the traffic: efforts have to be done to take the dynamic component in account.

4.3. New transportation systems

The problems related to the abusive use of the individual car in large cities led the populations and the political leaders to support the development of public transport. A demand exists for a transport of people and goods which associates quality of service, environmental protection and access to the greatest number. Thus the tram and the light subways of VAL type recently introduced into several cities in France conquered the populations, in spite of high financial costs.

However, these means of mass transportation are only possible on lines on which there is a keen demand. As soon as one moves away from these "lines of desire" or when one deviates from the rush hours, these modes become expensive and offer can thus only be limited in space and time.

To give a more flexible offer, it is necessary to plan more individual modes which approach the car as we know it. However, if one wants to enjoy the benefits of the individual car without suffering from their disadvantages, it is necessary to try to match several criteria: availability anywhere and anytime to all, lower air and soils pollution as well as sound levels, reduced ground space occupation, security, low cost.

Electric or gas vehicles available in self-service as in the Praxitèle system bring a first response to these criteria. To be able to still better meet the needs, it is however necessary to re-examine the design of the vehicles on the following points:

- ease empty car moves to better distribute them;
- better use of information systems inboard and on ground;
- better integrate this system in the global transportation system.

These systems are now operating (i.e. in La Rochelle). The challenge is to bring them to an industrial phase by transferring technologies to these still experimental projects.

4.4. Cybercars

The long term effort of the project is to put automatically guided vehicles (cybercars) on the road. It seems too early to mix cybercars and traditional vehicles, but data processing and automation now make it possible to consider in the relatively short term the development of such vehicles and the adapted infrastructures. IMARA aims at using these technologies on experimental platforms (vehicles and infrastructures) to accelerate the technology transfer and to innovate in this field.

Other application can be precision docking systems that will allow buses to be automatically maneuvered into a loading zone or maintenance area, allowing easier access for passengers, or more efficient maintenance operations. Transit operating costs will also be reduced through decreased maintenance costs and less damage to the breaking and steering systems.

Regarding technical topics, several aspects of Cybercars have been developed at IMARA this year. First, we have stabilized a generic Cycab architecture involving Inria Syndex tool and CAN communications. The critical part of the vehicle is using a real time Syndex application controlling the actuators via two Motorola's MPC555. Today, we have decided to migrate to the new dsPIC architecture for more efficiency and ease of use.

This application has a second feature, it can receive commands from an external source (Asynchronously this time) on a second CAN bus. This external source can be a PC or a dedicated CPU, we call it high level. To work on the high level, in the past years we have been developing a R&D framework called (Taxi) which used to take control of the vehicle (Cycab and Yamaha) and process data such as gyro, GPS, cameras, wireless communications and so on. Today, in order to rely on a professional and maintained solution, we have chosen to migrate to the RTMAPS SDK development platform. Today, all our developments and demonstrations are using this efficient prototyping platform. Thanks to RT-MAPS we've been able to do all the demonstrations on our cybercars: cycabs, Yamaha AGV and new Cybus platforms. These demonstrations include: reliable SLAMMOT algorithm using 2 to 4 laser sensors simultaneously, automatic line/road following techniques, PDA remote control, multi sensors data fusion, collaborative perception via ad-hoc network.

The second main topic is inter-vehicle communications using ad-hoc networks. We have worked with the HIPERCOM team for setting and tuning OLSR, a dynamic routing protocol for vehicles communications (see Section 3.2). Our goal is to develop a vehicle dedicated communication software suite, running on a specialized hardware. It can be linked also with the Taxi Framework for getting data such GPS information's to help the routing algorithm.

LAGADIC Project-Team

4. Application Domains

4.1. Overview

The natural applications of our research are obviously in robotics. In fact, researches undertaken in the Lagadic group can apply to all the fields of robotics implying a vision sensor. They are indeed conceived to be independent of the system considered (and the robot and the vision sensor can even be virtual for some applications).

Currently, we are mostly interested in using visual servoing for aerial and space application, micromanipulation, autonomous vehicle navigation in large urban environments or for disabled or elderly people.

We also address the field of medical robotics. The applications we consider turn around new functionalities of assistance to the clinician during a medical examination: visual servoing on echographic images, needle insertion, compensation of organ motions, etc.

Robotics is not the only possible application field to our researches. In the past, we were interested in applying visual servoing in computer animation, either for controlling the motions of virtual humanoids according to their pseudo-perception, or for controlling the point of view of visual restitution of an animation. In both cases, potential applications are in the field of virtual reality, for example for the design of video games, or virtual cinematography.

Applications also exist in computer vision and augmented reality. It is then a question of carrying out a virtual visual servoing for the 3D localization of a tool with respect to the vision sensor, or for the estimation of its 3D motion. This field of application is very promising, because it is in full rise for the realization of special effects in the multi-media field or for the design and the inspection of objects manufactured in the industrial world.

<div align="center" style="color:red">**AYIN Team**</div>

4. Application Domains

4.1. Remote sensing

The first application domain of the AYIN team concerns analysis and classification of remote sensing images. The very high spatial, spectral and temporal resolution of the last generation of imaging sensors (for instance, GeoEye, Ikonos, Pleiades, COSMO-SkyMed, TerraSAR-X, ...) provides rich information about environment and is very useful in a range of applications, such as investigating urban environments, precision agriculture, natural disasters and mineralogy. The development of these applications presents new challenges of high-dimensional and high-volume data analysis. The methods proposed by the AYIN team are applied for analysis of SAR, multi- and hyperspectral remote sensing images. In particular, the team develops approaches for image segmentation and classification, change detection, extraction of structures and object tracking.

4.2. Skin care

The second application domain of the AYIN team is skin care imaging which mainly consists in image analysis and classification for dermatology and cosmetology. Here we also deal with very high spatial, spectral and temporal resolution of the most recent imaging sensors. In dermatology we are particularly interested in hyperpigmentation detection and disorders severity evaluation (for instance, for melasma, acne, ...). In cosmetology our main goals are analysis, modeling and characterization of the condition of human skin, as well as evaluation of means to influence that condition. Some of the changes in skin over time have to do with chronological aging (such as pinheads for teenagers or wrinkles for mature people), others with extrinsic aging, caused for instance by sun exposure and smoking.

IMEDIA2 Team

4. Application Domains

4.1. Scientific applications

Examples: environmental images databases: fauna and flora; satellite images databases: ground typology; medical images databases: find images of a pathological character for educational or investigation purposes.

We are developing tools enabling multimedia access to biodiversity collections for species identifications inside the Pl@ntNet project.

In fact, almost all IMEDIA2 team members are involved in the Pl@ntNet project and develop methods that can be used in this context.

4.2. Audio-visual applications

Examples: Look for a specific shot in a movie, documentary or TV news, present a video summary. Help archivists to annotate the contents. Retrieve copies of a given material (photo or video) in a TV stream or on the web.

Our team has a collaboration with AFP press agency in the context of GLOCAL project.

<p style="text-align:center;color:red">LEAR Project-Team</p>

4. Application Domains

4.1. Application Domains

A solution to the general problem of visual recognition and scene understanding will enable a wide variety of applications in areas including human-computer interaction, retrieval and data mining, medical and scientific image analysis, manufacturing, transportation, personal and industrial robotics, and surveillance and security. With the ever expanding array of image and video sources, visual recognition technology is likely to become an integral part of many information systems. A complete solution to the recognition problem is unlikely in the near future, but partial solutions in these areas enable many applications. LEAR's research focuses on developing basic methods and general purpose solutions rather than on a specific application area. Nevertheless, we have applied our methods in several different contexts.

Semantic-level image and video access. This is an area with considerable potential for future expansion owing to the huge amount of visual data that is archived. Besides the many commercial image and video archives, it has been estimated that as much as 96% of the new data generated by humanity is in the form of personal videos and images [1], and there are also applications centering on on-line treatment of images from camera equipped mobile devices (e.g. navigation aids, recognizing and answering queries about a product seen in a store). Technologies such as MPEG-7 provide a framework for this, but they will not become generally useful until the required mark-up can be supplied automatically. The base technology that needs to be developed is efficient, reliable recognition and hyperlinking of semantic-level domain categories (people, particular individuals, scene type, generic classes such as vehicles or types of animals, actions such as football goals, etc). In a collaboration with Xerox Research Center Europe, supported by a CIFRE grant from ANRT, we study cross-modal retrieval of images given text queries, and vice-versa. In the context of the Microsoft-Inria collaboration we concentrate on retrieval and auto-annotation of videos by combining textual information (scripts accompanying videos) with video descriptors. In the EU FP7 project AXES we will further mature such video annotation techniques, and apply them to large archives in collaboration with partners such as the BBC, Deutsche Welle, and the Netherlands Institute for Sound and Vision.

Visual (example based) search. The essential requirement here is robust correspondence between observed images and reference ones, despite large differences in viewpoint or malicious attacks of the images. The reference database is typically large, requiring efficient indexing of visual appearance. Visual search is a key component of many applications. One application is navigation through image and video datasets, which is essential due to the growing number of digital capture devices used by industry and individuals. Another application that currently receives significant attention is copyright protection. Indeed, many images and videos covered by copyright are illegally copied on the Internet, in particular on peer-to-peer networks or on the so-called user-generated content sites such as Flickr, YouTube or DailyMotion. Another type of application is the detection of specific content from images and videos, which can, for example, be used for finding product related information given an image of the product. Transfer of such techniques is the goal of the start-up MilPix, to which our current technologies for image search are licensed. In a collaboration with Technosens we transfer face recognition technology, which they exploit to identify users of a system and adapt the interface to the user.

Automated object detection. Many applications require the reliable detection and localization of one or a few object classes. Examples are pedestrian detection for automatic vehicle control, airplane detection for military applications and car detection for traffic control. Object detection has often to be performed in less common imaging modalities such as infrared and under significant processing constraints. The main challenges are the relatively poor image resolution, the small size of the object regions and the changeable appearance of the objects. Our industrial project with MBDA is on detecting objects under such conditions in infrared images.

---

[1]http://www.sims.berkeley.edu/research/projects/how-much-info/summary.html

<div align="center">

MAGRIT Project-Team

</div>

4. Application Domains

4.1. Augmented reality

We have a significant experience in the AR field especially through the European project ARIS (2001–2004) which aimed at developing effective and realistic AR systems for e-commerce and especially for interior design. Beyond this restrictive application field, this project allowed us to develop nearly real time camera tracking methods for multi-planar environments. Since then, we have amplified our research on multi-planar environments in order to obtain effective and robust AR systems in such environments. We currently investigate both automatic and interactive techniques for scene reconstruction/structure from motion methods in order to be able to consider large and unknown environments.

4.2. Medical Imaging

For 15 years, we have been working in close collaboration with University Hospital of Nancy and GE Healthcare in interventional neuroradiology. Our common aim is to develop a multimodality framework to help therapeutic decisions and interventional gestures. In particular, we aim at developing tools allowing the physicians to take advantage of the various existing imaging modalities on the brain in their clinical practice: 2D subtracted angiography (2DSA), 3D rotational angiography (3DRA), fluoroscopy, MRI,...Recent works concern the use of AR tools for neuronavigation and the development of simulation tools of the interventional act for training or planning. This last project is developed in collaboration with the EPI Shacra.

4.3. Augmented head

Visual information on a speaker, especially jaws and lips but also tongue position, noticeably improves speech intelligibility. Hence, having a realistic augmented head displaying both external and internal articulators could help language learning technology progress in giving the student a feedback on how to change articulation in order to achieve a correct pronunciation. The long term aim of the project is the acquisition of articulatory data and the design of a 3D+t articulatory model from various image modalities: external articulators are extracted from stereovision data, the tongue shape is acquired through ultrasound imaging, 3D images of all articulators can be obtained with MRI for sustained sounds, magnetic sensors are used to recover the tip of the tongue.

MORPHEO Team

4. Application Domains

4.1. 4D modeling

Modeling shapes that evolve over time, analyzing and interpreting their motion has been a subject of increasing interest of many research communities including the computer vision, the computer graphics and the medical imaging communities. Recent evolutions in acquisition technologies including 3D depth cameras (Time-of-Light and Kinect), multi-camera systems, marker based motion capture systems, ultrasound and CT scans have made those communities consider capturing the real scene and their dynamics, create 4D spatio-temporal models, analyze and interpret them. A number of applications including dense motion capture, dynamic shape modeling and animation, temporally consistent 3D reconstruction, motion analyses and interpretation have therefore emerged.

4.2. Shape Analysis

Most existing shape analysis tools are local, in the sense that they give local insight about an object's geometry or purpose. The use of both geometry and motion clues makes it possible to recover more global information, in order to get extensive knowledge about a shape. For instance, motion can help to decompose a 3D model of a character into semantically significant parts, such as legs, arms, torso and head. Possible applications of such high-level shape understanding include accurate feature computation, comparison between models to detect defects or medical pathologies, and the design of new biometric models or new anthropometric datasets.

4.3. Human Motion Analysis

The recovery of dense motion information enables the combined analyses of shapes and their motions. Typical examples include the estimation of mean shapes given a set of 3D models or the identification of abnormal deformations of a shape given its typical evolutions. The interest arises in several application domains where temporal surface deformations need to be captured and analysed. It includes human body analyses for which potential applications with are anyway numerous and important, from the identification of pathologies to the design of new prostheses.

4.4. Interaction

The ability to build models of humans in real time allows to develop interactive applications where users interact with virtual worlds. The recent Kinect proposed by Microsoft illustrates this principle with game applications using human inputs perceived with a depth camera. Other examples include gesture interfaces using visual inputs. A challenging issue in this domain is the ability to capture complex scenes in natural environments. Multi-modal visual perception, e.g. depth and color cameras, is one objective in that respect.

<div align="center">**PERCEPTION Team**</div>

4. Application Domains

4.1. Human action recognition

We are particularly interested in the analysis and recognition of human actions and gestures. The vast majority of research groups concentrate on isolated action recognition. We address continuous recognition. The problem is difficult because one has to simultaneously address the problems of recognition and segmentation. For this reason, we adopt a per-frame representation and we develop methods that rely on dynamic programming and on hidden Markov models. We investigate two type of methods: one-pass methods and two-pass methods. One-pass methods enforce both within-action and between-action constraints within sequence-to-sequence alignment algorithms such as dynamic time warping or the Viterbi algorithm. Two-pass methods combine a per-action representation with a discriminative classifier and with a dynamic programming post-processing stage that find the best sequence of actions. These algorithms were well studied in the context of large-vocabulary continuous speech recognition systems. We investigate the modeling of various per-frame representations for action and gesture analysis and we devise one-pass and two-pass algorithms for recognition.

4.2. 3D reconstruction using TOF and color cameras

TOF cameras are active-light range sensors. An infrared beam of light is generated by the device and depth values can be measured by each pixel, provided that the beam travels back to the sensor. The associated depth measurement is accurate if the sensed surface sends back towards the sensor a fair percentage of the incident light. There is a large number of practical situations where the depth readings are erroneous: specular and bright surfaces (metal, plastic, etc.), scattering surfaces (hair), absorbing surfaces (cloth), slanted surfaces, e.g., at the bounding contours of convex objects which are very important for reconstruction, mutual reflections, limited range, etc. The resolution of currently available TOF cameras is of 0.3 to 0.5MP. Modern 2D color cameras deliver 2MP images at 30FPS or 5MP images at 15FPS. It is therefore judicious to attempt to combine the active-range and the passive-stereo approaches within a mixed methodology and system. Standard stereo matching methods provide an accurate depth map but are often quite slow because of the inherent complexity of the matching algorithms. Moreover, stereo matching is ambiguous and inaccurate in the presence of weakly textured areas. We develop TOF-stereo matching and reconstruction algorithms that are able to combine the advantages of the two types of depth estimation technologies.

4.3. Sound-source separation and localization

We explore the potential of binaural audition in conjunction with modern machine learning methods in order to address the problems of sound source separation and localization. We exploit the spectral properties of interaural cues, namely the interaural level difference (ILD) and the interaural phase difference (IPD). We have started to develop a novel supervised framework based on a training stage. During this stage, a sound source emits a broadband random signal which is perceived by a microphone pair embedded into a dummy head with a human-like head related transfer function (HRTF). The source emits from a location parameterized by azimuth and elevation. Hence, a mapping between a high-dimensional interaural spectral representation and a low-dimensional manifold can be estimated from these training data. This allows the development of various single-source localization methods as well as multiple-source separation and localization methods.

4.4. Audio-visual fusion for human-robot interaction

Modern human-robot interaction systems must be able to combine information from several modalities, e.g., vision and hearing, in order to allow high-level communication via gesture and vocal commands, multimodal dialogue, and recognition-action loops. Auditory and visual data are intrinsically different types of sensory data. We have started the development of a audio-visual mixture model that takes into account the heterogenous nature of visual and auditory observations. The proposed multimodal model uses modality specific mixtures (one mixture model for each modality). These mixtures are tied through latent variables that parameterize the joint audiovisual space. We thoroughly investigate this novel kind of mixtures with their associated efficient parameter estimation procedures.

PRIMA Project-Team

4. Application Domains

4.1. The Smart Spaces Research Plateform

Participants: Rémi Barraquand, James Crowley [correspondant], Augustin Lux, Amaury Nègre, Patrick Reignier, Lukas Rummelhard, Dominique Vaufreydaz.

Smart Spaces, Observation of human activity, context aware systems and services.

Project PRIMA has recently moved to a new Smart Spaces Research Plateform in order to develop and test components and services for context aware human centered services. The Smart Spaces Research Plateform is a 50 Square Meter space equipped with a microphone array, wireless lapel microphones, wide angle surveillance cameras, panoramic cameras, steerable cameras, scanning range sensors and two camera-projector video-interaction devices, and a KNX smart electrical system. The microphone array is used as an acoustic sensor to detect, locate and classify acoustic signals for recognizing human activities. The wide-angle and panoramic cameras provide fields of view that cover the entire room, and allows detection and tracking of individuals. Steerable cameras are used to acquire video of activities from any viewing direction.

Context aware human centered services may categorized as tools, advisors, or media. Tool services are designed to perform a specific task or function as robustly as possible. If any adaptation is involved, it should serve to adapt the function to a changing environment. The user interface, and any interaction with users should be perfectly predictable. The degree to which the operation of a tool should be transparent, visible or hidden from the user is an open research question. Advisor services observe the users actions and environment in order to propose information on possible courses of actions. Advisors should be completely obedient and non-disruptive. They should not take initiatives or actions that cannot be overridden or controlled by the user. Media services provide interpersonal communications, entertainment or sensorial extension.

Examples of human centered tool services include:

> An activity log recording system that records the events and activities of an individual's daily activities.

> A service that integrates control of heating, air-conditioning, lighting, windows, window-shades, exterior awnings, etc to provide an optimum comfort level defined in terms of temperature, humidity, CO2, acoustic noise and ambient light level.

> A service that manages the available stock of supplies in a home and orders supplies over the Internet to assure that the appropriate level of supplies are always available.

> A service to measure the walking rate, step size and posture of an elderly person to estimate health and predict the likelihood of a fall.

Some examples of advisor services include:

> A service that provides shopping advice about where and when to shop.

> A service that can propose possible menus based on the available food stuffs in the kitchen.

> A service that observes the activities of humans and appliances within the home and can suggest ways to reduce the cost of heating, electricity or communications.

> A service that observes lifestyle and can offer advice about improving health.

Some examples of media services include

> A service that maintains a sense of informal non-disruptive presence with distant family members.

> A robot device that communicates affection.

> A device that renders the surface temperature of wall, floors and windows to show energy consumption and loss within a house.

> Services that enable seamless tele-presence for communication with others

4.2. Ambient Services for Assisted Living

Participants: Rémi Barraquand, James Crowley, Frédéric Devernay, Amaury Nègre.

Ambient Assisted Living, Monitoring Services, Presence awareness.

The continued progress in extending life-span, coupled with declining birth rates have resulted in a growing number of elderly people with varying disabilities who are unable to conduct a normal life at home, thereby becoming more and more isolated from society. Governmental agencies including hospitals, healthcare institutions and social care institutions are increasingly overburdened with care of this growing population. Left unchecked, economic and man-power requirements for care of the elderly could well trigger a societal and economic crisis. There is an urgent societal need for technologies and services that allow elderly people to live autonomously in their own environments for longer periods. Smart environments provide a promising new enabling technology for such services.

Adapting smart environments to enhance the autonomy and quality of life for elderly require:

Robust, plug-and-play sensor technologies monitor the activities and health of elderly in their own home environments.

Easy to use communications services that allow people to maintain a sense of presence to avoid isolation without disrupting privacy or distracting attention from normal daily activities.

Architectural frameworks that allow ad hoc composition of services from distributed heterogeneous components scattered throughout the environment.

Distributed system architectures which allow the cooperation of independent emergency services to work together to provide emergency care,

Technologies interpret activity to warn of loss of mobility or cognitive function.

Engineering approaches for the customization/personalization/adaptation of living assistance systems at installation and run time,

Social, privacy, ethical and legal safeguards for privacy and control of personal data.

4.3. 3-D video processing

Participants: Frédéric Devernay [correspondant], Sylvain Duchêne, Sergi Pujades-Rocamora, Matthieu Volat.

3-D display, Stereoscopy, View Interpolation, Auto-calibration

Stereoscopic cinema has seen a surge of activity in recent years, and for the first time all of the major Hollywood studios released 3-D movies in 2009. This is happening alongside the adoption of 3-D technology for sports broadcasting, and the arrival of 3-D TVs for the home. Two previous attempts to introduce 3-D cinema in the 1950s and the 1980s failed because the contemporary technology was immature and resulted in viewer discomfort. But current technologies such as accurately-adjustable 3-D camera rigs with onboard computers to automatically inform a camera operator of inappropriate stereoscopic shots, digital processing for post-shooting rectification of the 3-D imagery, digital projectors for accurate positioning of the two stereo projections on the cinema screen, and polarized silver screens to reduce cross-talk between the viewers left- and right-eyes mean that the viewer experience is at a much higher level of quality than in the past. Even so, creation of stereoscopic cinema is an open, active research area, and there are many challenges from acquisition to post-production to automatic adaptation for different-sized display [41], [43].

Until recently, in order to view stereoscopic 3-D video, the user had to wear special glasses. Recent advances in 3-D displays provide true 3-D viewing experience without glasses. These screens use either a micro-lenticular network or a parallax barrier placed in front of a standard LCD, plasma, or LED display, so that different viewpoints provide different images. If the characteristics of the network and the screen are carefully chosen, the user will perceive two different images from the viewpoints of the left and right eyes. Such glasses-free 3-D screens usually display between 8 and a few dozen different viewpoints.

When the 3-D scene which has to be displayed is computer-generated, it is usually not a problem to generate a few dozen viewpoints. But when a real scene has to displayed, one would have to shoot it through the same number of synchronized cameras as there are viewpoints in order to display it properly. This makes 3-D shooting of real scenes for glasses-free 3-D displays mostly unpractical. For this reason, we are developing high-quality view-interpolation techniques, so that the many different viewpoints can be generated from only a few camera positions [42].

Our research focuses on algorithms derived from Computer Vision and Computer Graphics, applied to live-action stereoscopic 3-D content production or post-production, including [40]:

Live monitoring of stereoscopic video: geometric image misalignment, depth budget (i.e. limits on horizontal disparity), left-right color balance, left-right depth-of-field consistency [59].

Live correction of stereoscopic video: correct the above defects in real-time when it is possible, with the help of GPU-based architectures.

Adaptation of the stereoscopic content to the display size and distance, to avoid divergence or geometric deformations [42].

Novel camera setups and algorithms for unconstrained stereoscopic shooting (especially when using long focal length).

Novel camera setups and algorithms for glasses-free 3D displays.

Stereoscopic inpainting.

Stereoscopic match-moving.

Compositing stereoscopic video and matte painting without green screen.

Relighting of stereoscopic video, especially when videos are composited.

4.4. User localization in large-scale Smart Spaces

Participants: Amaury Nègre, Lukas Rummelhard, Dominique Vaufreydaz [correspondant].

Multi-modal perception, Smart Spaces, Localisation.

Ad-hoc assemblies of mobile devices embedding sensing, display, computing, communications, and interaction provide an enabling technology for smart environments. In the PRIMA project we have adopted a component oriented programming approach to compose smart services for such environments. Common services for smart spaces include

Services to manage energy in building, including regulating temperature, illumination, and acoustic noise,

Ambient assisted living services to extend the autonomy of elderly and infirm,

Logistics management for daily living,

Communication services and tools for collaborative work,

Services for commercial environments,

Orientation and information services for public spaces, and

Services for education and training.

We are pursuing development of components based on the concept of "large-scale" smart space that is an intelligent environment which will be deployed on a large surface containing several buildings (as a university campus for example). We also define the "augmented man" concept as a human wearing one or many mobile intelligent wireless devices (telephone, Smartphone, pda, notebook). Using all these devices, one can use many different applications (read emails, browse the Internet, file exchange, etc.). By combining the concepts of large-scale perceptive environments and mobile computing, we can create intelligent spaces, it becomes possible to propose services adapted to individuals and their activities. We are currently focussing on two aspects of this problem: the user profile and the user location within a smart space.

A fundamental requirement for such services is the ability to perceive the current state of the environment. Depending on the nature of the service, environment state can require sensing and modeling the physical properties of the environment, the location, identity and activity of individuals within the environment, as well as the set of available computing devices and software components that compose the environment. All of these make up possible elements for context modeling.

Observing and tracking people in smart environments remains a challenging fundamental problem. Whether it is at the scale of a campus, of a building or more simply of a room, we can combine several additional localization levels (and several technologies) to allow a more accurate and reliable user perception system. Within the PRIMA project, we are currently experimenting with a multi-level localization system allowing variable granularity according to the available equipment and the precision required for the targeted service.

SIROCCO Project-Team

4. Application Domains

4.1. Introduction

The application domains addressed by the project are:

- Compression with advanced functionalities of various image modalities (including multi-view, medical images such as MRI, CT, WSI, or satellite images)
- Networked multimedia applications via their various needs in terms of image and 2D and 3D video compression, or in terms of network adaptation (e.g., resilience to channel noise)
- Content editing and post-production

4.2. Compression with advanced functionalities

Compression of images and of 2D video (including High Definition and Ultra High Definition) remains a widely-sought capability for a large number of applications. The continuous increase of access network bandwidth leads to increasing numbers of networked digital content users and consumers which in turn triggers needs for higher core bandwidth and higher compression efficiencies. This is particularly true for mobile applications, as the need for wireless transmission capacity will significantly increase during the years to come. Hence, efficient compression tools are required to satisfy the trend towards mobile access to larger image resolutions and higher quality. A new impulse to research in video compression is also brought by the emergence of new formats beyond High Definition TV (HDTV) towards high dynamic range (higher bit depth, extended colorimetric space), super-resolution, formats for immersive displays allowing panoramic viewing and 3DTV.

Different video data formats and technologies are envisaged for interactive and immersive 3D video applications using omni-directional videos, stereoscopic or multi-view videos. The "omni-directional video" set-up refers to 360-degree view from one single viewpoint or spherical video. Stereoscopic video is composed of two-view videos, the right and left images of the scene which, when combined, can recreate the depth aspect of the scene. A multi-view video refers to multiple video sequences captured by multiple video cameras and possibly by depth cameras. Associated with a view synthesis method, a multi-view video allows the generation of virtual views of the scene from any viewpoint. This property can be used in a large diversity of applications, including Three-Dimensional TV (3DTV), and Free Viewpoint Video (FTV). The notion of "free viewpoint video" refers to the possibility for the user to choose an arbitrary viewpoint and/or view direction within a visual scene, creating an immersive environment. Multi-view video generates a huge amount of redundant data which need to be compressed for storage and transmission. In parallel, the advent of a variety of heterogeneous delivery infrastructures has given momentum to extensive work on optimizing the end-to-end delivery QoS (Quality of Service). This encompasses compression capability but also capability for adapting the compressed streams to varying network conditions. The scalability of the video content compressed representation, its robustness to transmission impairments, are thus important features for seamless adaptation to varying network conditions and to terminal capabilities.

In medical imaging, the large increase of medical analysis using various image sources for clinical purposes and the necessity to transmit or store these image data with improved performances related to transmission delay or storage capacities, command to develop new coding algorithms with lossless compression algorithms or *almost* lossless compression characteristics with respect to the medical diagnosis.

4.3. Networked visual applications

3D and Free Viewpoint TV: The emergence of multi-view auto-stereoscopic displays has spurred a recent interest for broadcast or Internet delivery of 3D video to the home. Multiview video, with the help of depth information on the scene, allows scene rendering on immersive stereo or auto-stereoscopic displays for 3DTV applications. It also allows visualizing the scene from any viewpoint, for scene navigation and free-viewpoint TV (FTV) applications. However, the large volumes of data associated to multi-view video plus depth content raise new challenges in terms of compression and communication.

Internet and mobile video: Broadband fixed (ADSL, ADSL2+) and mobile access networks with different radio access technologies (RAT) (e.g. 3G/4G, GERAN, UTRAN, DVB-H), have enabled not only IPTV and Internet TV but also the emergence of mobile TV and mobile devices with internet capability. A major challenge for next internet TV or internet video remains to be able to deliver the increasing variety of media (including more and more bandwidth demanding media) with a sufficient end-to-end QoS (Quality of Service) and QoE (Quality of Experience).

Mobile video retrieval: The Internet has changed the ways of interacting with content. The user is shifting its media consumption from a passive to a more interactive mode, from linear broadcast (TV) to on demand content (YouTubes, iTunes, VoD), and to user-generated, searching for relevant, personalized content. New mobility and ubiquitous usage has also emerged. The increased power of mobile devices is making content search and retrieval applications using mobile phones possible. Quick access to content in mobile environments with restricted bandwidth resources will benefit from rate-efficient feature extraction and description.

Wireless multi-camera vision systems: Our activities on scene modelling, on rate-efficient feature description, distributed coding and compressed sensing should also lead to algorithmic building blocks relevant for wireless multi-camera vision systems, for applications such as visual surveillance and security.

4.4. Medical Imaging (CT, MRI, Virtual Microscopy)

The use of medical imaging has greatly increased in recent years, especially with *magnetic resonance images (MRI) and computed tomography (CT)*. In the medical sector, lossless compression schemes are in general used to avoid any signal degradation which could mask a pathology and hence disturb the medical diagnosis. Nethertheless, some discussions are on-going to use near-lossless coding of medical images, coupled with a detection and segmentation of region-of interest (ROIs) guided by a modeling stage of the image sensor, a precise knowledge of the medical imaging modalities and by the diagnosis and expertise of practitioners. New application domains using these new approaches of telemedecine will surely increase in the future. The second aspect deals with the legal need of biomedical images storage. The legacy rules of such archives are changing and it could be interesting to propose adaptive compression strategies, i.e to explore reversible lossy-to-lossless coding algorithms and new storage modalities which use, in a first stage, the lossless representation and continuously introduce controlled lossy degradations for the next stages of archives. Finally, it seems promising to explore new representation and coding approaches for 3D biological tissue imaging captured by *3D virtual microscopy*. These fields of interest and scientific application domains commonly generate terabytes of data. Lossless schemes but also lossy approaches have to be explored and optimized, and interactive tools supporting scalable and interactive access to large-sized images such as these virtual microscopy slides need to be developed.

4.5. Editing and post-production

Video editing and post-production are critical aspects in the audio-visual production process. Increased ways of "consuming" video content also highlight the need for content repurposing as well as for higher interaction and editing capabilities. Content captured at very high resolutions may need to be repurposed in order to be adapted to the requirements of actual users, to the transmission channel or to the terminal. Content repurposing encompasses format conversion (retargeting), content summarization, and content editing. This processing requires powerful methods for extracting condensed video representations as well as powerful inpainting techniques. By providing advanced models, advanced video processing and image analysis tools,

more visual effects, with more realism become possible. Other applications such as video annotation/retrieval, video restoration/stabilization, augmented reality, can also benefit from the proposed research.

STARS Team

4. Application Domains

4.1. Introduction

While in our research the focus is to develop techniques, models and platforms that are generic and reusable, we also make effort in the development of real applications. The motivation is twofold. The first is to validate the new ideas and approaches we introduce. The second is to demonstrate how to build working systems for real applications of various domains based on the techniques and tools developed. Indeed, Stars focuses on two main domains: **video analytics** and **healthcare monitoring**.

4.2. Video Analytics

Our experience in video analytics [7], [1], [9] (also referred to as visual surveillance) is a strong basis which ensures both a precise view of the research topics to develop and a network of industrial partners ranging from end-users, integrators and software editors to provide data, objectives, evaluation and funding.

For instance, the Keeneo start-up was created in July 2005 for the industrialization and exploitation of Orion and Pulsar results in video analytics (VSIP library, which was a previous version of SUP). Keeneo has been bought by Digital Barriers in August 2011 and is now independent from Inria. However, Stars continues to maintain a close cooperation with Keeneo for impact analysis of VSIP and for exploitation of new results.

Moreover new challenges are arising from the visual surveillance community. For instance, people detection and tracking in a crowded environment are still open issues despite the high competition on these topics. Also detecting abnormal activities may require to discover rare events from very large video data bases often characterized by noise or incomplete data.

4.3. Healthcare Monitoring

We have initiated a new strategic partnership (called CobTek) with Nice hospital [62], [81] (CHU Nice, Prof P. Robert) to start ambitious research activities dedicated to healthcare monitoring and to assistive technologies. These new studies address the analysis of more complex spatio-temporal activities (e.g. complex interactions, long term activities).

To achieve this objective, several topics need to be tackled. These topics can be summarized within two points: finer activity description and longer analysis. Finer activity description is needed for instance, to discriminate the activities (e.g. sitting, walking, eating) of Alzheimer patients from the ones of healthy older people. It is essential to be able to pre-diagnose dementia and to provide a better and more specialised care. Longer analysis is required when people monitoring aims at measuring the evolution of patient behavioural disorders. Setting up such long experimentation with dementia people has never been tried before but is necessary to have real-world validation. This is one of the challenge of the European FP7 project Dem@Care where several patient homes should be monitored over several months.

For this domain, a goal for Stars is to allow people with dementia to continue living in a self-sufficient manner in their own homes or residential centers, away from a hospital, as well as to allow clinicians and caregivers remotely proffer effective care and management. For all this to become possible, comprehensive monitoring of the daily life of the person with dementia is deemed necessary, since caregivers and clinicians will need a comprehensive view of the person's daily activities, behavioural patterns, lifestyle, as well as changes in them, indicating the progression of their condition.

The development and ultimate use of novel assistive technologies by a vulnerable user group such as individuals with dementia, and the assessment methodologies planned by Stars are not free of ethical, or even legal concerns, even if many studies have shown how these Information and Communication Technologies (ICT) can be useful and well accepted by older people with or without impairments. Thus one goal of Stars team is to design the right technologies that can provide the appropriate information to the medical carers while preserving people privacy. Moreover, Stars will pay particular attention to ethical, acceptability, legal and privacy concerns that may arise, addressing them in a professional way following the corresponding established EU and national laws and regulations, especially when outside France.

As presented in 3.1 , Stars aims at designing cognitive vision systems with perceptual capabilities to monitor efficiently people activities. As a matter of fact, vision sensors can be seen as intrusive ones, even if no images are acquired or transmitted (only meta-data describing activities need to be collected). Therefore new communication paradigms and other sensors (e.g. accelerometers, RFID, and new sensors to come in the future) are also envisaged to provide the most appropriate services to the observed people, while preserving their privacy. To better understand ethical issues, Stars members are already involved in several ethical organizations. For instance, F. Bremond has been a member of the ODEGAM - "Commission Ethique et Droit" (a local association in Nice area for ethical issues related to older people) from 2010 to 2011 and a member of the French scientific council for the national seminar on "La maladie d'Alzheimer et les nouvelles technologies - Enjeux éthiques et questions de société" in 2011. This council has in particular proposed a chart and guidelines for conducting researches with dementia patients.

For addressing the acceptability issues, focus groups and HMI (Human Machine Interaction) experts, will be consulted on the most adequate range of mechanisms to interact and display information to older people.

<p style="text-align:center"><strong style="color:red">TEXMEX Project-Team</p>

4. Application Domains

4.1. Copyright protection of images and videos

With the proliferation of high-speed Internet access, piracy of multimedia data has developed into a major problem and media distributors, such as photo agencies, are making strong efforts to protect their digital property. Today, many photo agencies expose their collections on the web with a view to selling access to the images. They typically create web pages of thumbnails, from which it is possible to purchase high-resolution images that can be used for professional publications. Enforcing intellectual property rights and fighting against copyright violations is particularly important for these agencies, as these images are a key source of revenue. The most problematic cases, and the ones that induce the largest losses, occur when "pirates" steal the images that are available on the Web and then make money by illegally reselling those images.

This applies to photo agencies, and also to producers of videos and movies. Despite the poor image quality, thousands of (low-resolution) videos are uploaded every day to video-sharing sites such as YouTube, eDonkey or BitTorrent. In 2005, a study conducted by the Motion Picture Association of America was published, which estimated that their members lost 2,3 billion US$ in sales due to video piracy over the Internet. Due to the high risk of piracy, movie producers have tried many means to restrict illegal distribution of their material, albeit with very limited success.

Photo and video pirates have found many ways to circumvent even the protection mechanisms. In order to cover up their tracks, stolen photos are typically cropped, scaled, their colors are slightly modified; videos, once ripped, are typically compressed, modified and re-encoded, making them more suitable for easy downloading. Another very popular method for stealing videos is cam-cording, where pirates smuggle digital camcorders into a movie theater and record what is projected on the screen. Once back home, that goes to the web.

Clearly, this environment calls for an automatic content-based copyright enforcement system, for images, videos, and also audio as music gets heavily pirated. Such a system needs to be effective as it must cope with often severe attacks against the contents to protect, and efficient as it must rapidly spot the original contents from a huge reference collection.

4.2. Video database management

The existing video databases are generally little digitized. The progressive migration to digital television should quickly change this point. As a matter of fact, the French TV channel TF1 switched to an entirely digitized production, the cameras remaining the only analogical spot. Treatment, assembly and diffusion are digital. In addition, domestic digital decoders can, from now on, be equipped with hard disks allowing a storage initially modest, of ten hours of video, but larger in the long term, of a thousand of hours.

One can distinguish two types of digital files: private and professional files. On one hand, the files of private individuals include recordings of broadcasted programs and films recorded using digital camcorders. It is unlikely that users will rigorously manage such collections; thus, there is a need for tools to help the user: Automatic creation of summaries and synopses to allow finding information easily or to have within few minutes a general idea of a program. Even if the service is rustic, it is initially evaluated according to the added value brought to a system (video tape recorder, decoder), must remain not very expensive, but will benefit from a large diffusion.

On the other hand, these are professional files: TV channel archives, cineclubs, producers... These files are of a much larger size, but benefit from the attentive care of professionals of documentation and archiving. In this field, the systems can be much more expensive and are judged according to the profits of productivity and the assistance which they bring to archivists, journalists and users.

A crucial problem for many professionals is the need to produce documents in many formats for various terminals from the same raw material without multiplying the editing costs. The aim of such a *repurposing* is for example to produce a DVD, a web site or an alert service by mobile phone from a TV program at the minimum cost. The basic idea is to describe the documents in such a way that they can be easily manipulated and reconfigured easily.

4.3. Textual database management

Searching in large textual corpora has already been the topic of many researches. The current stakes are the management of very large volumes of data, the possibility to answer requests relating more on concepts than on simple inclusions of words in the texts, and the characterization of sets of texts.

We work on the exploitation of scientific bibliographical bases. The explosion of the number of scientific publications makes the retrieval of relevant data for a researcher a very difficult task. The generalization of document indexing in data banks did not solve the problem. The main difficulty is to choose the keywords, which will encircle a domain of interest. The statistical method used, the factorial analysis of correspondences, makes it possible to index the documents or a whole set of documents and to provide the list of the most discriminating keywords for these documents. The index validation is carried out by searching information in a database more general than the one used to build the index and by studying the retrieved documents. That in general makes it possible to still reduce the subset of words characterizing a field.

We also explore scientific documentary corpora to solve two different problems: to index the publications with the help of meta-keys and to identify the relevant publications in a large textual database. For that, we use factorial data analysis, which allows us to find the minimal sets of relevant words that we call meta-keys and to free the bibliographical search from the problems of noise and silence. The performances of factorial correspondence analysis are sharply greater than classic search by logical equation.

WILLOW Project-Team

4. Application Domains

4.1. Introduction

We believe that foundational modeling work should be grounded in applications. This includes (but is not restricted to) the following high-impact domains.

4.2. Quantitative image analysis in science and humanities

We plan to apply our 3D object and scene modeling and analysis technology to image-based modeling of human skeletons and artifacts in anthropology, and large-scale site indexing, modeling, and retrieval in archaeology and cultural heritage preservation. Most existing work in this domain concentrates on image-based rendering—that is, the synthesis of good-looking pictures of artifacts and digs. We plan to focus instead on quantitative applications. We are engaged in a project involving the archaeology laboratory at ENS and focusing on image-based artifact modeling and decorative pattern retrieval in Pompeii. This effort is part of the MSR-Inria project mentioned earlier and that will be discussed further later in this report. Application of our 3D reconstruction technology is now being explored in the field of cultural heritage and archeology by the start-up Iconem, founded by Y. Ubelmann, a Willow collaborator.

4.3. Video Annotation, Interpretation, and Retrieval

Both specific and category-level object and scene recognition can be used to annotate, augment, index, and retrieve video segments in the audiovisual domain. The Video Google system developed by Sivic and Zisserman (2005) for retrieving shots containing specific objects is an early success in that area. A sample application, suggested by discussions with Institut National de l'Audiovisuel (INA) staff, is to match set photographs with actual shots in film and video archives, despite the fact that detailed timetables and/or annotations are typically not available for either medium. Automatically annotating the shots is of course also relevant for archives that may record hundreds of thousands of hours of video. Some of these applications will be pursued in our MSR-Inria project, in which INA is one of our partners.