INRIA

Project-Team TEMICS

TraitEment, Modélisation d'Images et CommunicationS

Rennes - Bretagne-Atlantique

Theme : Vision, Perception and Multimedia Understanding

Activity Report

2010

Table of contents

1. Team

Research Scientists

Christine Guillemot [Team Leader, DR, Inria, HdR]
Teddy Furon [CR, INRIA]
Cédric Herzet [CR, INRIA, until Sept. 2010]
Claude Labit [DR, INRIA, 20%, HdR]
Aline Roumy [CR, INRIA]

Faculty Members

Jean-Jacques Fuchs [Professor, SPM, University of Rennes 1, HdR]
Olivier Le Meur [MdC, IFSIC, University of Rennes 1]
Simon Malinowski [ATER, until August 2010]

Technical Staff

Laurent Guillo [IR, CNRS]

PhD Students

Ana Charpentier [INRIA grant, MEDIEVALS project]
Thomas Colleu [CIFRE contract with France Télécom, until Oct. 2010]
Safa Cherigui [CIFRE contract with Technicolor]
Angélique Drémeau [INRIA grant, until Oct. 2010]
Mounira Ebdelli [INRIA grant, Since Dec. 2010]
Josselin Gauthier [Ministry of research grant]
Vincent Jantet [INRIA grant, Futurimages project]
Jonathan Taquet [Inria-Région ARED grant]
Velotiaray Toto-Zarasoa [Ministry of research grant]
Mehmet Turkan [INRIA grant, CORDIS]
Fuchun Xie [INRIA grant, CORDIS, until Oct. 2010]
Joaquin Zepeda [INRIA grant, ANR-ICOS-HD project, also with TexMex, until Oct. 2010]

Post-Doctoral Fellows

Fethi Smach [INRIA grant, ANR-ICOS-HD project, until July 2010]
François Lecellier [INRIA grant, ANR-ICOS-HD project, until August 2010]
Peter Meerwald [INRIA grant, ANR-MEDIEVALS project, since Sept. 2010]

Administrative Assistant

Huguette Béchu [TR, INRIA, shared with Fluminance and Serpico]

Others

Ronan Boitard [INRIA, since 15th Sept. 2009]
Yann Serrand [INRIA, until 14th Sept. 2010]
Tanguy Poirier [INRIA, since 15th Sept. 2009]
Gael Sourimant [INRIA grant, Fururimages project, until Feb. 2010]

2. Overall Objectives

2.1. Overall Objectives

The goal of the TEMICS project-team is the design and development of algorithms and practical solutions in the areas of analysis, modelling, coding, communication and watermarking of images and video signals.The TEMICS project-team activities are structured and organized around the following research directions :

- *3D modelling and representations of multi-view video sequences.*

The emergence of new video formats, allowing panoramic viewing, free viewpoint video (FTV), and Three-Dimensional TV (3DTV) on immersive displays are creating new scientific and technological problems in the area of video content modelling and representation. Omni-directional video, free viewpoint video and stereoscopic or multi-view video are formats envisaged for interactive and 3DTV. Omni-directional video refers to a 360-degree view from one single viewpoint or a spherical video. 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. A multi-view video together with depth information allows, by using view synthesis techniques, the generation of virtual views of the scene from any viewpoint. This property can be used in a large diversity of applications, including 3DTV, FTV, security monitoring and tracking. This type of 3D content representation is also known as MVD (Multi-View plus Depth). The TEMICS project-team focuses on several algorithmic problems to analyze, represent, compress and render multi-view video content. The team first addresses the problem of depth information extraction. The depth information is associated with each view as a depth map, and transmitted in order to perform virtual view generation and allow the inter-operability between capture (with N cameras) and display (of P views) devices. The huge amount of data contained in multi-view sequences motivates the design of efficient representation and compression algorithms.

- *Sparse representations, compression, feature extraction and texture description.*

  Low rate as well as scalable compression remains a widely sought capability. Scalable video compression is essential to allow for optimal adaptation of compressed video streams to varying network characteristics (e.g. to bandwidth variations) as well as to heterogeneous terminal capabilities. Wavelet-based signal representations are well suited for such scalable signal representations. Special effort is thus dedicated to the study of motion-compensated spatio-temporal expansions making use of complete or overcomplete transforms, e.g. wavelets, curvelets and contourlets, and more genrally of sparse signal approximation and representation techniques. The sparsity of the signal representation depends on how well the bases match with the local signal characteristics. Anisotropic waveforms bases, based on directional transforms or on sets of bases optimized in a sparsity-distortion sense are studied. Methods for texture analysis and synthesis, for prediction and for inpainting, which are key components of image and video compression algorithms, based on sparse signal representations are also developed. The amenability of these representations for image texture description is also investigated and measures of distance between sparse vectors are designed for approximate nearest neighbors search and for image retrieval. Beyond sparse image and video signals representations, the problem of quantization of the resulting representations taking into account perceptual models and measures, in order to optimize a trade-off between rate and perceptual quality, is studied.

- *Joint source-channel coding.* The advent of Internet and wireless communications, often characterized by narrow-band, error and/or loss prone, heterogeneous and time-varying channels, is creating challenging problems in the area of source and channel coding. Design principles prevailing so far and stemming from Shannon's source and channel separation theorem must be re-considered. The separation theorem holds only under asymptotic conditions where both codes are allowed infinite length and complexity. If the design of the system is heavily constrained in terms of complexity or delay, source and channel coders, designed in isolation, can be largely suboptimal. The project objective is to develop theoretical and practical solutions for image and video transmission over heterogeneous, time-varying wired and wireless networks. Many of the theoretical challenges are related to understanding the tradeoffs between rate-distortion performance, delay and complexity for the code design. The issues addressed encompass the design of error-resilient source codes, joint source-channel codes and multiply descriptive codes, minimizing the impact of channel noise (packet losses, bit errors) on the quality of the reconstructed signal, as well as of turbo or iterative decoding techniques.

- *Distributed source and joint source-channel coding.* Current compression systems exploit correlation on the sender side, via the encoder, e.g. making use of motion-compensated predictive or filtering techniques. This results in asymmetric systems with respectively higher encoder and lower decoder

complexities suitable for applications such as digital TV, or retrieval from servers with e.g. mobile devices. However, there are numerous applications such as multi-sensors, multi-camera vision systems, surveillance systems, with light-weight and low power consumption requirements that would benefit from the dual model where correlated signals are coded separately and decoded jointly. This model, at the origin of distributed source coding, finds its foundations in the Slepian-Wolf and the Wyner-Ziv theorems. Even though first theoretical foundations date back to early 70's, it is only recently that concrete solutions have been introduced. In this context, the TEMICS project-team is working on the design of distributed prediction and coding strategies based on both source and channel codes. Although the problem is posed as a communication problem, classical channel decoders need to be modified. Distributed joint source-channel coding refers to the problem of sending correlated sources over a common noisy channel without communication between the senders. This problem occurs mostly in networks, where the communication between the nodes is not possible or not desired due to its high energy cost (network video camera, sensor network...). For independent channels, source channel separation holds but for interfering channels, joint source-channel schemes (but still distributed) performs better than the separated scheme. In this area, we work on the design of distributed source-channel schemes.

- *Data hiding and watermarking.*

  The distribution and availability of digital multimedia documents on open environments, such as the Internet, has raised challenging issues regarding ownership, users rights and piracy. With digital technologies, the copying and redistribution of digital data has become trivial and fast, whereas the tracing of illegal distribution is difficult. Consequently, content providers are increasingly reluctant to offer their multimedia content without a minimum level of protection against piracy. The problem of data hiding has thus gained considerable attention in the recent years as a potential solution for a wide range of applications encompassing copyright protection, authentication, steganography, or a mean to trace illegal usage of the content. This latter application is referred to as fingerprinting. Depending on the application (copyright protection, traitor tracing or fingerprinting, hidden communication), the embedded signal may need to be robust or fragile, more or less imperceptible. One may need to only detect the presence of a mark (watermark detection) or to extract a message. The message may be unique for a given content or different for the different users of the content, etc. These different applications place various constraints in terms of capacity, robustness and security on the data hiding and watermarking algorithms. The robust watermarking problem can be formalized as a communication problem : the aim is to embed a given amount of information in a host signal, under a fixed distortion constraint between the original and the watermarked signal, while at the same time allowing reliable recovery of the embedded information subject to a fixed attack distortion. Applications such as copy protection, copyright enforcement, or steganography also require a security analysis of the privacy of this communication channel hidden in the host signal.

Given the strong impact of standardization in the sector of networked multimedia, TEMICS, in partnership with industrial companies, seeks to promote its results in standardization (JPEG, MPEG). While aiming at generic approaches, some of the solutions developed are applied to practical problems in partnership with industry (Thomson, France Télécom) or in the framework of national projects (RIAM ESTIVALE, ANR ESSOR, ANR ICOS-HD, ANR MEDIEVALS, ANR PERSEE, DGE/REGION FUTURIMAGES) and European projects (IST-NEWCOM++). The application domains addressed by the project are networked multimedia applications (on wired or wireless Internet) via their various requirements and needs in terms of compression, of resilience to channel noise, or of advanced functionalities such as navigation, protection and authentication.

3. Scientific Foundations

3.1. 3D scene modelling based on projective geometry

3D reconstruction is the process of estimating the shape and position of 3D objects from views of these objects. TEMICS deals more specifically with the modelling of large scenes from monocular video sequences. 3D reconstruction using projective geometry is by definition an inverse problem. Some key issues which do not have yet satisfactory solutions are the estimation of camera parameters, especially in the case of a moving camera. Specific problems to be addressed are e.g. the matching of features between images, and the modelling of hidden areas and depth discontinuities. 3D reconstruction uses theory and methods from the areas of computer vision and projective geometry. When the camera $\mathcal{C}_i$ is modelled as a *perspective projection*, the *projection equations* are :

$$\widetilde{p}_i = P_i \widetilde{x}, \tag{1}$$

where $\widetilde{x}$ is a 3D point with homogeneous coordinates $\widetilde{x} = (x \ y \ z \ 1)^t$ in the scene reference frame $\mathcal{R}_0$, and where $\widetilde{p}_i = (X_i \ Y_i \ 1)^t$ are the coordinates of its projection on the image plane I_i. The *projection matrix* P_i associated to the camera $\mathcal{C}_i$ is defined as $P_i = K(R_i|t_i)$. It is function of both the *intrinsic parameters* K of the camera, and of transformations (rotation R_i and translation t_i) called the *extrinsic parameters* and characterizing the position of the camera reference frame $\mathcal{R}_i$ with respect to the scene reference frame $\mathcal{R}_0$. Intrinsic and extrinsic parameters are obtained through calibration or self-calibration procedures. The *calibration* is the estimation of camera parameters using a calibration pattern (objects providing known 3D points), and images of this calibration pattern. The *self-calibration* is the estimation of camera parameters using only image data. These data must have previously been matched by identifying and grouping all the image 2D points resulting from projections of the same 3D point. Solving the 3D reconstruction problem is then equivalent to searching for $\widetilde{x}$, given $\widetilde{p}_i$, i.e. to solve Eqn. (1) with respect to coordinates $\widetilde{x}$. Like any inverse problem, 3D reconstruction is very sensitive to uncertainty. Its resolution requires a good accuracy for the image measurements, and the choice of adapted numerical optimization techniques.

3.2. Frame expansions

Signal representation using orthogonal basis functions (e.g., DCT, wavelet transforms) is at the heart of source coding. The key to signal compression lies in selecting a set of basis functions that compacts the signal energy over a few coefficients. Frames are generalizations of a basis for an overcomplete system, or in other words, frames represent sets of vectors that span a Hilbert space but contain more numbers of vectors than a basis. Therefore signal representations using frames are known as overcomplete frame expansions. Because of their inbuilt redundancies, such representations can be useful for providing robustness to signal transmission over error-prone communication media. Consider a signal $\mathbf{x}$. An overcomplete frame expansion of $\mathbf{x}$ can be given as $F\mathbf{x}$ where F is the frame operator associated with a frame $\Phi_F \equiv \{\varphi_i\}_{i \in I}$, φ's are the frame vectors and I is the index set. The ith frame expansion coefficient of $\mathbf{x}$ is defined as $(F\mathbf{x})_i \equiv \langle \varphi_i, \mathbf{x} \rangle$, for all $i \in I$. Given the frame expansion of $\mathbf{x}$, it can be reconstructed using the dual frame of Φ_F which is given as $\widetilde{\Phi}_F \equiv \{(F^h F)^{-1} \varphi_i\}_{i \in I}$. Tight frame expansions, where the frames are self-dual, are analogous to orthogonal expansions with basis functions. Frames in finite-dimensional Hilbert spaces such as $\mathbf{R}^K$ and $\mathbf{C}^K$, known as discrete frames, can be used to expand signal vectors of finite lengths. In this case, the frame operators can be looked upon as redundant block transforms whose rows are conjugate transposes of frame vectors. For a K-dimensional vector space, any set of N, $N > K$, vectors that spans the space constitutes a frame. Discrete tight frames can be obtained from existing orthogonal transforms such as DFT, DCT, DST, etc by selecting a subset of columns from the respective transform matrices. Oversampled filter banks can provide frame expansions in the Hilbert space of square summable sequences, i.e., $l_2(\mathbf{Z})$. In this case, the time-reversed and shifted versions of the impulse responses of the analysis and synthesis filter banks constitute the frame and its dual. Since overcomplete frame expansions provide redundant information, they can be used as joint source-channel codes to fight against channel degradations. In this context, the recovery of a message signal from the corrupted frame expansion coefficients can be linked to the error correction in infinite fields. For example, for discrete frame expansions, the frame operator can be looked upon as the generator matrix of a block code in the real or complex field. A parity check matrix for this code can be obtained from the singular

value decomposition of the frame operator, and therefore the standard syndrome decoding algorithms can be utilized to correct coefficient errors. The structure of the parity check matrix, for example the BCH structure, can be used to characterize discrete frames. In the case of oversampled filter banks, the frame expansions can be looked upon as convolutional codes.

3.3. Rate-distortion theory

Coding and joint source channel coding rely on fundamental concepts of information theory, such as notions of entropy, memoryless or correlated sources, of channel capacity, or on rate-distortion performance bounds. Compression algorithms are defined to be as close as possible to the optimal rate-distortion bound, $R(D)$, for a given signal. The source coding theorem establishes performance bounds for lossless and lossy coding. In lossless coding, the lower rate bound is given by the entropy of the source. In lossy coding, the bound is given by the rate-distortion function $R(D)$. This function $R(D)$ gives the minimum quantity of information needed to represent a given signal under the constraint of a given distortion. The rate-distortion bound is usually called OPTA (*Optimum Performance Theoretically Attainable*). It is usually difficult to find close-form expressions for the function $R(D)$, except for specific cases such as Gaussian sources. For real signals, this function is defined as the convex-hull of all feasible (rate, distortion) points. The problem of finding the rate-distortion function on this convex hull then becomes a rate-distortion minimization problem which, by using a Lagrangian formulation, can be expressed as

$$\frac{\partial J}{\partial Q} = 0 \quad \text{where} \quad J = R + \lambda D \quad \text{with} \quad \lambda > 0.$$

The Lagrangian cost function J is derivated with respect to the different optimisation parameters, e.g. with respect to coding parameters such as quantization factors. The parameter λ is then tuned in order to find the targeted rate-distortion point. When the problem is to optimise the end-to-end Quality of Service (QoS) of a communication system, the rate-distortion metrics must in addition take into account channel properties and channel coding. Joint source-channel coding optimisation allows to improve the tradeoff between compression efficiency and robustness to channel noise.

3.4. Distributed source coding

Distributed source coding (DSC) has emerged as an enabling technology for sensor networks. It refers to the compression of correlated signals captured by different sensors which do not communicate between themselves. All the signals captured are compressed independently and transmitted to a central base station which has the capability to decode them jointly. DSC finds its foundation in the seminal Slepian-Wolf (SW) and Wyner-Ziv (WZ) theorems. Let us consider two binary correlated sources X and Y. If the two coders communicate, it is well known from Shannon's theory that the minimum lossless rate for X and Y is given by the joint entropy $H(X,Y)$. Slepian and Wolf have established in 1973 that this lossless compression rate bound can be approached with a vanishing error probability for long sequences, even if the two sources are coded separately, provided that they are decoded jointly and that their correlation is known to both the encoder and the decoder. The achievable rate region is thus defined by $R_X \geq H(X|Y)$, $R_Y \geq H(Y|X)$ and $R_X + R_Y \geq H(X,Y)$, where $H(X|Y)$ and $H(Y|X)$ denote the conditional entropies between the two sources.

In 1976, Wyner and Ziv considered the problem of coding of two correlated sources X and Y, with respect to a fidelity criterion. They have established the rate-distortion function $R *_{X|Y}(D)$ for the case where the side information Y is perfectly known to the decoder only. For a given target distortion D, $R *_{X|Y}(D)$ in general verifies $R_{X|Y}(D) \leq R *_{X|Y}(D) \leq R_X(D)$, where $R_{X|Y}(D)$ is the rate required to encode X if Y is available to both the encoder and the decoder, and R_X is the minimal rate for encoding X without SI. Wyner and Ziv have shown that, for correlated Gaussian sources and a mean square error distortion measure, there is no rate loss with respect to joint coding and joint decoding of the two sources, i.e., $R *_{X|Y}(D) = R_{X|Y}(D)$.

3.5. Watermarking as a problem of communication with side information

Digital watermarking aims at hiding discrete messages into multimedia content. The watermark must not spoil the regular use of the content, i.e., the watermark should be non perceptible. Hence, the embedding is usually done in a transformed domain where a human perception model is exploited to assess the non perceptibility criterion. The watermarking problem can be regarded as a problem of creating a communication channel within the content. This channel must be secure and robust to usual content manipulations like lossy compression, filtering, geometrical transformations for images and video. When designing a watermarking system, the first issue to be addressed is the choice of the transform domain, i.e., the choice of the signal components that will *host* the watermark data. An extraction function $E(.)$ going from the content space $\mathcal{C}$ to the components space, isomorphic to $\mathbf{R}^N$, must then first be defined.

The embedding process actually transforms a host vector $\mathbf{V}$ into a watermarked vector $\mathbf{V_w}$. The perceptual impact of the watermark embedding in this domain must be quantified and constrained to remain below a certain level. The measure of perceptual distortion is usually defined as a cost function $d(\mathbf{V_w} - \mathbf{V})$ in $\mathbf{R}^N$ constrained to be lower than a given distortion bound d_w. Attack noise will be added to the watermark vector. In order to evaluate the robustness of the watermarking system and design counter-attack strategies, the noise induced by the different types of attack (e.g. compression, filtering, geometrical transformations, ...) must be modelled. The distortion induced by the attack must also remain below a distortion bound $d(\mathbf{V_a} - \mathbf{V}) < d_a$. Beyond this distortion bound, the content is considered to be non usable any more. Watermark detection and extraction techniques will then exploit the knowledge of the statistical distribution of the vectors $\mathbf{V}$. Given the above mathematical model, one has then to design a suitable communication scheme. Direct sequence spread spectrum techniques are often used. The chip rate sets the trade-off between robustness and capacity for a given embedding distortion. This can be seen as a labelling process $S(.)$ mapping a discrete message $m \in \mathcal{M}$ onto a signal in $\mathbf{R}^N$.

The decoding function $S^{-1}(.)$ is then applied to the received signal $\mathbf{V_a}$ in which the watermark interferes with two sources of noise: the original host signal ($\mathbf{V}$) and the attack ($\mathbf{A}$). The problem is then to find the pair of functions $\{S(.), S^{-1}(.)\}$ that will allow to optimise the communication channel under the distortion constraints $\{d_t, d_a\}$. This amounts to maximizing the probability to decode correctly the hidden message:

$$\max \mathrm{Prob}[S^{-1}(S(m) + \mathbf{V} + \mathbf{A}) = m] \quad \text{under constraints } d_t, d_a$$

A new paradigm stating that the original host signal $\mathbf{V}$ shall be considered as a *channel state* only known at the embedding side rather than a source of noise, appeared recently. The watermark signal thus depends on the channel state: $\mathbf{S} = S(m, \mathbf{V})$. This new paradigm known as communication with side information, sets the theoretic foundations for the design of new communication schemes with increased capacity.

3.6. Traitor tracing as a compressed sensing problem over binary data

Multimedia security witnesses an increasing interest on traitor tracing, also known as active fingerprinting, aiming at pinpointing the origin of the leak within a distribution framework. A server distributes a copy of a video to n users. A pirated copy emerges on an illegal P2P file sharing network. The issue is to identify the dishonest users even if a collusion, ie. a group of c attackers $\{j_1, ..., j_c\}$, merged their individual copies to forge this pirated copy.

Traitor tracing uses a watermarking technique to embed a unique codeword $\mathbf{x}_j$ into the individual copy given to user j. The collusion process mixes the codewords of the colluders into a pirated sequence $\mathbf{y} = f(\mathbf{x}_{j_1}, ..., \mathbf{x}_{j_c})$. Fingerprinting is the art of designing a binary code $\mathbf{X} = [\mathbf{x}_1, ..., \mathbf{x}_n]$, ie. set of unique binary codewords, and a tracing algorithm such that it can identify the codewords that have been used to forge the pirated sequence. The utmost requirement is the probability of false positive ϵ, ie. the probability of accusing an innocent user. A criterion of comparison is the length of the code m. It has been proven that the shortest length is asymptotically $m = O(c^2 \log(n\epsilon^{-1}))$. G. Tardos was the first to exhibit a practical construction of such an optimal code.

This problem can be seen as compressed sensing over binary data. Define $\mathbf{a}$ the indicator vector where $a(j) = 1$ if user j is a colluder, 0 otherwise. A possible collusion strategy is to insert symbol '1' whenever possible (ie. when at least one colluder has such a symbol), which reads as $\mathbf{y} = \mathbf{X} \otimes \mathbf{a}$ ($\otimes$ defines the matrices multiplication over the binary field). In other words, $\mathbf{X}$ is the dictionary of 'atoms', and the accusation process aims at recovering the active 'atoms' coded by $\mathbf{a}$. This makes the connection with compressed sensing explicit. However, traitor tracing is somehow a more involved issue than compressed sensing since the collusion is free to choose the process $f(.)$ to forge $\mathbf{y}$ from $\mathbf{x}_{j_1}, ..., \mathbf{x}_{j_c}$.

4. Application Domains

4.1. Introduction

The application domains addressed by the project are networked multimedia applications via their various needs in terms of image and 2D and 3D video compression, network adaptation (e.g., resilience to channel noise), or in terms of advanced functionalities such as navigation, copy and copyright protection, or tracing of illegal content usage.

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 multti-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. Multimedia communication

Networked multimedia is expected to play a key role in the development of 3G and beyond 3G (i.e. all IP-based) networks, by leveraging higher bandwidth, IP-based ubiquitous service provisioning across heterogeneous infrastructures, and capabilities of rich-featured terminal devices. However, networked multimedia presents a number of challenges beyond existing networking and source coding capabilities. Among the problems to be addressed is the transmission of large quantities of information with delay constraints on heterogeneous, time-varying communication environments with non-guaranteed quality of service (QoS). It is now a common understanding that QoS provisioning for multimedia applications such as video or audio does require a loosening and a re-thinking of the end-to-end and layer separation principle. In that context, the joint source-channel coding and the cross-layer paradigms set the foundations for the design of efficient solutions to the above challenges.

In parallel, emerging multimedia communication applications such as wireless video (e.g. mobile cameras), multi-sensors, multi-camera vision systems, surveillance systems are placing additionnal constraints on compression solutions, such as limited power consumption due to limited handheld battery power. The traditional balance of complex encoder and simple decoder may need to be reversed for these particular applications. In addition, wireless camera sensors capture and need to large volume of redundant data wighout information exhange between the sensors. The redundancy and correlation between the captured data can then only be removed on the receiving end. Distributed source coding is a recent research area which aims at addressing these needs.

4.4. Copy protection, copyright enforcement and traitor tracing

Data hiding has gained attention as a potential solution for a wide range of applications placing various constraints on the design of watermarking schemes in terms of embedding rate, robustness, invisibility, security, complexity. Here are two examples to illustrate this diversity. In copy protection, the watermark is just a flag warning compliant consumer electronic devices that a pirated piece of content is indeed a copyrighted content whose cryptographic protection has been broken. The priorities are a high invisibility, an excellent robustness, and a very low complexity at the watermark detector side. The security level must be fair, and the payload is reduced to its minimum (this is known as zero-bit watermarking scheme). In the fingerprinting (or traitor tracing) application, user identifying codes are embedded in the host signal to dissuade dishonest users to illegally give away the copyrighted contents they bought. The embedded data must be non perceptible not to spoil the entertainment of the content, and robust to a collusion attack where several dishonest users mix their copies in order to forge an untraceable content. This application requires a high embedding rate as anti-collusion codes are very long and a great robustness, however embedding and decoding can be done off-line affording for huge complexity.

5. Software

5.1. New software and former software still in use or evolving

5.1.1. Libit

Participant: Laurent Guillo [contact person].

Libit is a C library initially developed by Vivien Chappelier and Hervé Hégou former Ph.D students in the TEMICS project-team. It extends the C language with vector, matrix, complex and function types, and provides some common source coding, channel coding and signal processing tools. The goal of libit is to provide easy to use yet efficient tools commonly used tools to build a communication chain, from signal processing and source coding to channel coding and transmission. It is mainly targeted at researchers and developers in the fields of compression and communication. The syntax is purposely close to that of other tools commonly used in these fields, such as MATLAB, octave, or IT++. Therefore, experiments and applications can be developped, ported and modified simply. As examples and to ensure the correctness of the algorithms with respect to published results, some test programs are also provided. (http://libit.sourceforge.net). The library is made available under the GNU Library General Public Licence and is widely used in the software developments of the project-team.

5.1.2. *Oriented wavelet based image codec*

Participant: Christine Guillemot [contact person].

This still image codec is based on oriented wavelet transforms developed in the team. The transform is based on wavelet lifting locally oriented according to multiresolution image geometry information. The lifting steps of a 1D wavelet are applied along a discrete set of local orientations defined on a quincunx sampling grid. To maximize energy compaction, the orientation minimizing the prediction error is chosen adaptively. This image codec outperforms JPEG-2000 for lossy compression. This software has been registered at the APP (Agence de Protection des Programmes) under the number IDDN.FR.001.260024.000.S.P.2008.000.21000. The possibility to extract image descriptors in the transform domain making use of the orientation maps inherent to the coding algorithm has been studied in 2010.

5.1.3. *Distributed video codec*

Participants: Aline Roumy, Velotiaray Toto-Zarasoa [contact person].

A distributed video coding software has been developed within the DISCOVER European research project (http://www.discoverdvc.org) with contributions of the TEMICS project-team. In particular, the TEMICS project-team has contributed to the following modules of the DISCOVER codec: side information extraction with the mesh-based and hybrid block/mesh based motion estimation techniques, Rate control, Optimal MMSE reconstruction at the decoder, and on the exploitation of the source memory with a distributed DPCM solution. The DISCOVER codec is one of the most (or even the most) efficient distributed video coding solutions nowadays. Its executable files, along with sample configuration can be downloaded from http://www.discoverdvc.org/. The results of a comprehensive performance evaluation of the DISCOVER codec can be found on the web-page http://www.img.lx.it.pt/~discover/home.html. In 2010, the codec has evolved by enhancing the Slepian-Wolf decoder in order to account for the memory present in the input sources. A two-state Gilbert-Elliott finite state process is used for modelling the sources and a novel correlation model has been introduced. The model parameters are estimated with an Expectation-Maximization algorithm and the LDPC decoder has been modified in order to take into account the Gilbert-Eliott source model. The Distributed Video Coding system that integrates the enhancement that we propose here demonstrates a quality-versus-rate improvement by up to 10.14%, with respect to its elder version. A simplified model with non-uniform sources is also proposed that achieves an improvement by up to 5.7% (see Section 6.3.1).

5.1.4. *M3DPlayer: 3D video player*

Participant: Vincent Jantet [contact person].

A 3D player - named M3DPlayer - supporting rendering of a 3D scene and navigation within the scene has been developed. It integrates as a plug-in the 3D model-based video codec of the team. From a video sequence of a static scene viewed by a monocular moving camera, the 3D model-based video codec allows the automatic construction of a representation of a video sequence as a stream of textured 3D models. 3D models are extracted using stereovision and dense matching maps estimation techniques. A virtual sequence is reconstructed by projecting the textured 3D models on image planes. This representation enables 3D functionalities such as synthetic objects insertion, lightning modification, stereoscopic visualization or interactive navigation. The codec allows compression at very low bit-rates (16 to 256 kb/s in 25Hz CIF format) with a satisfactory visual quality. It also supports scalable coding of both geometry and texture information. The first version of the software was registered at the Agency for the Protection of Programmes (APP) under the number IDDN.FR.001.130017.000S.P.2003.000.41200. A second version of the player has been registered at the APP (Agence de Protection des Programmes) under the number IDDN.FR.001.090023.000.S.P.2008.000.21000. In 2009-2010, we focused on improving the rendering engine, based on recent OpenGL extensions, to be able to render the viewed scenes on an auto-stereoscopic display with low-end graphic cards. In our case, auto-stereoscopic display requires the rendering of eight 1920x1200 frames instead of just one for a standard display. This player is also used to render LDI (Layered Depth Images) and LDV (Layered Depth Videos) and to visualize 3D scenes on autostereoscopic displays taking multiple input views rendered from the LDI representation.

5.1.5. *Depth maps extractor*

Participant: Josselin Gauthier [contact person].

This software aims at estimating depth maps from multi-view videos, to provide Multi-View plus Depth (MVD) videos. MVD videos can be used to synthesize virtual views of the scene, or to render a new multi-view video with a different number of views than the original video, for instance in an auto-stereoscopic display setup. This software has been developed in the context of the DGE/Region research project Futurim@ges. This software has been compared to the Depth Estimation Reference Software from the MPEG 3DV group. The depth maps prouced are much smoother and the depth map extraction is 75 times faster. Figure 1 shows depth maps extracted by the DERS software on the left and our depth maps extractor on the right.

(a)

Figure 1. Depth maps extracted for the kendo sequence with DERS (on the left) and TEMICS software (on the right.)

5.1.6. *WSVC: Wavelet-based Scalable Video Codec*

Participants: Ronan Boitard, Christine Guillemot, Laurent Guillo [contact person], Aline Roumy.

WSVC is a wavelet based video codec based on a motion compensated t+2D wavelet analysis. It is about to be registered at the Agency for the Protection of Programmes (APP). The codec supports three forms of scalability: temporal via motion-compensated temporal wavelet transforms, spatial scalability enabled by a spatial wavelet transform and SNR scalability enabled by a bit-plane encoding technique. A so-called /extractor/ allows the extraction of a portion of the bitstream to suit a particular receiver temporal and spatial resolution or the network bandwidth. This codec has been provided to Alcatel-Lucent and will be used in the context of the collaboration with Alcatel-Lucent (see Section 7.1.3) but also in the just starting ANR-ARSSO project (see Section 7.2.9). The software has been registered at the APP.

5.1.7. *Watermarking platform*

Participant: Teddy Furon [contact person].

This software platform aims at integrating as plug-ins a set of functions, watermark detection and robust watermarking embedding and extraction, for different applications (fingerprinting and copyright protection). These plug-ins include the Broken Arrows software and the Chimark2 software. The Broken Arrows software has been developed in collaboration with the CNRS-Gipsa-lab in Grenoble, in the context of the international challenge BOWS-2 (Break Our Watermarking System - 2nd Edition). The source code has been registered at the Agency for the Protection of Programmes (APP) under the number IDDN.FR.001.170012.000.S.P.2008.000.41100. The software is available as an open-source code distributed under the INRIA-CNRS license CECILL (see http://bows2.gipsa-lab.inpg.fr). The Chimark2 software developed in the context of the RNRT-Diphonet project is a robust image watermarking tool and registered at the Agency for the Protection of Programmes (APP) under the number

IDDN.FR.001.480027.001.S.A.2002.000.41100. The robust watermarking tool has then be extended in the context of the IST-Busman project for marking video signals. The platform also aims at supporting different types of attacks (collusion, compression, crop, de-synchronization, ...) and at demonstrating the performance of the watermarking tools in presence of these attacks. The platform is built with the CMake cross platform "make", and we can provide autonomous installers to install the platform on new computers. To enable the portability under Linux OS and Mac OS, new OpenCV functionnalities have been developed, based on the use of ffmpeg. A set of benchmarking scripts have been written in Python, to provide a benchmarking platform for today and future watermarking techniques. Our software now offers a multi-tab player which is able to open AVI, CIF and QCIF movies, but also many image formats as bmp, pgm, or jpeg.

5.2. ADT PICOVIN

Participants: Ronan Boitard, Laurent Guillo [contact person], Tanguy Poirier, Yann Serrand.

The ADT Picovin is a technological development action, which works closely with the project-team TEMICS. This is a development structure which gives its support to the project-team to integrate new and relevant algorithms into the state-of-the-art codec and to take part in standardization.

In January 2010 in Kyoto, the ISO/MPEG and ITU-T/VCEG groups issued a join call for proposals about a future video codec that would be a possible successor of H264. Responses were revealed in April 2010 in Dresden. Rate gains over 40% have been achieved for a quality equivalent to H.264. As there was not a solution definitely better than the others, a temporary solution, the Test Model under Consideration (TMuC), emerged step by step in gathering the best tools that will be evaluated.

In 2010, the ADT mainly focused on the development and integration of algorithms dedicated to intra and inter prediction. A part of our work was submitted and presented as a proposal in Geneva in July 2010. It dealt with a new intra method prediction based on a linear combination of template matching predictors. It performs well when integrated in KTA2.7 and we were encouraged to further study its behaviour once integrated in the TMuC, as all our new tools will be from now.

Since July, the ADT also took part in tool experiments (TE) which aims at evaluating tools gathered in the TMuC. We ran conjointly with Technicolor Rennes tests mostly on asymmetric motion partitioning. We presented our results in October 2010 in Guangzhou.

This new development structure started in October 2008 and will last three years. During most of this year, three junior engineers and one permanent engineer from the SED Rennes (development and experimentation department of INRIA Rennes) take part to the ADT. It is supported by the technological development department of INRIA.

6. New Results

6.1. 3D representation and compression of video sequences

3DTV and Free Viewpoint Video (FVV) are now emerging video formats expected to offer an anhanced user experience. The 3D experience consists either in 3D relief rendering called 3DTV, or interactive navigation inside the content called FTV (Free viewpoint TV). 3D information can be easily computed for synthetic movies. However, so far, no professional nor consumer video cameras capable of capturing the 3D structure of the scene are available (except of course for Z-cameras prototypes). As a consequence, 3D information representing real content have to be estimated from acquired videos, using computer vision-based algorithms. This is the scope of the first research axis described below which focuses on depth maps extraction. Once the depth information has been extracted, the resulting videos with the associated depth maps must be coded to be stored or transmitted to the rendering device. 3D representations of the scene as well as associated compression techniques must be developed. The choice of the representation of the 3D is of central importance. On one hand, it sets the requirements for acquisition and signal processing. On the other hand, it determines the rendering algorithms, degree and mode of interactivity, as well as the need and means for compression and transmission. This is the scope of the next two research axis below.

6.1.1. *3d representation for multi-view video sequences based on a soup of polygons*
Participants: Thomas Colleu, Claude Labit.

This study is carried out in collaboration with INSA-Rennes (Luce Morin) and with France Telecom under a CIFRE contract. The aim is to study a new data representation for multi-view sequences. The data representation must allow real-time rendering of good quality images for different viewpoints, possibly virtual (ie non acquired) viewpoints. Moreover, the representation should be compact and efficiently compressed so as to limit the data overload compared with encoding each of the N video sequences using traditional 2D video codec. A new representation that takes as an input multi-view video plus depth (MVD), and results in a polygon soup has been developed. A polygon soup is a set of polygons that are not necessarily connected to each other. Each polygon is defined with image plus depth data and rendered as a 3D primitive thanks to a graphics processing unit (GPU). The polygons are actually quadrilaterals (aka quads) and they are extracted with a quadTree decomposition of the depth maps. The advantages of using polygonal primitives instead of point primitives traditionally used in other representations were shown. Also many redundancies across the viewpoints were reduced so as to obtain a compact representation.

In 2010, in order to evaluate this representation, a new compression method has been developed and view synthesis methods have been implemented. The image quality and bitrate have been compared with another existing method based on a multi-view plus depth (MVD) data representation. The results have shown that this representation provides slightly higher image quality at medium and high bit-rates. The results have also shown that, in both methods, some artifacts appear in the form of texture misalignments. They are due to geometry errors or inconsistencies across the views. A new method for reducing texture misalignments has been studied and evaluated. This method is called 'floating geometry'. It consists in deforming the geometry of the representation depending on the desired view point such that texture misalignments are reduced. In order to compute this deformation, views are synthesized at original viewpoints and compared with original images. Then 3D deformation is guided by the 2D motion estimation between synthesized and original views. Finally, for synthesizing virtual viewpoints, then the geometry deformation is interpolated between 3D positions already computed. This floating geometry method has been evaluated with the polygon soup representation mentioned above. The results have shown that some typical artifacts can be reduced. The results have shown that a gain in PSNR of up to 1.5 dB could be obtained with this method.

6.1.2. *3d representation for multi-view video sequences based LDI representations*
Participants: Christine Guillemot, Vincent Jantet.

This study is carried out in collaboration with INSA/IETR (Luce Morin). A multi-view video is a collection of video sequences captured for the same scene, synchronously by many cameras at different locations. 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), Free Viewpoint Video (FTV), security monitoring, tracking and 3D reconstruction. The huge amount of data contained in a multi-view sequence motivates the design of efficient compression algorithms.

The compression algorithm strongly depends on the data representation, which in turn very much depends on the view synthesis methods. View synthesis approaches can be classified in two classes: Geometry-based rendering (GBR) approaches and Image-based rendering (IBR) approaches. GBR methods use a detailed 3D model of the scene. These methods are useful with synthetic video data but they become inadequate with real multi-view videos, where 3D models are difficult to estimate. IBR approaches are an attractive alternative to GBR. They allow the generation of photo-realistic virtual views. The Layer Depth Image (LDI) representation is one of these IBR approaches. In this representation, pixels are no more composed by a single color and a single depth value, but can contain several colors and associated depth values. This representation reduces efficiently the multi-view video size, and offers a fast photo-realistic rendering, even with complex scene geometry.

Extending the work done in 2009 on the Incremental LDI construction, we have developed in 2010 an efficient LDI compression algorithm and a fast virtual view rendering method. The LDI is compressed using the MVC coder on both texture and depth maps. Each layer is considered as a temporal sequence. The first layer is compressed as a normal video, and extra layers are predicted from this first layer using the MVC encoder. Thanks to the LDI construction, the background layer is temporally fixed, and each extra layer is easy to predict from the first one. To synthesize virtual views of the scene, we have developed an ordered projection, permitting fast detection and easy inpainting of cracks and other small disocclusions. Results are shown in figure 3.
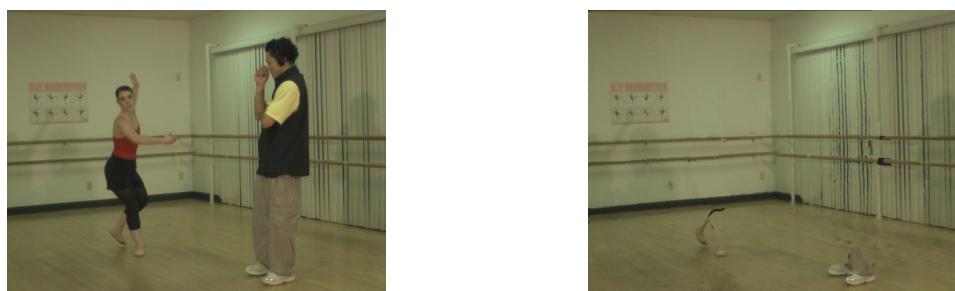
(a) LDI Foreground   (b) LDI Background

Figure 2. Textures of a 2 layers LDI
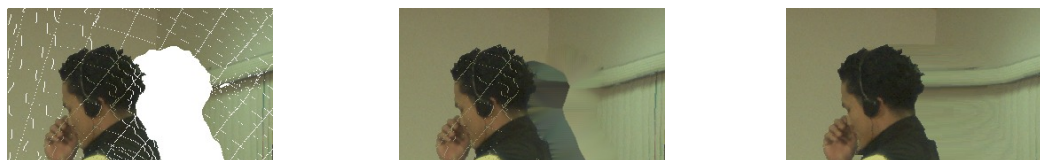
(a) Disocclusions and cracks   (b) Inpainting (Telea's method)   (c) Ordered Projection

Figure 3. Rendering results

6.1.3. Inpainting

Participants: Josselin Gauthier, Christine Guillemot, Olivier Le Meur.

The inpainting challenge is to estimate missing parts of an image or a video, e.g. replacing foreground objects with a visually pleasing and plausible background. Two kinds of algorithms have been proposed in the past. In a first category, inpainting methods fill holes by propagating linear structures. These algorithms are inspired by the partial differential equations of physical heat flows (diffusion-scheme). The main drawback is that the diffusion introduces some blur, which is noticeable and annoying when the hole to be filled is large. To cope with this drawback, a second kind of inpainting algorithm uses textures sampled from the known parts of the picture in order to fill in unknown parts of the picture. These methods are called examplar-based techniques. Since 2010, we have developed a new inpainting algorithm which combines the strengths of diffusion-schemes and examplar-based schemes. This algorithm has been used for solving problems of disocclusions in multi-view processing and virtual view synthesis for free view-point navigation. The inpainting algorithm is in this case aided by the depth information. The results are illustrated in Figure 4 which shows the region of

disocclusion in a synthesized virtual view as well as the inpainting results obtained with a method mixing tensor based diffusion and examplar-based texture synthesis, but which does not take into account the depth information, and the depth-aided approach developed in the team.

| (a) Disocclusion region | (b) Inpainting result with a classical method (Criminisi's) | (c) Inpainting result with the depth-aided method developed in the team |

Figure 4. Inpainting results.

6.1.4. *Computational modelling of visual attention*
Participants: Josselin Gauthier, Olivier Le Meur.

In the past few years, there has been a growing interest in observer-centric applications. In this context, the computational modelling of visual attention plays an important role. Since 2010, we have worked on the design of a new visual attention model. The proposed method is based on both low-level visual features and higher-level visual information. Indeed, when people gaze at real scenes, eye movements are influenced both by a set of bottom-up processes and by top-down effects such as the task, prior knowledge or the semantic context. We are more interested in prior knowledge which is strongly related to visual inferences. Visual inferences that are unconscious stem from the perceptual learning. The goal is then to infer from the visual properties of the scene prior knowledge. In 2010, we have focused on three visual inferences: the dominant depth, the type of the scene and the position of the horizon line. Three specialized detectors have been designed. These three features will be used in the model in order to improve the prediction of salient areas. For instance, given the type of the scene (indoor or outdoor), a particular strategy might be proposed to define salient areas.

6.2. Sparse representation, compression and interaction with indexing

6.2.1. *High-level feature extraction*
Participants: Safa Cherigui, Christine Guillemot.

Significant research effort has been, in the past years, dedicated to low-level video signal features analysis, e.g. motion analysis, segmentation, multi-resolution analysis. However, accounting for higher-level features appears necessary for further progress. In 2010, we have worked on the design of a new method for extracting a global condensed representation of the scene, based on computer vision techniques tracking self similarities within and across images. This method aims at constructing a condenses representation known as epitome in the literature. Although earlier work has shown that epitomes could be powerful tools for segmentation, denoising, recognition, indexing, and texture synthesis, further considerations, as their mean description length, or robustness, need to be taken into account in a compression and communication context, likely to lead to specialized construction methods. A novel method has been developed for constructing the epitome representation of an image. The epitome has been used for image compression showing significant performance gain with respect to H.264 Intra coding.

6.2.2. *Sparse signal representation algorithms*
Participants: Angélique Drémeau, Jean-Jacques Fuchs, Cédric Herzet.

Sparse representations and compressed sensing continue to be fashionable domains in the signal processing community and numerous pre-existing areas are now labeled this way. Sparse approximation methods aim at finding representations of a signal with a small number of components taken from an overcomplete dictionary of elementary functions. The problem basically involves solving an under-determined system of equations with the constraint that the solution vector has the minimum number of non-zero elements. Except for the exhaustive combinatorial approach, there is no known method to find the exact solution under general conditions on the dictionary. Among the various algorithms that find approximate solutions, pursuit algorithms (matching pursuit, orthogonal matching pursuit or basis pursuit) are the most well-known.

In 2010, we have applied convex optimization ideas to array synthesis and observed that sparsity is present in the optimal design. More precisely if one considers the synthesis of the basic narrow main beam low sidelobe linear array with constrained length, the optimal array is indeed the sparsest one, the one having the smallest number of elements necessary to prevent the appearance of grating lobes. A posteriori this is probably not too surprising since minimizing the ℓ_∞-norm, one can expect the emergence of its dual the ℓ_1-norm and hence this sparseness result.

We have also considered means to convey available a priori information on the observations to the optimization criterion when seeking a sparse representation of a signal on a redundant basis. More or less efficient or elegant solutions to these problems have been proposed, modifying the penalty term in a sparse representation criterion is one of them. We show investigated ways to translate prior information by modifying the penalization term of the usual $\ell_2 - \ell_1$ regularized criterion and analyzed how to tune the corresponding hyper-parameters by forming the dual of these modified criterion. We have evaluated the associated performance on the sum of harmonics example where taking into account the structure of each individual harmonic signal definitely improves the efficiency of an estimator of the fundamental frequencies. Less trivial applications include the case where two or more measurement vectors are available that weight differently a same sparse set of atoms to be identified.

6.2.3. *Dictionary learning methods for sparse representations*

Participants: Christine Guillemot, Joaquin Zepeda.

Closely related to the sparse representation problem is the design of dictionaries adapted to "sparse" problems. The sparsity of the signal representation indeed depends on how well the bases match with the local signal characteristics. The adaptation of the transform to the image characteristics can mainly be made at two levels: i) in the spatial domain by adapting the support of the transform; ii) in the transformed domain by adapting the atoms of the projection basis to the signal characteristics.

In 2010, in collaboration with Ewa Kijak (TexMex), we have developed methods for learning dictionaries to be used for sparse signal representations. These methods lead to dictionaries which have been called Iteration-Tuned Dictionaries (ITDs), Basic ITD (BITD), Tree-Structured ITD (TSITD) and Iteration-Tuned and Aligned Dictionaries (ITAD). Iteration-Tuned Dictionaries (ITDs) generalize traditional overcomplete dictionaries for sparse representations by adapting them to the iterative nature of practical decomposition schemes. The general setup is based on the Matching Pursuit (MP) algorithm: The MP algorithm selects a single atom from a fixed dictionary in each iteration. MP schemes based on ITDs instead choose an atom from an iteration-dependent dictionary that varies from one iteration to the next. In the Basic ITD (BITD) setup, a single dictionary is made available for any given MP iteration. This scheme can nonetheless be generalized by considering instead that many possible "candidate" dictionaries are available for use in a given iteration.

The Tree-Structured ITD (TSITD) setup is one such case wherein the candidate dictionaries are organized into a tree, with each tree layer corresponding to one MP iteration. Each node of the TSITD tree contains a single dictionary matrix, and each atom of the matrix in turn gives rise to one child node. The large size of the TSITD tree poses a problem concerning TSITD training and storage, and this issue is addressed by the Iteration-Tuned and Aligned Dictionary (ITAD), wherein all the candidate dictionaries of a given layer are forced to be a rotated version of a single, layer-dependent, prototype dictionary.

All three proposed ITD schemes (BITD, TSITD and ITAD) have been shown to outperform the state-of-the-art learned dictionaries in terms of PSNR versus sparsity. The performance of these dictionaries has also been assessed for both compression and denoising applications. ITAD in particular has been used to produce a new image codec that outperforms JPEG2000 for a fixed image class.

6.2.4. *Analysis and feature extraction for indexing*
Participants: Christine Guillemot, François Lecellier, Fethi Smach.

In collaboration with TexMex (H. Jégou) and Technicolor (P. Perez), we have worked on a trajectory descriptor for video search in the context of the ANR ICOS-HD project. This descriptor is based on tracklets, which can been seen as short trajectories. A tracklet descriptor is produced and exploited according to the following steps:

- interest point detection on the first video frame, seen as a image sequence. These points are tracked in time. If a point is lost, then a new point is detected and tracked;

- transform trajectories to produce descriptors invariant to different transformations ;

- comparison methods for tracklets descriptors.

We have used the Kanade Lucas Tomasi (KLT) method for interest point tracking. In brief, salient properties are detected and tracked based on the Newton-Raphson method, which minimizes an error cost between two successive frames by considering local windows around an interest point. At this stage, the trajectories are not invariant to scaling, rotation and translation, and therefore can not be straightforwardly used for our application. As rotations are rare, we focused on the translation and scaling transformations. For this purpose, we extract sub-trajectories pour a small group of frames and subsequently normalize the tracklets in scale and position to obtain the desire invariance. For each frame group, we then obtain a set of tracklets. Each set is aggregated using a method inspired by the Fisher Kernel representation of Perronnin. As a result, the output descriptor is more compact than the original set, and is amenable to be compared using a standard metric, such as the Euclidean distance. The results obtained on the evaluation set considered in the context of the ICOS-HD project shows a success rate of 100 % when the query video is a scaled or cropped version of the reference video. Note however that this approach, in its current form, is not invariant to rotation and flipping.

6.2.5. *Texture synthesis and prediction based on sparse approximations*
Participants: Christine Guillemot, Mehmet Turkan.

The problem of texture prediction can be regarded as a problem of texture synthesis. Methods based on sparse approximations, and using orthogonal matching pursuit and basis pursuit algorithms, have been investigated for this texture synthesis and prediction problem. The problem is looked at as a problem of texture synthesis (or inpainting) from noisy observations taken from a causal neighborhood. The goal of sparse approximation techniques is to look for a linear expansion approximating the analyzed signal in terms of functions chosen from a large and redundant set (dictionnary). In the methods developed, the sparse signal approximation is run in a way that allows for the same operation to be done at the decoder, i.e. by taking the previously decoded neighborhood as the known support. The sparse signal approximation is thus run with a set of *masked* basis functions, the masked samples corresponding to the location of the pixels to be predicted. The decoder proceeds in a similar manner by running the algorithm with the *masked* basis functions and by taking the previously decoded neighborhood as the known support.

Since a good representation of the support region does not necessarily lead to a good approximation of the block to be predicted, the 'sparsity level minimizing a chosen criterion (e.g. mean square error (MSE) or a rate-distortion (RD) cost function) needs to be transmitted to the decoder. The other drawback of this approach is that the atoms may not span the signal residue space "of the block to be predicted at the each iteration" even though the dictionary has been well adapted in the spatial domain. As a result, the minimization of encoded residual information has not been optimized even if the signal prediction seems sufficient in terms the chosen criterion. In order to overcome above described problems, in 2010, a novel spatial texture prediction method based on non-negative matrix factorization (NMF) is considered, and assessed comparatively to the template

matching and sparse approximation based techniques. It has been proved that the NMF based prediction method can be an effective solution to above described drawbacks of sparse approximation approach, and it offers a powerful and efficient alternative for minimizing the encoded information of an image or an intra frame.

A fast and efficient solution to iterative texture prediction problem is also highly dependent on the choice of the dictionary. For this task, the main consideration in employing a dictionary construction model is basically based on two approaches: the analytic approach and the learning (training) approach. In practice, for natural image signals, it has already proven that learning the dictionary A from a set of training samples leads to much better reconstruction results since the resulting trained dictionary has been well-adapted and fine-tuned to the training set of samples of a given particular structural data or an image. Some examples of learning based approach include the K-SVD, the method of optimal directions (MOD), (Generalized) Principle Component Analysis (PCA), and Sparse Orthonormal Transforms (SOT), and so on. Here, the main consideration on constructing a dictionary A for the prediction purposes rather than image transformation, or image denoising. For this purpose, we have proposed a fast and online dictionary learning structure, so called On the Fly Dictionaries (OFD), which can be seen as an application of above introduced learning based methods into texture prediction such as block based image prediction (and therefore image compression) and texture synthesis such as image inpainting. The OFD structure is based on a combinational set of level-based dictionaries that require an underlying sparsity model. The advantages of this novel structure include stability in adaptation of dictionaries under quantization noise in the case of the image compression problem, flexibility in adaptation of dictionaries to variable number of pixels to be synthesized (especially in the inpainting problem) and, low-complexity and locally-adaptive fast online training.

6.2.6. *Lossless coding for medical images*
Participants: Claude Labit, Jonathan Taquet.

Many techniques, both for lossy and lossless compression of biomedical images, have been introduced in previous years : a survey of these methods can be found in [68], including some experimental results on MRI and computed tomography (CT) images. In 2010, we have developed two algorithms for coding medical images.

The first algorithm is defined for lossless and near-lossless compression of 2D images, in order to allow a fast access to randomly selected slices [46], [58]. It uses a resolution scalable representation that provide quick-view navigation facilities. This approach combines DPCM techniques with hierarchical interpolation in order to provide a hierarchical oriented prediction (HOP) with adaptive capabilities. Its lossless performances are nearly equivalent to DPCM, but the algorithm also provides resolution scalability. The HOP algorithm is also well suited for near-lossless compression which is maximum peak of absolute error (PAE) controlled lossy compression. It provides interesting rate-distortion trade-off compared to JPEG-LS and equivalent or better PSNR than JPEG-2000 for high bit-rate (slight losses) on noisy native CT and MRI. For a PAE equal to 4 : on native CT images, the PSNR of HOP is 0.15dB better than JPEG-LS and 1dB better than JPEG-2000 for a same rate; on native MRI images, the PSNR is equivalent to tghe one obtained with JPEG-LS for a rate reduced by 0.2bpp, and equivalent to JPEG-2000 (that do not provide PAE controlled losses).

The second algorithm is based on optimized wavelet decomposition (OWD) [47], and can be used both in 2D or volumic images (sequence of slices). It adjusts the decomposition structure and associated supports lengths of a filter family (Lanczos, or Deslauriers-Dubuc, ...) in order to better adapt to the content of each biomedical image. This content is not always native: it may be post-processed by the radiologist, or smoothed by techniques used during the reconstruction. On this kind of images, the lossless compression can be improved by up to 30% using the proposed approach while staying equivalent on noisy images (compared to standard lossless wavelet decomposition). Since wavelets are used, the compressed data can be easily structured in progressive data flows either in terms of quality or resolution. Note that HOP is better suited for random access to noisy images (native MRI or CT) and for their (near-)lossless archiving, while OWD, also exploiting inter-slice redundancy might be more suitable for lossless (or lossy) archiving and progressive transmission.

6.3. Joint source-channel and distributed coding

6.3.1. *Slepian-Wolf coding of memory sources and application to Distributed Video coding*

Participants: Christine Guillemot, Aline Roumy, Velotiaray Toto-Zarasoa.

In 2010, we addressed the problem of distributed source coding for sources with memory. The challenge is to exploit both inter- and intra-(memory) correlation in a distributed coder i.e. without any cooperation between the encoders. This is still a challenging problem because if the two correlation types are processed separately, it leads to sub-optimality. Therefore we worked on the construction of codes and their related decoding algorithm in order to jointly exploit all correlations of the signal. Building on the optimality of channel codes for the distributed compression of memoryless sources, we proposed a scheme based on channel codes. As an application, we considered Distributed Video Coding, a video compression system that builds upon the idea of distributed source coding in order to achieve efficient video compression while *(i) maintaining a low complexity at the encoder and (ii) being robust to noise.*

For this application to Distributed Video Coding, we first propose an efficient model for the sources: the original images to be compressed and their prediction available at the decoder. Introducing memory and non-uniformity in the bitplanes was straightforward. As for the correlation model, we showed that the usual additive model did not fit well the data. More precisely, in the additive model, the prediction noise is assumed to be independent of the original image which is not always true. We therefore proposed a combined model, where the correlation noise can be either independent of the original signal or of its prediction. For this model, we derived the achievable compression rates and the gain with respect to the case where the sources are modeled as uniform i.i.d. sources.

To summarize, we introduced a new model well suited for the sources in Distributed Video coding. The new model includes (i) a new (inter) correlation model (i.e. between the sources), (ii) takes into account the non uniformity and the memory of each source (Hidden Markov models are used). We derived the achievable compression rates for these models and derived an algorithm that can jointly estimate the source parameters, the source symbols, the correlation parameter and the correlation type [14]. The Distributed Video Coding system that integrates the enhancement that we propose here demonstrates a quality-versus-rate improvement by up to 10.14% [49], with respect to its elder version. A simplified model with non-uniform sources is also proposed that achieves an improvement by up to 5.7% [51].

In DSC it is usually assumed that the source parameters (source and correlation parameter) is known at both encoder and decoder. However, as in the work mentioned above [51], [49], [14], these parameters have to be estimated. In this problem called multiterminal estimation theory, only lower bounds of the Cramer Rao bound exist. In [51], [49], [14] joint parameter and data estimation has been performed with an EM (Expectation Maximization) algorithm. However this algorithm is very sensitive to the initialization. Therefore, a novel estimate to initialize this EM algorithm has been proposed. When the correlation model is a Binary Symmetric Channel (BSC), we proposed a ML estimate based on the subset of the data available at the decoder. This estimator can be performed prior to the decoding and can therefore initialize the EM algorithm. The efficiency of the estimator and the convergence of the estimator to the lower bound of the Cramer Rao bound has been shown [14]. It therefore shows that the lower bound is tight. This work has been submitted to IEEE Communication Letter.

We also addressed the problem of non-asymmetric Slepian-Wolf (SW) coding of two correlated non-uniform Bernoulli sources. We first showed that the problem is not symmetric in the two sources, contrarily to the case of uniform sources, due to the asymmetry induced by two underlying channel models, namely additive and predictive BSC. That asymmetry has to be accounted for during the decoding. In view of that result, we described the implementation of a joint nonasymmetric decoder of the two sources based on Low-Density Parity-Check (LDPC) codes and Message Passing (MP) decoding. We also gave a necessary and sufficient condition for the recovery of the two sources, that imposes a triangular structure of a subpart of the equivalent matrix representation of the code [50].

6.3.2. *Distributed compressed sensing*

Participants: Christine Guillemot, Aline Roumy, Velotiaray Toto-Zarasoa.

We considered the distributed compression of correlated sources with memory (a similar problem as in Section 6.3.1) but with a compressed sensing (CS) approach. CS attempts to avoid the usual approach where a very large number of samples is first acquired and then most of them are discarded. Instead, in CS sampling is performed by taking a number of projections of the signal onto pseudorandom sequences, which yields a universal and low-complexity way to represent a signal. In CS and Distributed CS (DCS), perfect reconstruction of the signals are considered. Here, we addressed the DCS problem in a rate distortion framework and proposed a system for coding of intra- and inter-correlated sources in application scenarios such as sensor networks. It is based on the joint use of CS to capture memory of a signal, and DSC to take advantage of intersensor correlations. The proposed system has satisfactory rate-distortion performance in comparison with a system that does not employ CS but only DSC, showing that linear measurements represent a viable, universal and low-complexity signal representation. We showed that quantizing the measurements performs better than quantizing sources and then applying measurements, due to the resilience of CS to quantization noise. Moreover, the optimal number of measurements can be chosen as the one guaranteeing (close-to-)perfect reconstruction. In addition, using joint decoding, dequantization and reconstruction techniques allows to boost performance even further, making the proposed scheme an attractive choice for environments such as sensor networks, in which the devices performing acquisition and processing are severely constrained in terms of energy and computations. This work has been performed in collaboration with E. Magli and G. Coluccia (Politecnico di Torino) in the framework of the NewCom++ project. It has been submitted to the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 11).

6.3.3. *Information theory and covert channels*
Participant: Aline Roumy.

In [40], we considered the problem of steganography in legitimate channels and proposed a novel method to detect covert channels prior to the use of the forbidden communication. First, we modeled the global system to be tested in terms of transition systems and used information theory to measure a possible information leak. Previous attempts to use information theory in this context, assumed a simple memoryless model for the channel. Here, we showed that the system has to be modeled as a finite state channel model, where the input/output alphabets depends on the state. For the case of memoryless states, independent of the input/output, we provided capacity computation in order to measure the information leak. This allowed us to show that the usual notion of interference does not characterize the notion of deliberate information flow of covert channels. This work has been performed in collaboration with L. Helouet (INRIA, EPI DistribCom).

6.4. Image and video watermarking and security

6.4.1. *Watermarking for fingerprinting*
Participants: Ana Charpentier, Teddy Furon, Fuchun Xie.

On top of being robust and invisible, a watermarking for video fingerprinting must be also secure. The management of the watermarking secret key might be painful in practical applications and therefore this key might not be changed very frequently. This leads to the following security threat: the collusion has a big amount of video blocks, all watermarked with the same technique and the same watermarking secret key. By a careful analysis, colluders might disclose this key. Once compromised, they can erase the watermark so that the accusation fails tracing them despite the use of a good traitor tracing code. Even worse, they can create new symbols embedded in the pirated copy leading to accusing an innocent.

Our approach was to start from a state-of-the-art robust and invisible zero-bit watermarking technique. This technique has been deeply benchmarked during the international challenge BOWS-2 (The 2nd edition of the 'Break Our Watermarking Scheme competition). During this challenge, two security attacks were mounted succeeding in a good enough estimation of the secret key within 5,000 watermarked images. We achieved to patch these attacks raising the security level up to 100,000 images while almost maintaining the same robustness [53]. However, this attacks / counter-attacks cycle doesn't mean that the technique is now secure since other attacks might be discovered. This is the reason why we did try to tackle the class of attacks based

on second-order statistics (with tools like PCA, ICA, OPAST, etc). But our results were very mitigated because a perfect security level against this class of attacks is possible but with either a big loss of robustness, either a loss of reliability (ie. the probability of false alarm of the watermarking detection is bigger of several orders of magnitude) [39].

6.4.2. *Protocols for fingerprinting*
Participants: Ana Charpentier, Teddy Furon.

A key assumption of the traitor tracing schemes developed so far is that the colluders may know their own codewords but they ignore the codeword of any other innocent user. Otherwise, the collusion can very easily forge a pirated content framing an innocent user because it contains a sequence close enough to his/her codeword. This puts a lot of pressure on the versioning mechanism which creates the personal copy of the content in accordance to a codeword. For instance, suppose that the versioning is done in the user's setup box, the unique codeword being loaded into this device at the manufacture. If the code matrix ends up in the hands of an untrustworthy employee, then the whole fingerprinting system is pulled down. This is one argument of the motivation for designing cryptographic protocols for the construction, the versioning and the accusation. We have proposed a new asymmetric fingerprinting protocol dedicated to the state-of the-art Tardos codes. We believe that this is the first such protocol, and that it is practically efficient. The construction of the fingerprints and their embedding within pieces of content is based on oblivious transfer and do not need a trusted third party. Note, however, that during the accusation stage, a trusted third party, like a Judge, is necessary like in any asymmetric fingerprinting scheme we are aware of. Further work is needed to determine if such a third-party can be eliminated. In particular, we anticipate that some form of secure multi-party computation can be applied.

6.4.3. *New decoders for fingerprinting*
Participants: Ana Charpentier, Teddy Furon.

So far, the accusation process of a Tardos code is based on single decoders which compute a score per user. Users with the highest score or whose scores is above a threshold are then deemed guilty. In the past years, we have contributed to this approach bringing two improvements: the 'learn and match' strategy aims at estimating the collusion process and using the matched score function; a rare event analysis translates this score into a more meaningful probability of being guilty. A fast implementation computes the scores of one million of users within 0.2 second on a regular laptop. Therefore, contrary to common belief, although a single decoder is exhaustive with a linear complexity in $O(n)$, it is not slow.

This fast implementation allows us to propose iterative decoders. A first idea is that conditioning by the identities of some colluders bring more discrimination power to the score function. The first iteration is thus a single decoder, users we are extremely confident to accuse are enrolled as side information. The next iteration computes new scores for the remaining users etc. A second idea is that information theory proves that a joint decoder computing scores for pairs, triplets, or in general t-tuples is more powerful than single decoders working with scores for single users. However, nobody did try them since the number of t-tuples is in $O(n^t)$. We propose in a first iteration to use a single decoder, to prune out users who are definitively innocents (because their scores are low) and keeping $O(\sqrt{n})$ individual suspects. The second iteration is a joint decoding working on pairs of users etc. Iteratively, we prune out enough users such that it is manageable to run a joint decoder on bigger t-tuples. A paper applying these ideas on group testing has been submitted to IEEE ICASSP, its version dedicated to traitor tracing will be submitted to Information Hiding in december.

7. Contracts and Grants with Industry

7.1. Industrial contracts

7.1.1. *CIFRE contract with France Telecom on the problem of 3D video representation*
Participants: Thomas Colleu, Claude Labit.

- Title : 3D video representation for 3DTV and Free viewpoint TV.
- Research axis : § 6.1.1.
- Partners : France Télécom, Irisa/Inria-Rennes.
- Funding : France Télécom.
- Period : Oct.07-Sept.10.

This contract with France Telecom R&D (started in October 2007) aims at investigating the data representation for 3DTV and Free Viewpoint TV applications. 3DTV applications consist in 3D relief rendering and Free Viewpoint TV applications consist in an interactive navigation inside the content. As an input, multiple color videos and depth maps are given to our system. The goal is to process this data so as to obtain a compact representation suitable for 3D TV and Free Viewpoint functionalities. We have developed a multi-view video representation based on a global geometric model composed with a soup of polygons. In 2009, the construction of the representation has been improved in order to reduce the number of quads and eliminate artifacts around depth discontinuities as well as the corresponding rendering approach (see Section 6.1.1). A coding algorithm adapted to the polygon soup representation has been developed. The performance of the coding scheme has been compared with the encoding of the depth maps using JPEG2000. In 2010, a multi-view video coding scheme based on the soup of polygon representation has been developed.

7.1.2. CIFRE contract with Thomson on the problem of spectral deconvolution for video compression and protection

Participants: Safa Cherigui, Christine Guillemot.

- Title : Sparse modeling of spatio-temporal scenes
- Research axis : § 6.2.1.
- Partners : Thomson, Irisa/Inria-Rennes.
- Funding : Thomson, ANRT.
- Period : Nov.09- Oct.12.

This CIFRE contract concerns the Ph.D of Safa Cherigui. The objective is to investigate texture and scene characterization methods and models based on concepts of spatio-temporal epitomes and/or signatures for different image processing problems focusing in particular on video compression and editing. A novel method has been developed for constructing the epitome representation of an image. The epitome has been used for image compression showing significant performance gain with respect to H.264 Intra coding.

7.1.3. Collaboration with Alcatel on robust video compression

Participants: Christine Guillemot, Laurent Guillo, Aline Roumy, Velotiaray Toto-Zarasoa.

- Title: Self adaptive video codec
- Funding: Joint research laboratory between INRIA and Alcatel
- Period: Oct. 2010 - May. 2011.

In the framework of the joint research lab between Alcatel-Lucent and INRIA, we participate in the ADR (action de recherche) Selfnets (or Self optimizing wireless networks). More precisely, we collaborate with the Alcatel-Lucent team on a self adaptive video codec. The goal is to design a video codec, which is able to self-adapt to the existing underlying transport network. Therefore the video codec has to include

- Means at the encoder to adapt dynamically the output bitrate to the estimated channel throughput and to the effective transport QoS while maintaining the video quality requirements.
- Means at the decoder to be resilient to any remaining packet losses.

Adaptation to the network and the receiver will be achieved thanks to the usage of the scalable video codec WSVC (see Section 5.1.6). As for the robustness against losses, the approach is to introduce dedicated video redundancies and therefore outperform typical FEC schemes. A novel approach based on Wyner-Ziv coding (or source coding with side information) has shown substantial gain compared to the typical FEC scheme. The challenge is to make this approach scalable and also adaptable to any erasure rate.

7.2. National contracts

7.2.1. *DGE/Region-FUTURIMAGE*

Participants: Christine Guillemot, Vincent Jantet, Claude Labit.

- Title : 3D multi-view video transmission and restitution.
- Research axis : § 6.1.2.
- Partners : Thomson R&d France, Thomson Grass Valley, Orange Labs, Alcatel-Lucent, Irisa/Inria-Rennes, IRCCyN, Polymorph Software, BreizhTech, Artefacto, Bilboquet, France 3 Ouest, TDF.
- Funding : Region.
- Period : Oct.08-Aug.10.

The Futurim@ages project studies coding, distribution and rendering aspects of future television video formats: 3DTV, high-dynamic range videos, and full-HD TV. In this context, TEMICS focuses on compact representations and restitution of multi-view videos. Multi-view videos provide interesting 3D functionalities, such as 3DTV (visualization of 3D videos on auto-stereoscopic screen devices) or Free Viewpoint Video (FFV, i.e. the ability to change the camera point of view while the video is visualized). However, multi-view videos represent a huge amount of redundant data compared with standard videos, hence the need to develop efficient compression algorithms. Stereoscopic or auto-stereoscopic devices display very specific camera viewpoints, which should be generated even if they do not correspond to acquisition viewpoints. Artifacts such as ghosting or bad modelled occlusions must be dealt with to render high quality 3D videos.

In the Futurim@ages project, we have addressed the problem of depth map retrieval from multi-view and from monocular videos of static scenes. For the latter case, we rebuilt the Structure from Motion part of the 3D video codec developed in the team by integrating the lattest state-of-the-art vision algorithms. As for the multi-view case, we developed a new depth map estimation algorithm which aims at preserving depth discontinuities, which is necessary to render correct virtual views of the scene. Methods for constructing LDI representations of multi-view plus depth video sequences have also been developed (see Section 6.1.2).

7.2.2. *RIAM-MEDIEVALS*

Participants: Caroline Fontaine, Teddy Furon, Ana Charpentier.

- Title : Watermarking and Visual Encryption for Video and Audio on Demand legal diffusion
- Research axis : § 6.4.1.
- Partners : MEDIALIVE, LSS (Univ. Paris XI/Supelec), GET-INT, THOMSON Software et Technologies Solutions, AMOSSYS SAS.
- Funding : ANR.
- Period : 28/12/2007-28/12/2010

MEDIEVALS is a project dealing with the diffusion of video or audio on demand. MEDIALIVE developed a software to secure this delivery through visual encryption, and the goal of the project is to add watermarking/fingerprinting in the process to improve the security of the delivered content. In 2010, Temics developed cryptographic protocols for secure deployment of traitor tracing codes. We also proposed more efficient decoders which allows shorter fingerprinting codes.

7.2.3. NAVIMED

Participants: Claude Labit, Jonathan Taquet.

- Title : Compression algorithms for volumetric medical images
- Research axis : § 6.2.6.
- Partners : C2 innovativ'Systems, ETIAM.
- Funding : Brittany Region.
- Period : 06/2010-06/2012

The objective of the project is to develop lossless and near-lossles compression algorithms for volumetric images (like computed tomography or magnetic resonance imaging), an optimized storage and archiving system, as well as a tool allowing navigation within the medical images from a thin web-based client. Two algorithms for coding MRI and CT images have been developed showing interesting performance gains compared with state of the art methods used for compressing medical images (see Section 6.2.6)

7.2.4. ANR-PERSEE

Participants: Angélique Drémeau, Josselin Gauthier, Christine Guillemot, Laurent Guillo, Olivier Le Meur.

- Title : Perceptual coding for 2D and 3D images.
- Research axis : § 6.1.4.
- Partners : IRCCYN-Polytech Nantes, INSA-Rennes, Telecom Paris Tech.
- Funding : ANR.
- Period : 10/2009-09/2012

The objective of the project is to develop perceptually driven coding solutions for mono-view and multi-view video. TEMICS contributes on different problems relevant for mono-view and multi-view video coding: visual attention modeling (see Section 6.1.4), on texture synthesis and inpainting for both 2D and 3D content (see Sections 6.2.5 and 6.1.3).

7.2.5. ANR-COHDEQ

Participant: Jean-Jacques Fuchs.

The RNRT project COHDEQ 40 "COHerent DEtection for QPSK 40GHz/s systems" whose coordinator is Alcatel has started in January 2007. Its aim was to establish the feasibility of coherent detection in optical fibers transmission systems. As far as Irisa is concerned, the work is done by ASPI in collaboration with TEMICS. It covers all the signal processing aspects of this specific digital communication system that will be able to achieve a 100 Gbit/s channel rate.

This project ended on June 16, 2010. It is worth noting that G. Charlet the project leader from Alcatel-Lucent (Nozay, France) was selected by the MIT magazine 'Technology Review' among its '2010-Young-Innovators' for his "Record-breaking optical fibers for global communications", an outcome of this project.

7.2.6. ANR-TCHATER

Participant: Jean-Jacques Fuchs.

The RNRT project TCHATER "Terminal Coherent Hétérodyne Adaptatif TEmps Réél" whose coordinator is Alcatel has started in January 2008. It aims to fully implement coherent detection in an optical fibers transmission systems, with among others the real time implementation on dedicated FPGA's that will be taken care off by the Inria-Arenaire team. As far as Irisa is concerned, the work is done by ASPI in collaboration with TEMICS. To adapt the extremely high channel rate, 4 ADC's (analog-to-digital converters) are needed and to accommodate the FPGA's to their output rate, temporal multiplexing of order 40 is required. In collaboration with the Arenaire project, the signal processing algorithms have been adapted and tuned to fit within the stringent real time constraints. This project has been delayed, by side effects of the financial crisis, and will end in 2011.

7.2.7. ANR-STRADE

Participant: Jean-Jacques Fuchs.

The project STRADE "Réseaux du Futur et Services" whose coordinator is Alcatel-Lucent started on November 1, 2009. It will run over a 3-year period and aims to investigate the potentialities of optical fibers with higher effective area than those used nowadays. The overall objective is to increase the global transmission capacity of a single fiber. As far as Irisa is concerned, the work is done by ASPI in collaboration with TEMICS and concerns the signal processing aspects. The fiber and the end-equipments are still under development and so far the signal processing algorithms have only been tested on simulated data whose validity and proximity to the real data is questionable.

7.2.8. ANR-ICOS-HD

Participants: Christine Guillemot, François Lecellier, Fethi Smach, Joaquin Zepeda.

- Title : Scalable Indexing and Compression Scalable for High Definition TV;
- Research axis : § 6.2.4.
- Partners : Université de Bordeaux, CNRS/I3S;
- Funding : ANR.
- Period : 01/01/2007-31/12/2010

The objective of the project was to develop new solutions of scalable description for High Definition video content to facilitate their editing, their acces via heterogeneous infrastructures (terminals, networks). The introduction of HDTV requires adaptations at different levels of the production and delivery chain. The access to the content for editing or delivery requires associating local or global spatio-temporal descritors to the content. In 2010, the TEMICS project-team contributed in particular on the study of new forms of signal representation amenable to both compression and feature extraction. The dictionary training methods described in Section 6.2.3 are part of TEMICS contributions to the ICOS-HD project together with the tracklet video descriptor mentioned in Section 6.2.4.

7.2.9. ANR-ARSSO

Participants: Mounira Ebdelli, Christine Guillemot, Laurent Guillo, Aline Roumy.

- Title : Adaptable, Robust, Streaming SOlutions.
- Partners : INRIA/Planète, TESA-ISAE, CEA-LETI/LNCA, ALCATEL LUCENT BELL LABS, THALES Communications, EUTELSAT SA.
- Funding : ANR.
- Period : 06/2010-11/2013

The ARSSO project focuses on multimedia content communication systems, characterized by more or less strict real-time communication constraints, within highly heterogeneous networks, and toward terminals potentially heterogeneous too. It follows that the transmission quality can largely differ in time and space. The solutions considered by the ARSSO project must therefore integrate robustness and dynamic adaptation mechanisms to cope with these features. The overall goal is to provide new algorithms, develop new streaming solutions and study their performances. TEMICS contribute on the development and improvement of scalable video coding techniques and components to make the video codec robust to losses. More specifically loss concealments methods will be developed.

7.3. European contracts

7.3.1. FP6-IST NOE NEWCOM++

Participants: Christine Guillemot, Aline Roumy.

- Title: NEWCOM: Network of Excellence in Wireless Communication.
- Research axis: 6.3.2
- Funding: CEE.
- Period: Jan. 2008 - Dec. 2010.

The NEWCOM++ project proposal (Network of Excellence in Wireless COMmunication) intends to create a trans-European virtual research centre on the topic "The Network of the Future". It was submitted to Call 1 of the VII Framework Program under the Objective ICT-2007.1.1: The Network of the Future, mainly in its target direction "Ubiquitous network infrastructure and architectures". We participate in the workpackage WPR7 - Joint source and channel co-decoding which we now coordinate together with the task TR7.3 Tools for multi-terminal JSCC/D. WPR7 addresses issues related to the robust transmission of multimedia, and essentially video, over wireless channels (possibly terminating a wired IP network). Such issues are : (i) solving the compatibility problem with the classical OSI layers separation (to what extent can we keep this separation ?) (ii) providing new tools (and expanding existing ones) for Joint Source and Channel Coding/decoding (JSCC/D) in classical one to one, one to many (broadcast), or distributed contexts (iii) providing new tools for analysing the efficiency of these tools (iv) working on practical, long term situations, which will be used as test-beds.

8. Dissemination

8.1. Awards

- M. Turkan has received a best student paper award at IEEE-ICIP, Oct. 2010, for the paper [52].
- J. Zepeda has received the second best student paper award at IEEE-MMSP, Oct. 2010, for the paper [54].

8.2. Patents

- A patent on "Méthodes de représentations parcimonieuses" (J. Zepeda, C. Guillemot, E. Kijak) has been filed in April 2010.

8.3. PhD and HDR committees outside the team

- C. Guillemot has been rapporteur of the PhD thesis of:
  - T. Maugey, Telecom ParisTech, 19 Nov. 2010;
  - F. Bassi, Univ. of Paris Sud 11, 3 Dec. 2010;
  - A. Zribi, Telecom Bretagne, 7 Dec. 2010;
  - A.S. Bacquet, Univ. of Valenciennes, 10 Dec. 2010;
- C. Guillemot has been rapporteur of the HDR of:
  - C. Delpha, Univ. of Paris Sud 11, 3 Dec. 2010;
- C. Labit has been rapporteur of the PhD thesis of:
  - Zafar Shashid, Univ. de Montpellier, 8 Oct. 2010;
  - Olivier Brouard, Univ. de Nantes, 20 July 2010.
- C. Labit has been rapporteur of the HDR of:
  - Titus Zaharia, Telecom SudParis, HDR Univ Paris 6, 22 Nov. 2010;
- O. Le Meur has been examiner of the PhD thesis of:
  - M. Perreira Da Silva, Univ. de la Rochelle, 10/12/2010.

8.4. Leadership within the scientific community

- T. Furon is associate editor of the EURASIP journal on Information Security.

- T. Furon is associate editor of the IET Journal on Information Security.
- T. Furon is a member of the IEEE Technical Committee of Information Forensics and Security.
- C. Guillemot is associate editor of the Eurasip International Journal on Image Communication (since Oct. 2010).
- C. Guillemot serves as a member in the award committee of the Eurasip Image communication journal (2007-2010).
- C. Guillemot serves as a member of the "Specif- Gilles Kahn Thesis Award" committee (2008-2010).
- C. Guillemot is a member of the Selection and Evaluation Committee of the "Pôle de Compétitivité" Images and Networks of the Region of Ouest of France (since Sept. 2007).
- C. Guillemot has served as expert in the evaluation of project proposals for OSEO-ANVAR, as well as in the evaluation panel of the Photonics department at DTU (Denmark Technical University, Copenhagen), Nov. 2-4, 2010.
- C. Guillemot is the coordinator of the ANR ICOS-HD project.
- C. Guillemot has been the general co-chair of the IEEE International Conf. on Multimedia Signal Processing, IEEE-MMSP 2010.
- C. Guillemot has been a member of the technical program committees of the Picture Coding Symposium (PCS'2010), and of the national conference CORESA'2010.
- C. Labit has served as a reviewer for the technical program committees of: Int. Conf of Image Processing, ICASSP, Eusipco.
- C. Labit is member of the GRETSI association board.
- C. Labit is, for the national INRIA's research department, scientific adviser of INRIA-SUPCOR (Support services for ANR collaborative research initiatives.
- C. Labit is the Scientific Board chairman of Rennes1 University (since June 1st, 2008).
- C. Labit is president of Rennes-Atalante Science Park.
- A. Roumy was a member of the technical program committee of CrownCom 2010.

8.5. Exhibitions and demonstrations

- The TEMICS project-team presented a demo at the exhibition held during the IEEE International conference on "MultiMedia Signal Processing, MMSP" in Saint-Malo, Oct. 2010.

8.6. Teaching

- Enic, Villeneuve-d'Ascq, (C. Guillemot: Video communication) ;
- Esat, Rennes, (C. Guillemot: Image and video compression) ;
- Engineer degree Diic-INC, Ifsic-Spm, university of Rennes 1 (O. Le Meur, L. Guillo : image processing, compression, communication);
- Engineer degree Diic-LSI, Ifsic-Spm, university of Rennes 1 (L. Guillo, O. Le Meur : compression, video streaming);
- Engineer degree ESIR, Université de Rennes 1: Jean Jacques Fuchs teaches several courses on basic signal processing and control ;
- Master Research 2: SISEA: Jean Jacques Fuchs teaches a course on optimization and sparse representations ; he also intervenes in the Joint International Program of the University of Rennes 1 and the SouthEast University of China (Nanjing) and teaches a course in Advanced Signal Processing in the International Master of Science in Electronics and Telecommunications.

- Master Research-2 SISEA: C. Guillemot and C. Labit teach a course on image and video compression ;

- Master Research-2 M2RI: C. Guillemot teaches a course on image and video compression ;

- Master 2 Psychologie de la cognition, University of Paris 8 (Psychologie des processus cognitifs, Mesure et modelisation): O. Le Meur teaches a course on "Selective visual attention: from experiments to computational models";

- Master MITIC, Univ. Rennes 1: O. Le Meur teaches a course on Acquisition/Image Processing/Compression;

- Master, Network Engineering, university of Rennes I: L. Guillo teaches a course on video streaming;

- Computer science and telecommunications magistère program, Ecole Normale Supérieure de Cachan, Ker Lann campus: A. Roumy teaches a course on Information theory.

9. Bibliography

Major publications by the team in recent years

[1] G. CAIRE, S. GUEMGHAR, A. ROUMY, S. VERDU. *Maximizing the spectral efficiency of coded CDMA under successive decoding*, in "IEEE Trans. on Information Theory", Jan. 2004, p. 152 - 164.

[2] F. CAYRE, C. FONTAINE, T. FURON. *Watermarking Security: Theory and Practice*, in "IEEE Transactions on Signal Processing", 2005, vol. 53, n° 10, p. 3976–3987, numéro spécial "Supplement on Secure Media III".

[3] V. CHAPPELIER, C. GUILLEMOT. *Oriented wavelet transform for image compression and denoising*, in "IEEE Transactions on Image Processing", Oct. 2006, vol. 15, n° 10, p. 2892-2903.

[4] J.-J. FUCHS. *On the application of the global matched filter to DOA estimation with uniform circular arrays.*, in "IEEE Trans. on Signal Processing", Apr. 2001, n° 49, p. 702-709.

[5] T. FURON, P. DUHAMEL. *An Asymmetric Watermarking Method*, in "IEEE Trans. on Signal Processing", April 2003, vol. 51, n° 4, p. 981-995, Special Issue on Signal Processing for Data Hiding in Digital Media and Secure Content Delivery.

[6] T. GUIONNET, C. GUILLEMOT. *Soft decoding and synchronization of arithmetic codes: Application to image transmission over noisy channels*, in "IEEE Transactions on Image Processing", Dec. 2003, vol. 12, p. 1599-1609.

[7] H. JÉGOU, C. GUILLEMOT. *Robust multiplexed codes for compression of heterogeneous data*, in "IEEE Trans. on Information Theory", Apr. 2005, vol. 51, n° 4, p. 1393-1403.

[8] L. OISEL, E. MÉMIN, L. MORIN, F. GALPIN. *One-dimensional dense disparity estimation for three-dimensional reconstruction*, in "IEEE Transactions on Image Processing", Sept. 2003, vol. 12, n° 9, p. 1107–1119.

[9] S. PATEUX, G. LE GUELVOUIT. *Practical watermarking scheme based on wide spread spectrum and game theory*, in "Signal Processing : Image Communication", Apr. 2003, n° 18, p. 283–296.

[10] G. RATH, C. GUILLEMOT. *Subspace-Based Error and Erasure Correction with DFT Codes for Wireless Channels*, in "IEEE Trans on Signal Processing", Nov. 2004, vol. 52, n⁰ 11, p. 3241-3252.

Publications of the year

Doctoral Dissertations and Habilitation Theses

[11] T. COLLEU. *Data representation, compression and rendering for 3D video*, Université Rennes 1, Dec. 2010.

[12] A. DRÉMEAU. *Décompositions parcimonieuses: approches Bayésiennes et application à la compression d'images*, Université Rennes 1, Nov. 2010.

[13] A. MARTIN. *Représentations parcimonieuses adaptées à la compression d'images*, Université Rennes 1, Apr 2010, http://hal.inria.fr/tel-00482804.

[14] V. TOTO-ZARASOA. *Codage de sources distribuées: Outils et applications à la compression vidéo*, Université Rennes 1, Nov. 2010, http://hal.inria.fr/tel-00539044.

[15] F. XIE. *Robust and Secure Watermarking for Multimedia Traitor Tracing*, Université Rennes 1, Sept. 2010.

[16] J. ZEPEDA. *Nouvelles méthodes de représentations parcimonieuses: application à la compression et à l'indexation d'images*, Université Rennes 1, Oct. 2010.

Articles in International Peer-Reviewed Journal

[17] Z. CHEN, C. GUILLEMOT. *Perceptually-friendly H.264/AVC Video Coding based on Foveated Just-Noticeable-Distortion Model*, in "IEEE Trans. on Circuits and Systems for Video Technology", June 2010, vol. 20, n⁰ 6, p. 806-819, http://hal.inria.fr/inria-00504260.

[18] T. COLLEU, S. PATEUX, L. MORIN, C. LABIT. *A polygon soup representation for multiview coding*, in "Journal of Visual Communication and Image Representation", Feb 2010, p. 1–32, http://hal.inria.fr/hal-00457634.

[19] O. CRAVE, B. PESQUET-POPESCU, C. GUILLEMOT. *Robust video coding based on multiple description scalar quantization with side information*, in "IEEE Trans. on Circuits and Systems for Video Technology, accepted", June 2010, vol. 20, n⁰ 6, p. 669-679, http://hal.inria.fr/inria-00504261.

[20] J.-J. FUCHS. *Des moindres carrés aux moindres déviations*, in "Traitement du Signal", 2010, vol. 27, p. 109-119.

[21] B. FUCHS, J.-J. FUCHS. *Optimal Narrow Beam Low Sidelobe Synthesis for Arbitrary Arrays*, in "IEEE Transactions on Antennas and Propagation", Jun 2010, vol. 58, n⁰ 6, p. 2130-2135, http://hal.inria.fr/hal-00488891.

[22] J.-J. FUCHS, B. FUCHS. *Synthesis of optimal narrow beam low sidelobe linear array with constrained length*, in "Progress in Electromagnetics Research B", 2010, vol. 25, p. 315-330, http://hal.archives-ouvertes.fr/hal-00519584/en/.

[23] T. Furon, G. Doërr. *Tracing pirated content on the Internet: Unwinding Ariadne's thread*, in "IEEE Security & Privacy Magazine", 2010, vol. 8, nᵒ 5, p. 69-71, http://hal.inria.fr/inria-00505851.

[24] O. Le Meur, J.-C. Chevet. *Relevance of a feed-forward model of visual attention for goal-oriented and free-viewing tasks*, in "IEEE Trans. on Image Processing", Nov. 2010, vol. 19, nᵒ 11, p. 2801-2813, http://hal.inria.fr/inria-00504259.

[25] O. Le Meur, S. Cloarec, P. Guillotel. *Automatic content repurposing for Mobile applications*, in "SMPTE Motion Imaging Journal", Feb. 2010, vol. 119, nᵒ 1, http://hal.inria.fr/inria-00504258.

[26] O. Le Meur, A. Ninassi, P. Le Callet, D. Barba. *Do video coding impairments disturb the visual attention deployment?*, in "Signal Processing Image Communication", Sept. 2010, vol. 25, nᵒ 8, p. 597-609, http://hal.inria.fr/hal-00494469.

[27] O. Le Meur, A. Ninassi, P. Le Callet, D. Barba. *Overt visual attention for free-viewing and quality assessment tasks: Impact of the regions of interest on a video quality metric*, in "Signal Processing Image Communication", Aug. 2010, vol. 25, nᵒ 7, p. 547-558, http://hal.inria.fr/hal-00494466.

[28] C. Morand, J. Benois-Pineau, J.-P. Domenger, J. Zepeda, E. Kijak, C. Guillemot. *Scalable Object-based Video Retrieval in HD Video DataBases*, in "Eurasip Journal on Signal Processing: Image Communication", July 2010, vol. 25, nᵒ 6, p. 450-465, http://hal.inria.fr/inria-00504257.

[29] N. Sellami, M. Siala, A. Roumy, I. Kammoun. *MAP sequence equalization for imperfect frequency selective channel knowledge*, in "ETT European Transactions on Telecommunications", 2010, vol. 21, p. 121-130, http://hal.inria.fr/inria-00504262.

International Peer-Reviewed Conference/Proceedings

[30] T.-T. Do, E. Kijak, T. Furon, L. Amsaleg. *Challenging the Security of Content Based Image Retrieval Systems*, in "IEEE MMSP", France Saint-Malo, IEEE, Oct 2010, http://hal.inria.fr/inria-00505846.

[31] T.-T. Do, E. Kijak, T. Furon, L. Amsaleg. *Deluding Image Recognition in SIFT-based CBIR Systems*, in "ACM Multimedia in Forensics, Security and Intelligence", Italie Firenze, ACM, Oct 2010, http://hal.inria.fr/inria-00505845.

[32] T.-T. Do, E. Kijak, T. Furon, L. Amsaleg. *Understanding the Security and Robustness of SIFT*, in "ACM Multimedia", Italie Firenze, Oct 2010, http://hal.inria.fr/inria-00505843.

[33] A. Drémeau, C. Herzet. *An EM-Algorithm Approach for the Design of Orthonormal Bases adapted to Sparse Representations*, in "Proc. IEEE Int'l Conf. on Acoustics, Speech and Signal Processing (ICASSP)", États-Unis Dallas, 2010, p. 2046-2049, http://hal.inria.fr/inria-00504511.

[34] A. Drémeau, C. Herzet, C. Guillemot, J.-J. Fuchs. *Anisotropic Multi-scale Sparse Learned Bases for Image Compression*, in "Proc. SPIE Electronic Imaging Conference on Visual Information Processing and Communication", États-Unis San Francisco, 2010, http://hal.inria.fr/inria-00504514.

[35] A. Drémeau, C. Herzet, C. Guillemot, J.-J. Fuchs. *Sparse Optimization with Directional DCT Bases for Image Compression*, in "Proc. IEEE Int'l Conf. on Acoustics, Speech and Signal Processing (ICASSP)", États-Unis Dallas, 2010, p. 1290-1293, http://hal.inria.fr/inria-00504510.

[36] A. DRÉMEAU, M. TURKAN, C. HERZET, C. GUILLEMOT, J.-J. FUCHS. *Spatial intra-prediction based on mixtures of sparse representations*, in "IEEE Intl. Workshop on Multimedia Signal Processing (MMSP'10)", Saint-Malo, France, Oct. 2010.

[37] J.-J. FUCHS. *DOA estimation in the presence of unknown colored noise: the Global Matched Filter approach.*, in "Proc. European Signal Processing Conference (EUSIPCO)", Aalborg, Aug. 2010, p. 1369-1373, http://hal.inria.fr/hal-00538945.

[38] J.-J. FUCHS. *The generalized likelihood ratio test and the sparse representations approach*, in "Proc. Intl. Conference on Image and Signal Processing, Icisp", Canada Trois Rivières, 2010, http://hal.inria.fr/inria-00504509.

[39] T. FURON, F. XIE, C. FONTAINE. *Towards Robust and Secure Watermarking*, in "ACM Multimedia and Security", Italie Roma, Sep 2010, http://hal.inria.fr/inria-00505849.

[40] L. HELOUET, A. ROUMY. *Covert channel detection using Information Theory*, in "International Workshop on Security Issues in Concurrency (SecCo)", August 2010, http://hal.inria.fr/inria-00539634.

[41] C. HERZET, A. DRÉMEAU. *Bayesian Pursuit Algorithms*, in "Proc. European Signal Processing Conference (EUSIPCO)", Aalborg, Aug. 2010.

[42] C. HERZET, A. DRÉMEAU. *Sparse Representation Algorithms based on Mean-Field Approximations*, in "Proc. IEEE Int'l Conf. on Acoustics, Speech and Signal Processing (ICASSP)", États-Unis Dallas, 2010, p. 2034-2037, http://hal.inria.fr/inria-00504512.

[43] A. MARTIN, J.-J. FUCHS, C. GUILLEMOT, D. THOREAU. *Phase refinement for image prediction based on sparse prediction*, in "Proc. SPIE Electronic Imaging Conference on Visual Information Processing and Communication", États-Unis San Francisco, 2010, http://hal.inria.fr/inria-00504515.

[44] T. MAUGEY, J. GAUTHIER, B. PESQUET-POPESCU, C. GUILLEMOT. *Using an exponential power model for Wyner-Ziv Coding*, in "IEEE Intl. Conf. On Acoustic and Signal Processing, ICASSP", États-Unis Dallas, 2010, http://hal.inria.fr/inria-00504516.

[45] O. OGUZ, C. HERZET, L. VANDENDORPE. *Low Complexity Iterative Detection in the Presence of Nuisance Parameters*, in "Proc. IEEE Int'l Conf. on Acoustics, Speech and Signal Processing (ICASSP)", États-Unis Dallas, 2010, http://hal.inria.fr/inria-00504513.

[46] J. TAQUET, C. LABIT. *Near-lossless and scalable compression for medical imaging using a new adaptive hierarchical oriented prediction*, in "IEEE Intl. Conference on Image Processing, ICIP'10", Sept. 2010, http://hal.inria.fr/inria-00538794.

[47] J. TAQUET, C. LABIT. *Optimized decomposition basis using Lanczos filters for lossless compression of biomedical images*, in "IEEE Intl. Workshop on MultiMedia Signal Processing, MMSP'10", Oct. 2010, http://hal.inria.fr/inria-00538797.

[48] J. TAQUET, C. LABIT. *Optimized decomposition basis using Lanczos filters for lossless compression of biomedical images*, in "IEEE Intl. Workshop on Multimedia Signal Processing (MMSP'10)", Saint-Malo, France, Oct. 2010.

[49] V. TOTO-ZARASOA, A. ROUMY, C. GUILLEMOT. *Hidden Markov model for Distributed Video Coding*, in "International Conference on Image Processing (ICIP)", Sept. 2010, http://hal.inria.fr/inria-00538792.

[50] V. TOTO-ZARASOA, A. ROUMY, C. GUILLEMOT. *Non-asymmetric Slepian-Wolf coding of non-uniform Bernoulli sources*, in "International Symposium on Turbo Codes (ISTC)", Sept. 2010, http://hal.inria.fr/inria-00538793.

[51] V. TOTO-ZARASOA, A. ROUMY, C. GUILLEMOT. *Non-uniform source modeling for Distributed Video Coding*, in "European Signal Processing Conference (EUSIPCO)", August 2010, http://hal.inria.fr/inria-00538790.

[52] M. TURKAN, C. GUILLEMOT. *Image prediction: Template matching vs. Sparse approximation*, in "Proc. IEEE Int'l Conf. on Image Processing (ICIP)", Hong-Kong, 2010, p. 2034-2037, http://hal.inria.fr/hal-00538834.

[53] F. XIE, T. FURON, C. FONTAINE. *Better security levels for "Broken Arrows"*, in "Proceedings of the SPIE : International Symposium on Electronic Imaging 2010 : Media Forensics and Security XII", États-Unis San Jose, SPIE, 2010, vol. 7541, http://hal.inria.fr/hal-00472982.

[54] J. ZEPEDA, C. GUILLEMOT, E. KIJAK. *The Iteration-Tuned Dictionary for Sparse Representations*, in "IEEE Intl. Workshop On MultiMedia Signal Processing, MMSP", Saint-Malo, France, Oct. 2010, http://hal.inria.fr/inria-00539076.

[55] J. ZEPEDA, E. KIJAK, C. GUILLEMOT. *Approximate nearest neighbors search using sparse approximations*, in "IEEE Intl. Conf. On Acoustic and Signal Processing, ICASSP", Dallas, États-Unis, March 2010, http://hal.inria.fr/inria-00504517.

National Peer-Reviewed Conference/Proceedings

[56] E. BOSC, V. JANTET, L. MORIN, M. PRESSIGOUT, C. GUILLEMOT. *Vidéo 3D: quel debit pour la profondeur?*, in "CORESA'10", Oct. 2010, http://hal.archives-ouvertes.fr/hal-00533843/fr/.

[57] V. JANTET, L. MORIN, C. GUILLEMOT. *Génération, compression et rendu de LDI*, in "CORESA'10", Oct. 2010, http://hal.inria.fr/hal-00530752_v1/.

[58] J. TAQUET, C. LABIT. *Compression sans perte et presque sans perte d'images médicales à l'aide d'un prédicteur hiérarchique orienté et adaptatif*, in "CORESA'10", Oct. 2010, http://hal.inria.fr/inria-00538798.

Scientific Books (or Scientific Book chapters)

[59] T. FURON. *Chapitre V: Les mesures techniques de protection - Une brève introduction*, in "Les technologies de l'information au service des droits : opportunités, défis, limites", D. LE MÉTAYER (editor), Cahiers du Centre de recherches informatique et droit, Bruylant, 2010, vol. 32, p. 85-90, http://hal.inria.fr/inria-00538241/en/.

Research Reports

[60] T.-T. DO, E. KIJAK, T. FURON, L. AMSALEG. *Understanding the Security and Robustness of SIFT*, INRIA, May 2010, RR-7280, http://hal.inria.fr/inria-00482502.

[61] P. DÉRIAN, P. HÉAS, C. HERZET, E. MEMIN. *Wavelet Expansion and High-order Regularization for Multiscale Fluid-motion Estimation*, INRIA, Jul 2010, RR-7348, http://hal.inria.fr/inria-00505780.

[62] E. FRANÇOIS, L. GUILLO. *TE12.2: Technicolor and INRIA report on AMP evaluation*, Joint Collaborative Team on Video Coding, Oct 2010, JCTVC-C286.

[63] E. FRANÇOIS, L. GUILLO, A. ICHIGAYA, H. YU. *TE12 : report on AMP evaluation*, Joint Collaborative Team on Video Coding, Oct 2010, JCTVC-C030.

[64] E. FRANÇOIS, D. THOREAU, L. GUILLO. *TE 3 subset2: report on Simplified Geometry Block Partitioning*, Joint Collaborative Team on Video Coding, Oct 2010, JCTVC-C031.

[65] C. GUILLEMOT, L. GUILLO, P. TANGI. *Intra Prediction by a linear combination of Template Matching predictors*, Joint Collaborative Team on Video Coding, Jul 2010, JCTVC-B078.

[66] G. SOURIMANT. *Depth maps estimation and use for 3DTV*, INRIA, Feb 2010, RT-0379, http://hal.inria.fr/inria-00458221.

[67] J. TAQUET, C. LABIT. *Une introduction à la compression d'images médicales volumiques*, INRIA, Jun 2010, RR-7324, http://hal.inria.fr/inria-00494306.

[68] J. TAQUET, C. LABIT. *Une introduction à la compression d'images médicales volumiques*, INRIA, June 2010, RR-7324, http://hal.inria.fr/inria-00494306.