



INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Project-Team TEMICS

*Traitement, Modélisation d'Images et
CommunicationS*

Rennes - Bretagne-Atlantique

THEME COG

Activity
R *eport*

2008

Table of contents

1. Team	1
2. Overall Objectives	1
3. Scientific Foundations	3
3.1. 3D scene modelling based on projective geometry	3
3.2. Frame expansions	4
3.3. Rate-distortion theory	5
3.4. Distributed source coding	5
3.5. Watermarking as a problem of communication with side information	6
4. Application Domains	6
4.1. Introduction	6
4.2. Compression with advanced functionalities	7
4.3. Multimedia communication	7
4.4. Copy protection, copyright enforcement and enriched content	7
5. Software	8
5.1. Software of the past four years	8
5.1.1. Libit	8
5.1.2. Soft VLC decoding library	8
5.1.3. PICS: Inter Layer Predictors for Scalable Coding	8
5.1.4. Oriented wavelet based image codec	8
5.1.5. WAVIX: Wavelet based video codec	8
5.1.6. 3D model-based video codec	9
5.1.7. Video communication platform	9
5.1.8. Broken Arrows Software	10
5.2. Software with significant evolutions in 2008	10
5.2.1. Distributed video codec	10
5.2.2. M3DPlayer: 3D video player	10
5.2.3. Watermarking platform	10
5.3. ADT PICOVIN	11
6. New Results	11
6.1. 3D representation and compression of video sequences	11
6.1.1. 3d modelling from video, GPS and GIF data for urban environments	11
6.1.2. Depth map extraction for multi-view sequences	12
6.1.3. 3d representation for multi-view video sequences	13
6.2. Sparse representation, compression and interaction with indexing	14
6.2.1. Sparse signal representation algorithms	14
6.2.2. Spatial and temporal prediction based on sparse signal representations	15
6.2.3. Quantization and coding of sparse vectors	15
6.2.4. Perceptual coding	15
6.2.5. Lossless coding for medical images	16
6.2.6. Feature extraction for joint compression and local description	17
6.3. Joint source-channel coding	17
6.3.1. Multiple description coding with side information	17
6.3.2. Joint source-channel coding and Slepian-Wolf coding with arithmetic codes	18
6.3.3. Parameter estimation and design of receivers robust to model uncertainties	19
6.3.4. Direction of arrival estimation of correlated waveforms	20
6.4. Distributed source and source-channel coding	20
6.4.1. Robust rate-adaptive non-asymmetric Slepian-Wolf coding	20
6.4.2. Distributed video compression	21
6.5. Image and video watermarking and security	21

6.5.1. Fingerprinting	21
6.5.2. Reliability	22
6.5.3. Watermarking Security	23
7. Contracts and Grants with Industry	23
7.1. Industrial contracts	23
7.1.1. CIFRE contract with France Telecom on the problem of 3D video representation	23
7.1.2. CIFRE contract with Thomson on the problem of spectral deconvolution for video compression and protection	23
7.1.3. Collaboration with Alcatel on robust video compression	24
7.2. National contracts	24
7.2.1. DGE/Region-FUTURIMAGE	24
7.2.2. RIAM-ESTIVALE	25
7.2.3. RIAM-MEDIEVALS	25
7.2.4. ACI Nebbiano	25
7.2.5. ANR-COHDEQ	26
7.2.6. ANR-TCHATER	26
7.2.7. Alibaba	26
7.2.8. ANR-ESSOR	26
7.2.9. ANR-ICOS-HD	27
7.3. European contracts	27
7.3.1. NoE-ECRYPT	27
7.3.2. FP6-IST NOE NEWCOM++	27
8. Other Grants and Activities	28
9. Dissemination	28
9.1. Patents	28
9.2. Invited talks and book chapters	28
9.3. Visits	29
9.4. Leadership within the scientific community	29
9.5. Exhibitions and demonstrations	30
9.6. Teaching	30
10. Bibliography	31

1. Team

Research Scientist

Christine Guillemot [Team Leader, DR, Inria, HdR]
Teddy Furon [CR, INRIA, until 31st Aug. 2008]
Cédric Herzet [CR, INRIA]
Claude Labit [DR, INRIA, 20%]
Gagan Rath [CR, INRIA, until 31st Mar. 2008]
Aline Roumy [CR, INRIA]
Caroline Fontaine [CR, CNRS, 80%]

Faculty Member

Jean-Jacques Fuchs [Professor, SPM, University of Rennes 1, HdR]
Luce Morin [Associate professor, Ifsic, University of Rennes 1, until 31st Aug. 2008, HdR]

Technical Staff

Laurent Guillo [IR, CNRS]

PhD Student

Gael Sourimant [INRIA grant, Fururimages project]
Ana Charpentier [INRIA grant, MEDIEVALS project, since 1st Apr. 2008]
Thomas Colleu [CIFRE contract with France Télécom]
Olivier Crave [INRIA grant, ANR-ESSOR]
Angélique Drémeau [INRIA grant]
Vincent Jantet [INRIA grant, Futurimages project, since 1st Sept. 2008]
Denis Kubasov [INRIA grant, IST-DISCOVER project]
Simon Malinowski [Ministry of research grant]
Aurélien Martin [CIFRE contract with Thomson]
Jonathan Taquet [Inria-Région ARED grant, since 15th Oct. 2008]
Velotiaray Toto-Zarasoia [Ministry of research grant, since Oct. 1st, 2007]
Mehmet Turkan [INRIA grant, CORDIS, since 1st Oct. 2008]
Fuxun Xie [INRIA grant, CORDIS]
Joaquin Zepeda [INRIA grant, ANR-ICOS-HD project, also with TexMex]

Post-Doctoral Fellow

Zhenzhong Chen [ERCIM grant, until 9th Sept. 2008]

Administrative Assistant

Huguette Béchu [TR, INRIA, shared with VISTA]

Other

Jérôme Allasia [INRIA, until 31st Aug. 2008]
Mathieu Desoubes [INRIA]
Gabrielle Ombrouk [INRIA, since 15th Sept. 2008]
Yann Serrand [INRIA, since 15th Sept. 2008]

2. Overall Objectives

2.1. Overall Objectives

The goal of the TEMICS project-team is the design and development of theoretical frameworks as well as 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 :

- *Analysis and modelling of video sequences.* The support of advanced interaction functionalities such as video content manipulation, navigation, 3DTV or free view-point visualization requires the development of video analysis and modelling algorithms. The TEMICS project-team focuses on the design of algorithms for 3D scene modelling from monocular, multi-view and multi-modal video sequences with optimum trade off between model reliability and description cost (rate).
- *Sparse representations, compression and interaction with indexing.*

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. Frame expansions and in particular 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. Anisotropic waveforms have shown to be promising for a range of applications and in particular for compact representations of still images. Sparse signal representations are also very powerful tools for compression but also for texture analysis and synthesis, for prediction and for inpainting. While sparse representations currently used in image coding are all based on the l_2 error metric and the associated ubiquitous PSNR quality measure, it is well-known that this metric is not really appropriate from a perceptual point of view. The TEMICS project-team investigates sparse representations and dedicated fast algorithms which, besides the l_1 norm on the weights that ensures sparseness, would minimize the reconstruction error with norms different from the l_2 norm as is systematically the case nowadays. Spatial and temporal prediction and coding techniques based on sparse representations are also studied.

There is a relation between sparse representation and clustering (i.e. vector quantization). In clustering, a set of descriptive vectors is learned and each sample is represented by one of these vectors, the one closest to it, usually in the l_2 distance measure. In contrast, in sparse representations, the signal is represented as a linear combination of several vectors. In a way, it is a generalization of the clustering problem. The transformed versions of the signal lie on a low-dimension manifold in the high-dimensional space spanned by all pixel values. The amenability of these representations for image texture description is also investigated.

- *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 a theoretical and practical framework setting the foundations for optimal design of image and video transmission systems 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 source 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.

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, and steganography. Depending on the application (copyright protection, traitor tracing, 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 (ACI NEBBIANO, RIAM ESTIVALE), ANR ESSOR, ANR ICOS-HD, ANR MEDIEVALS,) 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

Keywords: *3D reconstruction, camera models, computer vision, epipolar constraints, fundamental matrix, perspective projection, 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 :

$$\tilde{p}_i = P_i \tilde{x}, \quad (1)$$

where \tilde{x} is a 3D point with homogeneous coordinates $\tilde{x} = (x \ y \ z \ 1)^t$ in the scene reference frame \mathcal{R}_0 , and where $\tilde{p}_i = (X_i \ Y_i \ 1)^t$ 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 \tilde{x} , given \tilde{p}_i , i.e. to solve Eqn. (1) with respect to coordinates \tilde{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

Keywords: *Wavelet transforms, error correcting codes, filter banks, multiple description coding, oversampled 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 i th 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 $\tilde{\Phi}_F \equiv \{(F^h F)^{-1} \varphi_i\}_{i \in I}$. 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

Keywords: *OPTA limit (Optimum Performance Theoretically Attainable), Rate allocation, Rate-Distortion optimization, channel modelization, error correcting codes, joint source-channel coding multiple description coding, lossy coding, oversampled frame expansions.*

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

Keywords: *Slepian-Wolf coding, Wyner-Ziv coding, coding with side information, correlated sources.*

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

Keywords: *capacity, discrimination, information theory, side information, watermarking.*

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 $E(\cdot)$ 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 $S(\cdot)$ mapping a discrete message $m \in \mathcal{M}$ onto a signal in \mathbf{R}^N .

The decoding function $S^{-1}(\cdot)$ 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 $\{S(\cdot), S^{-1}(\cdot)\}$ 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:

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

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.

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 video compression, network adaptation (e.g., resilience to channel noise), or in terms of advanced functionalities such as navigation, content copy and copyright protection, or authentication.

4.2. Compression with advanced functionalities

Notwithstanding the already large number of solutions, compression remains a widely-sought capability especially for audiovisual communications over wired or wireless IP networks, often characterized by limited bandwidth. The advent of these delivery infrastructures has given momentum to extensive work aiming at optimized end-to-end QoS (Quality of Service). This encompasses low rate compression capability but also capability for adapting the compressed streams to varying network conditions. Scalable coding solutions making use of mesh-representations and/or spatio-temporal frame expansions are developed for that purpose. At the same time, emerging interactive audiovisual applications show a growing interest for 3-D scene navigation, for creating intermediate camera viewpoints, for integrating information of different nature, (e.g. in augmented and virtual reality applications). Interaction and navigation within the video content requires extracting appropriate models, such as regions, objects, 3-D models, mosaics, shots... The signal representation space used for compression should also be preferably amenable to signal feature and descriptor extraction for fast and easy data base access purposes.

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 paradigm sets the foundations for the design of efficient solutions to the above challenges. Distributed source coding is driven by a set of emerging applications such as wireless video (e.g. mobile cameras) and sensor networks. Such applications are placing additional constraints on compression solutions, such as limited power consumption due to limited handheld battery power. The traditional balance of complex encoder and simple decoder needs to be reversed.

4.4. Copy protection, copyright enforcement and enriched content

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 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. Software of the past four years

5.1.1. *Libit*

Participant: Laurent Guillo [contact person].

Libit is a C library 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 developpers 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 developed, 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. (URL: <http://libit.sourceforge.net>). The library is made available under the GNU Library General Public Licence.

5.1.2. *Soft VLC decoding library*

Participant: Simon Malinowski [contact person].

This library contains a set of robust decoding tools for variable length codes (VLC) and for quasi-arithmetic codes. It contains tools for soft decoding with reduced complexity with aggregated state models for both types of codes. It also includes soft decoding tools of punctured quasi-arithmetic codes with side information used for Slepian-Wolf coding of correlated sources. This software requires the Libit library (see above) and the GMP (GNU Multiple Precision) library.

5.1.3. *PICS: Inter Layer Predictors for Scalable Coding*

Participant: Gagan Rath [contact person].

This library contains a set of tools for inter-layer prediction in a scalable video codec. In particular, it contains a tool for improved spatial prediction of the higher resolution layers based on the lower resolution layer. It also contains tools for orthogonal transforms of the enhancement layers, which were derived from the Laplacian pyramid structure in the scalable video codec. This software has been registered at the APP (Agence de Protection des Programmes) under the number IDDN.FR.01.140018.000.S.0.2007.000.21000.

5.1.4. *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. Extensions for lossless compression are being studied. 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.

5.1.5. *WAVIX: Wavelet based video codec*

Participant: Laurent Guillo [contact person].

The video codec called WAVIX (Wavelet based Video Coder with Scalability) is a low rate fine grain scalable video codec based on a motion compensated t+2D wavelet analysis. Wavix supports three forms of scalability: temporal via motion-compensated temporal wavelet transforms, spatial scalability enabled by a spatial wavelet transforms 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. A first version of the codec has been registered at the Agency for the Protection of Programmes (APP) under the number IDDN.FR.001.160015.000.S.P.2003.000.20100. Robust variable length codes decoding tools have been integrated in the decoder. Redundant temporal motion-compensated filtering has been added in order to increase the codec resilience to packet losses. The codec has been used for experiments in the RNRT-COSINUS project. This codec will be used to validate algorithm developments to be performed in collaboration with Alcatel (see Section 7.1.3).

5.1.6. 3D model-based video codec

Participant: Luce Morin [contact person].

The Temics project-team has in the past years developed a software for 3D modelling of video sequences which allows interactive navigation and view point modification during visualization on a terminal. From a video sequence of a static scene viewed by a monocular moving camera, this software 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 has been registered at the Agency for the Protection of Programmes (APP) under the number IDDN.FR.001.130017.000S.P.2003.000.41200.

5.1.7. Video communication platform

Participant: Laurent Guillo [contact person].

The TEMICS project-team pursues the development of a video communication platform, called VISIUM. This platform provides a test bed allowing the study and the assessment, in a realistic way, of joint source channel coding, video modelling or video coding algorithms. It is composed of a video streaming server "Protée", of a network emulator based on NistNet and of a streaming client "Pharos":

- The streaming server allows for the streaming of different types of content: video streams encoded with the WAVIX coder as well as streams encoded with the 3D-model based coder. The video streaming server is able to take into account information from the receiver about the perceived quality. This information is used by the server to estimate the bandwidth available and the protection required against bit errors or packet losses. The server can also take advantage of scalable video streams representations to regulate the sending rate.
- The streaming client, "Pharos", built upon a preliminary version called "Criqs", can interact with the server by executing scripts of RSTP commands. They can gather specific commands such as "play", "forward", "rewind", "pause", establish RTP/RTCP connections with the server and compute QoS information (jitter, packet loss rate,...). The client enables the plug-in of different players and decoders (video and 3D).

The server "Protée" and the client "Criqs" are respectively registered at the Agency for the Protection of Programmes (APP) under the number IDDN.FR.001.320004.000.S.P.2006.000.10200 and the number IDDN.FR.001.320005.000.S.P.2006.000.10800. This platform makes use of two libraries integrated in both the server and the client. The first one "Wull6" is an extension to IPv6 of the "Wull" library implementing the transport protocol UDP-Lite base on the RFC 3828. The second one "bRTP" implements a subset of the RTP/RTCP protocols based on the RFC3550. These two libraries are respectively registered at the Agency for the Protection of Programmes (APP) under the number IDDN.FR.001.270018.001.S.A.2004.000.10200 and the number IDDN.FR.001.320003.000.S.P.2006.000.10200.

5.1.8. *Broken Arrows Software*

Participant: Teddy Furon [contact person].

In collaboration with Patrick Bas (CNRS - Gipsa-lab - Grenoble), we have developed and benchmarked a new watermarking technique which has been the main element of the international challenge BOWS-2 (Break Our Watermarking System - 2nd Edition). This challenge is now over and the technique has been disclosed. 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 URL: bows2.gipsa-lag.inpg.fr).

5.2. Software with significant evolutions in 2008

5.2.1. *Distributed video codec*

Participant: Denis Kubasov [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 and test files, 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>

5.2.2. *M3DPlayer: 3D video player*

Participants: Jérôme Allasia, Laurent Guillo, Luce Morin [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 (see Section 5.1.6). The 3D player integrated as a component to the VISIUM communication platform (see Section 5.1.7) allows remote access to a 3D sequence. In this context the video streaming server "Protée" has been modified to support 3D streams. In the near future, we would like to introduce scalability capabilities of the coded representation of the 3D models and texture, in order to enable dynamic adaptation of the transmitted bitstream to both network and terminal capabilities. The player is registered at the APP (Agence de Protection des Programmes) under the number IDDN.FR.001.090023.000.S.P.2008.000.21000.

5.2.3. *Watermarking platform*

Participants: Mathieu Desoubreaux, 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 (see above) and the Chimark2 software. 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.

This software platform has significantly evolved in 2008. Many technical changes occurred: the use of Virtual Dub has been given up since it is not portable and it is distributed under a license which would have contaminated our source code. The platform is now based on the openCV computer vision library and the Wxwidget cross-platform GUI library. It is now a self-contained demonstrator, not just a plug-in to be used in a video edition software. The INRIA forge stores and manages all the versions of the source code. All the routines have been implemented and the demonstrator is now running in version 1.0. The latest results on watermarking and new accusation function providing more robustness to collusion attacks have been integrated. The estimator of the rare event probability that has been patented is also integrated in the demonstrator: we are able to trace back the most likely guilty users and to evaluate the risk of accusing innocent people with excellent accuracy.

5.3. ADT PICOVIN

Participants: Laurent Guillo [contact person], Gabrielle Ombrouck, Yann Serrand.

The ADT Picovin is a technological development action, which works closely with the project-team TEMICS. This new development structure gives its support to the project-team to carry out a new and efficient video codec. This support is twofold. The project-team TEMICS provides the ADT with innovative algorithms that are first adapted and integrated in the current codec. Then the resulting codec is evaluated thanks to a large base of videos in order to measure how significant these new contributions are and to detect as soon as possible unexpected behaviour. Hence, step by step, the emerging new codec will perform better and better and should be able to meet requirements of a possible call for proposition of existing organizations for standardization. Throughout this process, it will be compared with the still evolving competing codecs, for instance JM, KTA and DIRAC.

The very first activities of the ADT are focussed on the integration of new algorithms related to texture prediction and on the development of the integration and evaluation platform. This new development structure started in October 2008 and will last three years. Two 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

Keywords: *2D and 3D meshes, 3D modelling, motion modelling, projective geometry, tracking, triangulation.*

6.1.1. 3d modelling from video, GPS and GIF data for urban environments

Participants: Vincent Jantet, Luce Morin, Gael Sourimant.

This work is done in collaboration with the BUNRAKU project-team (Kadi Bouatouch). From a video sequence of a static scene viewed by a monocular moving camera, we have studied methods to automatically construct a representation of a video as a stream of textured 3D models. 3D models are extracted using stereo-vision and dense matching maps estimation techniques. However, this approach presents some limitations, in particular the presence of drift in the location and orientation of the 3D models. Drift is due to accumulation of uncertainties in the 3D estimation process along time. This is a strong limitation for virtual reality applications such as insertion of synthetic objects in natural environments. On the other hand, GIS (Geographic Information Systems) provide a complete, geo-referenced modelling of city buildings. However, they are far less realistic than video captures, due to artificial textures and lack of geometric details. Video and GIS data thus give complementary information: video provides photorealism, geometrical details, precision in the fronto-parallel axes; GIS provides a "clean" and complete geometry of the scene, structured into individual buildings. We have also added a GPS acquisition synchronized with video acquisition.

In order to combine these three types of data, the first step is to register the data in the same coordinate system. GPS data only provide a rough approximation of the camera position (but not its orientation) with regards to the GIS database, so GIS/Video registration can be seen as a two-steps procedure. First, a new fully automatic registration procedure is applied on the first video frame using vision-based algorithms and context analysis. In a second step, the pose is tracked for each frame, using a visual servoing approach based on the registration of 2D interest points extracted from the images, and 3D points that correspond to these feature points projection onto the building model provided by the GIS database. As a result, we get the pose of the camera, e.g. its position and orientation with regards to the GIS database, for each frame of the video sequence. These poses permit to project accurately the 3D models onto the images, and textures of the building façades are extracted and cleaned in a pixel-wise fusion procedure.

In 2008, we focused on the supervised automatic initialization part of the registration procedure. The method we developed in order to register the first image with the corresponding 3D model can be decomposed into two stages:

- *Rough pose estimation:* The camera approximate motion is computed using state of the art vision algorithms. The estimated translation is related to the GPS displacement so as to get an initial camera orientation for the first image, which permits to get the same reprojected 3D models than those visible in the video.
- *Precise pose estimation:* The rough camera pose computed in the previous step is used to detect and match 3D lines of the model with 2D lines extracted from the images. The matching procedure is constrained by the image context (ground, façades, sky) derived from image segmentation, and robust 2D-3D lines registration is ensured thanks to the RANSAC algorithm. Once lines correspondences are given, accurate pose is computed with visual servoing.

The central point of this pose computation is to select a *key image* in the video that permits to get the same reprojected 3D models than those visible in the video. This is a difficult problem since GPS measures are often poorly accurate, thus leading to a false rough pose estimation and then an unsolvable precise pose estimation. Further frames in the video are tested one after the other, and when the algorithm selects a frame with a supposed good pose, the user is asked to accept or reject the result: that is the supervised part. If the frame is rejected, another one is searched for further in the video.

Several criteria are used by our algorithm to determine whether a frame should be a good one to compute the pose for the first frame. First of all, we test the amount of epipolar residual induced by the estimated fundamental matrix F , using the correspondences tagged as valid by the RANSAC-based algorithm which computed F . If this residual is too high, the epipolar geometry is considered as too badly estimated to compute the rough pose. From a rough pose, the current frame is kept if we can extract enough projected 3D lines to compute the pose using a visual virtual servoing scheme. This constitutes the second criterion. The last criterion is the result of the robust pose computation. Using 2D lines extracted from the images and 3D projected lines computed from the GIS model and the rough pose, a RANSAC-based algorithm iterates over the possible line correspondences in order to discard outliers and find the right ones. The success of this phase can be measured by the number of RANSAC iterations performed to find a result. Indeed, this number decreases as the probability to find a good solution increases. In our tests, the proposed frames that passed these three tests were only rejected by the user when the GPS measures were completely false.

6.1.2. Depth map extraction for multi-view sequences

Participants: Claude Labit, Luce Morin, Gael Sourimant.

3DTV and Free Viewpoint Video (FVV) are now emerging video formats expected by end-users, because of their navigation capabilities and visualization enhancement, especially when displayed on stereoscopic or auto-stereoscopic screens. Though 3D information can be easily computed for synthetic movies, at present time no professional nor consumer video cameras capable of capturing the 3D structure of the acquired scene are available (except of course for Z-camera prototypes). As a consequence, 3D information representing real content have to be estimated from acquired images only, using computer vision-based algorithms.

In this study - supported by the Futurim@ges project - we focus both on extending our 3D model-based video codec (see Section 5.1.6) to multi-view image sequences, and on post-processing this estimated 3D information to provide attractive videos purified from usual 3D artifacts (badly modelled occlusions, texture stretching, etc.). For multi-view image sequences, the scene structure can be estimated not only over time, but also over space. At time t , the N acquired views can be used to compute the 3D structure of the scene more easily since the camera bank calibration is supposed to be known. We are now designing algorithms to integrate this space-based reconstruction to the classical time-based reconstruction, in order to produce higher quality 3D models. Once the 3D structure of the scene has been estimated (and stored as a pool of depth maps), it is post-processed using our Java-based software *M3dEncoder2*. One goal in this post-processing steps is to generate auto-stereoscopic-ready videos, so they could be displayed on available auto-stereoscopic screens (at present time we aim the Newsight™ and the Philips™ 3D screens). We identified four successive processing levels of the input videos plus depth maps:

1. *Pre-processing of input depth maps*: our codec generate depth maps that can be noisy on uniform zones, and rather smooth over discontinuities. In order to smooth the uniform zones and reinforce the discontinuities, we apply a statistical operator based on median absolute deviation on each point of the depth maps.
2. *3D representation*: the depth maps are then converted into 3D models. In our original 3D compression framework, they were modelled as uniform meshes with full connectivity. Here we propose to use a representation either based on a Delaunay triangulation of the map, or on a Quadtree representation. In both cases we use a thresholding scheme to detect depths discontinuities, which are used in the 3D representation. We are now exploring global representations of the scene, by integrating the complete depth maps information into a single 3D model.
3. *3D display*: depending on the chosen 3D representation, several display-specific treatments can be applied. For instance, in the case where we have one 3D model per GOP, we can blend the rendering of successive meshes as a normalized weighted sum, so as to have smooth transitions between the successive viewed models over time.
4. *Display post-processing*: in the case of 2D plus Depth video output (used for auto-stereoscopic display on Philips screens), we apply a last bilateral filtering stage on the depth part of the images. This results in noise reduction on low depth variation zones, but still depth discontinuities preservation.

6.1.3. 3d representation for multi-view video sequences

Participants: Thomas Colleu, Claude Labit, Luce Morin, Gael Sourimant.

The 3D experience consists either in 3D relief rendering called 3DTV, or interactive navigation inside the content called FTV (Free viewpoint TV). For the design of any 3D video system, the choice of a certain representation 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. During 2008, we studied existing 3D video representations and tested new representations that can provide 3DTV and FTV functionalities as well as efficient compression capabilities. The input of the system consists in multiple color videos and depth maps. An important point is to reduce the redundancies between the views so as to represent only new data.

Many representations were studied. They differ from the amount and type of geometry and texture that are used. The Image-based representation does not use geometry information at all. It provides a potentially high quality of virtual view synthesis avoiding any 3D scene reconstruction. However, this benefit has to be paid by dense sampling of the real world with a sufficiently large number of natural cameras. The depth image-based representation is more flexible than the image-based one. It includes geometrical information in the form of depth maps. The disadvantage is the high redundancy of the transmitted data when multiple views of the scene are captured. The surface-based representation reconstructs a global model of a scene or an object, therefore all redundancies are eliminated compared with the depth image-based representation and the compactness of the representation does not depend on the number of views. To this end, reliable computation of 3D scene models

remains difficult and is often limited to foreground objects only. The volumetric representation offers linear access time thanks to the octree structure, but voxels are rendered as cubes, resulting in poor visualization when the rendering viewpoint is close to the surface. The impostor-based representation provides a simplified model of the scene and is able to combine image quality and rendering speed. As for the textures, the multi-texturing approach is very popular in many contributions since the quality of the rendered views is much better than using a single texture. However, this solution requires to store all the original images and is therefore very memory expensive.

The representation that we tested is based on a polygon soup. This is a set of 3D polygons without connectivity. First, every depth maps were divided into a multi-resolution quad-tree structure and a polygon was created for every quad. This quad-tree structure allows minimizing the number of polygons while maintaining geometric accuracy. Second, a reference image was chosen as the starting point for the polygon soup, and polygons from the occluded areas at this point of view were added to the representation thanks to the other available point of views. As a result a global model without redundancies is obtained. Concerning the image quality, the main problems encountered come from errors in the depth maps and pixel "ghosting" artifacts at depth discontinuities. The goal is to eliminate these problems in the near future.

6.2. Sparse representation, compression and interaction with indexing

Keywords: *compression, frame expansions, rate-distortion theory, scalability, wavelet transforms.*

6.2.1. Sparse signal representation algorithms

Participants: Jean-Jacques Fuchs, Christine Guillemot, Gagan Rath.

Sparse representations has become an important topic with numerous applications in signal or image processing. It has now evolved in part toward the so-called Compressed Sensing area which has potentially many applications. 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 2008, in order to further widen its applicability and open potentially new domains of applications we have continued to evaluate the recovery conditions developed fast algorithms for new criterion. Instead of the now ubiquitous criterion

$$\min_x \frac{1}{2} \|Ax - b\|_2^2 + h\|x\|_1$$

which, by the way we introduced more than 10 years ago, we consider the cases where, in this criterion, the ℓ_2 -norm is replaced by an ℓ_1 or ℓ_∞ norm and develop the recovery conditions and fast algorithms for them [31], [34] While it appears that the corresponding representations are no longer parsimonious, they have different properties that might be of interest.

We have also introduced the concept of a complementary matching pursuit (CMP). The algorithm is similar to the classical matching pursuit (MP), but performs complementary actions. Instead of selecting one atom to be included in the sparse approximation, it selects $(N - 1)$ atoms to be excluded from the approximation at each iteration. Though these two actions seem apparently the same, they are actually performed in two different spaces. On a conceptual level, the MP searches for "the solution vector among sparse vectors" whereas the CMP searches for "the sparse vector among the solution vectors". As a consequence of the complementary action, the CMP does not minimize the residual error at each iteration, however it may converge faster yielding sparser solution vectors than matching pursuit. We have shown that when the dictionary is a tight frame, the CMP is equivalent to the MP, but in the general case they are not equivalent. We have also studied orthogonal extensions of the CMP and we have shown that they perform the complementary actions to those of their classical matching pursuit counterparts.

The above sparse coding problem implicitly assumes that the dictionary matrix A is known beforehand. This is, however, not true in many of the applications cited above. The sparse coding problem therefore follows the dictionary training problem where the objective is to design the optimal dictionary for sparsely encoding a set of observed signals. This problem is also known as dictionary training in sparse approximation literature. Among the many training algorithms, the recently proposed K-SVD method has been shown to be a simple, efficient, and optimal training algorithm. The algorithm adopts a procedure similar to the K-means algorithm for codebook design in vector quantization, but adapted to the sparse coding problem. The algorithm, however, suffers from a sub-optimality because of the sequential updating of the atoms. Simple modifications can improve the performance of the algorithm in terms of the convergence speed and/or the optimality of the solution atoms.

6.2.2. *Spatial and temporal prediction based on sparse signal representations*

Participants: Jean-Jacques Fuchs, Christine Guillemot, Aurélie Martin, Mehmet Turkan.

Closed-loop spatial prediction has been widely used for image compression in transform or spatial domains. The prediction is often done by simply “propagating” the pixel values along the specified direction. This approach fails in complex textured areas. In 2008, we have addressed the problem of spatial and temporal video prediction along two directions:

- We have on one hand pursued the study of spatial image prediction based on matching pursuit algorithms. The problem is looked at as a problem of texture synthesis (or inpainting) from noisy data 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 (dictionary). In the methods developed, the sparse signal approximation is 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 taking the previously decoded neighborhood as the known support. The analogy with spectral deconvolution method has been studied.
- One critical problem of sparse representations is the design of the dictionary. Early 2008, locally adaptive dictionary construction methods have been designed both for spatial and temporal prediction. Significant spatial prediction gains have been shown compared to static DCT or DFT dictionaries. Similarly, high temporal prediction gains have been obtained compared to classical block matching techniques. These methods are currently being validated in a complete video coder/decoder in the context of the ADT Picovin (see Section 5.3).
- Locally adaptive dictionary construction must account for the presence of discontinuities, of edges, and must account for local texture characteristics. Local texture analysis and classification techniques are currently being developed and assessed.

6.2.3. *Quantization and coding of sparse vectors*

Participants: Angélique Drémeau, Jean-Jacques Fuchs, Christine Guillemot, Cédric Herzet.

Sparse approximation methods aim at finding representations of a signal with a small number of components taken from an overcomplete dictionary of elementary functions. They may thus be naturally well-suited for signal compression. However, the use of sparse approximations in compression raises a number of issues related to the quantization and the entropy coding of the sparse solution vector, which may be of large dimension. Methods for quantizing the sparse vectors and for entropy coding the index of the non-zero components are being investigated. Their rate-distortion performance is being assessed. The sparsity of the solution vector, and the overall rate-distortion performance of the representation, depends on how well the dictionary is tailored to the local signal characteristics. The problem of dictionary learning with the ultimate goal of optimized rate-distortion performance is addressed.

6.2.4. *Perceptual coding*

Participants: Zhenzhong Chen, Christine Guillemot.

In addition to the redundancy of the visual signal statistical correlation, the perceptual redundancy should be exploited to provide perceptual-friendly visual signal representation. Due to the spatial and temporal masking effects, the human visual system has the limitation on the perceptibility of certain levels of noise. Since the human visual system is space-invariant where the fovea has the highest density of sensor cells on the retina, the visual acuity decreases with increased eccentricity relative to the fovea. We have investigated a foveated just-noticeable-distortion (JND) model. In contrast to traditional JND methods which exploit the visibility of the minimally perceptible distortion but assume the visual acuity to be consistent over the image, a foveation model is incorporated in the spatial and temporal JND models. The foveated JND model is developed by combining the spatial JND as a function of luminance contrast and spatial masking effect, the temporal JND to model the temporal masking effect, and a foveation model to describe the relationship between the visibility threshold and eccentricity relative to the fovea. Associated with the proposed foveated JND model, more imperceptible distortion can be tolerated in the contaminated image.

The foveated JND model has been used for H.264/AVC video coding. Bit allocation and rate-distortion optimization are performed according to the foveated JND profile. The regions with higher visibility thresholds are coded with larger quantizers since these regions can tolerate higher distortion. The saved bit rate can be used to improve the quality of the regions which cannot tolerate high distortion. Therefore, the subjective quality of the whole image is improved. The performance of the foveated JND model has been assessed with subjective tests following the corresponding protocols in Rec. ITU-R BT.500.

6.2.5. Lossless coding for medical images

Participants: Claude Labit, Jonathan Taquet.

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. Usual techniques, such as lossless JPEG, LS-JPEG or lossless JPEG2000, are currently proposed and have been compared in a preliminary study (with ETIAM as industrial partner). The first obtained results show the opportunity to launch a new prospective research study in order to

- increase the present performances of traditional lossless compression schemes which are quite low (the average compression ratio of 3.8:1 obtained over a large medical imaging database is usually mentioned in tutorial books);
- extend these techniques, which are essentially based on block decomposition and spatio-frequency transforms, to oriented wavelet approaches (such as oriented wavelet transforms and quincunx sampling *Lifting algorithm*) and their adaptation to telemedicine and medical imaging storage applications.

Afterwards, during the project (J. Taquet's thesis), we propose to take into account coding algorithms commonly used in multimedia domain and so, adapt these techniques formerly developed in the lossy compression framework:

- explore, taking out the too restrictive constraint of exact digital reconstruction, *almost* lossless compression schemes i.e compression techniques which generate no irreversible degradations for the medical diagnosis; these techniques have to be selectively evaluated with respect to the medical imaging sources processed;
- propose new functionalities such as Region-Of-Interest analysis and coding (ROI-Based compression) which enable a simultaneous optimization of several rate-distortion functions adapted to the a-priori defined ROIs.

This Ph-D thesis, partially supported by a research grant of Bretagne Council, will take also benefit of the IHE-Europe technical coordination (« Integrating the Health Care Enterprise ») hosted at Irisa/INRIA Rennes Bretagne-Atlantique research center and of the presence, at Rennes, of the industrial partner ETIAM, SME

leader in Europe developing innovative tools for multimedia connectivity and medical imaging communication.

6.2.6. Feature extraction for joint compression and local description

Participants: Christine Guillemot, Joaquin Zepeda.

The objective of the study initiated in 2007, in collaboration with Ewa Kijak from TEXMEX, is to design signal representation and approximation methods amenable to both image compression (that is with sparseness properties) and description. During the last two decades, image representations obtained with various transforms, e.g., Laplacian pyramid, separable wavelet transforms, curvelets and bandlets have been considered for compression and de-noising applications. Yet, these critically-sampled transforms do not allow the extraction of low level signal features (points, edges, ridges, blobs) or of local descriptors. Feature extraction requires the image representation to be covariant under a set of admissible transformations. The Gaussian scale space is thus often used for description, however it is not amenable to compression. One robust (non-sparse) descriptor based on the Gaussian scale space is the SIFT descriptor. A recent approach referred to as *Video Google* tackles the high dimensionality problem of this descriptor by forming a single sparse descriptor obtained from multiple input non-sparse SIFT descriptors. The approach consists in vector quantizing each SIFT descriptor and then taking a (weighted) histogram of codeword indices. It allows using the principles of *inverted files*. Inverted file indices provide a solution to the indexation of high dimensional data (specifically, textual documents) by representing the data as sparse vectors. Document similarity calculations are thus carried out efficiently using the scalar products between these sparse vectors. We have developed a related approach that applies a pursuit-based sparse decomposition to each SIFT descriptor to obtain a sparse vector for each input SIFT descriptor. The descriptors retains the local characteristic of the input descriptors rather than forming a single global descriptor, while still enabling the use of inverted file type indices.

Video Google's synthesis of a sparse SIFT vector begs the question of whether one can come up with a descriptor that is sparse by design, thus avoiding the performance loss due to the sparse synthesis while retaining the search benefits related to sparsity. Our aim is to adapt existing work in sparse decomposition (designed with compression and prediction in mind) to the construction of image descriptors displaying covariance to the set of admissible transformations. In this context we have identified three problems that are currently addressed: 1) dictionary design, 2) atom selection method and 3) descriptor comparison method. Regarding dictionary design, the set of atoms comprising the dictionary should be closed under the set of admissible transformations. For atom selection a new non-iterative method has been developed: let $\underline{a}(\underline{t})$ denote a (continuous) dictionary formed by transforming a generating function $\underline{a}(\cdot)$ using transformation parameters \underline{t} . We propose to select atoms corresponding to transformation parameters $\{\underline{t}_o\}$ that locally maximize (over \underline{t}) the correlation between the dictionary function $\underline{a}(\underline{t})$ and the considered image patch. Such a method yields atoms that vary predictably with the transformations from the set of admissible image transformations. These predictable variations are at the root of the third problem to address, as descriptor comparison methods must take these variations into account when calculating descriptor similarities. One current approach considered of accounting for these variations in atom selection is to compare atoms using a distance defined in the transformation space, and to include this atom similarity measure in the descriptor distance expression.

6.3. Joint source-channel coding

Keywords: *error and erasure resilient coding and decoding, generalized likelihood ratio test, information theory, maximum likelihood estimation, multiple description coding, robust detection, stochastic modelling.*

6.3.1. Multiple description coding with side information

Participants: Olivier Crave, Christine Guillemot.

This study is carried out in collaboration with ENST-Paris (Béatrice Pesquet-Popescu). Multiple description coding has been introduced as a generalization of source coding subject to a fidelity criterion for communication systems that use diversity to overcome channel impairments. Distributed source coding is related to the problem of separate encoding and joint decoding of correlated sources. This paradigm naturally imparts resilience to transmission noise. The duality between the two problems, that is multiple description coding (MDC) and distributed source coding (DSC), is being explored in order to design loss resilient video compression solutions.

In 2007, two-description coding schemes based on overcomplete temporal signal expansions and systematic lossy Wyner-Ziv coding of a subset of the information have been designed. Systematic lossy Wyner-Ziv coding can be seen as an error control system which can be used as an alternative to Automatic Repeat reQuest (ARQ) or Forward Error Correction (FEC). These techniques result in satisfactory RD performances at the side decoders, i.e. in presence of channel impairments. However, when the two descriptions are received, the Wyner-Ziv data is redundant and does not contribute to improving the quality of the reconstructed signal.

To address this limitation, in 2008, we have investigated the problem of multiple description coding with side information. The achievable rate-distortion region for this problem has been recently established in the literature, when common side information about a correlated random process is known to each multiple description decoder. A lossy source coding algorithm based on multiple description coding with side information at the receiver has been designed. It builds upon both multiple description coding principles and Slepian-Wolf (SW) coding principles. The input source is first quantized with a multiple description scalar quantizer (MDSQ) which introduces redundancy or correlation in the transmitted streams in order to take advantage of path diversity. The resulting sequences of indexes are Slepian-Wolf encoded, that is separately encoded and jointly decoded. While the first step (MDSQ) plays the role of a channel code, the second one (SW coding) plays the role of a source code, compressing the sequences of quantized indexes. In a second step, the cross-decoding (and iterative) of the two descriptions is proposed. This allows us to account for both the correlation with the side information as well as the correlation between the two descriptions, and thus reduce the bit rate needed for the central decoder. Slepian-wolf coding of both descriptions thus allows decreasing the bit rate while preserving the robustness inherent to MDC. This scheme has first been validated on theoretical Gaussian sources and then integrated in a wavelet-based video coding algorithm. The approach finds applications in video streaming over peer-to-peer networks with cooperative receivers, or for robust video transmission with light-weight encoding devices.

6.3.2. Joint source-channel coding and Slepian-Wolf coding with arithmetic codes

Participants: Christine Guillemot, Simon Malinowski.

In the TEMICS project-team, we have introduced a new set of state models to be used in soft-decision (or trellis) decoding of variable length codes and quasi-arithmetic codes. The approach consists in aggregating states of the bit/symbol trellis which are distant of T instants of the symbol clock. The values of this parameter allow us to trade complexity against estimation accuracy. The state aggregation leads to close-to-optimum estimations with significantly reduced complexity. This model is well suited for the introduction and exploitation of *a priori* (or side information) in the transmitted bitstream in order to favor the selection of synchronous and correct paths in the soft decoding process. This can then be seen as a joint source-channel coding strategy. In comparison with other methods, such as markers which help the resynchronization of the decoding process, the strategy proposed does not lead to any modification of the compressed bit-stream. The approach turns out to outperform widely used techniques, such as the popular approach based on the introduction of a forbidden symbol to quasi-arithmetic codes. This joint source-channel coding technique is currently being assessed against turbo codes. Turbo codes have indeed been also considered for source compression. The approach is then known as turbo compression.

Arithmetic codes can also be used for distributed source coding or Slepian-Wolf coding. Most Slepian-Wolf coding solutions are based on channel coding principles assuming memoryless sources. However, in practice the signals considered have memory. Building upon our results on robust and soft decoding of arithmetic and quasi-arithmetic codes, we have investigated two new Slepian-Wolf coding strategies which aim at

capturing the source memory in addition to the correlation between the sources. The first approach is based on a puncturing mechanism which then further compresses the stream but introduces some uncertainty in the decoding process. This can be regarded as transmitting the corresponding bitstream over an erasure channel. A second approach, based on overlapped quasi-arithmetic codes, has been developed in collaboration with the University Polytechnic of Catalonia (X. Artigas, L. Torres) in the context of the European IST-Discover project. The resulting codes are not uniquely decodable. The side information, which is a noisy version of the emitted symbol sequence, available to the decoder is used to remove the ambiguity induced by the puncturing or the overlapping technique. Turbo coding and decoding structures based on the punctured and overlapped quasi-arithmetic codes have been studied in order to reduce the gap with classical turbo-codes.

The synchronisation recovery properties of QA codes when used with an erasure channel have also been studied. In particular, we have proposed a technique to compute the average number of symbols on which an erased bit propagates and the probability mass function (pmf) of the random variable corresponding to the difference between the number of encoded and decoded symbols (called ΔS). These quantities are computed by calculating gain polynomials on a state diagram. It is then possible to estimate these two quantities when the bit-stream is sent over an erasure channel characterized by an erasure probability (and not only for one erased bit). The knowledge of the pmf of ΔS allows us to find out which states of the QA code are less sensitive to puncturing. In addition, if a side information is available (like for example in the asymmetric Slepian-Wolf problem), this information can be taken into account in the computation of the pmf of ΔS . This can sometimes modify the sensitivity of a state with respect to erasure. The puncturing can hence be adapted to the resynchronisation properties of QA codes when used with an erasure channel, leading to codes that are more robust to puncturing.

6.3.3. *Parameter estimation and design of receivers robust to model uncertainties*

Participants: Cedric Herzet, Aline Roumy.

Many source and channel codes used in practical communication systems are based on the assumption that the source and channel statistical models are perfectly known to the receiver. Unfortunately, this ideal assumption is rarely satisfied in practice and taking the model uncertainties into account in the receiver is of utmost importance to make reliable decisions on the transmitted data. We have therefore developed generic low-complexity tools in order to design robust receivers and study their performance.

Our work is based on the family of receivers which jointly make an estimate of the model parameters and the transmitted data. We first considered the implementation of such receivers based on the well-known expectation-maximization (EM) algorithm. In this context, we developed a low complexity procedure which enables to easily predict the average speed of convergence of the receiver to its fixed points, see [19]. This work has been carried out in collaboration with V. Ramon (UCL, Belgium), L. Vandendorpe (UCL, Belgium) and A. Renaux (ENS Cachan, France). Depending on the system parameters, the EM-based receivers can sometimes exhibit a very low speed of convergence. In such cases, the large number of iterations required to reach a fixed point can lead to an unacceptable complexity overhead. Therefore, we developed and studied different methodologies to limit this increase of complexity. First, we are currently studying the properties of a modified version of the EM algorithm which is based on the rescheduling of the messages in a factor graph. Secondly, we proposed a novel general procedure for the design of robust receivers. The proposed approach is based on the iterative maximization of an approximated likelihood function and exhibits a faster speed of convergence than the EM algorithm in the scenarios we have considered, see [36].

Parameter estimation is usually performed with the help of a training sequence. Intuitively the longer the training sequence, the better the performance but also the lower the throughput of the transmission. Therefore, a tradeoff exists and in 2008 we studied the optimal length of a training sequence. For this problem, we derived two criteria: an effective signal-to-noise ratio and an effective channel capacity for the training-based transmission scheme. The receiver considered is a turbo-equalizer composed of a maximum a posteriori (MAP) equalizer (to combat the effect of the multipath in the channel), a MAP decoder (to deal with the additive noise of the channel) and a decision-directed channel estimator. Based on our previous contribution which gives a closed form expression of the performance of such a receiver, we showed that the optimal training sequence

according to the two criteria is the shortest possible [37], [20]. This work has been carried out in collaboration with N. Sellami, I. Hadj Kacem (ENIS, Sfax, Tunisia) and I. Fijalkow (ENSEA, Cergy, France), in the context of a CNRS-DGRSRT project.

6.3.4. Direction of arrival estimation of correlated waveforms

Participant: Jean-Jacques Fuchs.

The research is performed in an underwater-acoustics domain but it is also of interest in some communication systems, where the information is carried by waveforms that undergo multi-path propagation. In that case it may be important to locate precisely the direction of arrivals of the different paths carrying the same information in order to adequately combine them to enhance the signal to noise ratio.

Note that spatial correlation is induced by temporal correlation between the signals carried by the waveforms. Typically a temporally correlated signal emitted by a source arrives at the array through different paths, with different delays and the temporal correlation thus induces spatial correlation.

The difficulty to locate such paths comes from the fact that in case of correlation the rank of the covariance matrix of the so-called snapshots becomes difficult to detect and that in case of full correlation (coherency) the rank of this matrix is no longer equal to the number of impinging paths. This makes all the sub-space based DOA estimation schemes, such as MUSIC, inapplicable.

We propose to use the Global Matched filter (GMF) to overcome this difficulty. Instead of applying it, as usual, to a set of beams, we propose [33] to apply to the first column of the estimated covariance matrix of the snapshots. This is a complex vector that, in the usual statistical setting used to describe this situation, contains most of the information present in the data. The corresponding redundant basis that adequately allows to describe this vector is then a redundant set of steering vectors and the usual $\ell_2 - \ell_1$ penalized criterion is used to both detect and locate the paths that are present.

To solve this complex optimization problem, one can either resort to the iterative scheme we developed last year or use second order cone programming subroutines that are freely available (SeDuMi, <http://sedumi.mcmaster.ca>).

6.4. Distributed source and source-channel coding

Keywords: MAC channels, Slepian-Wolf coding, Wyner-Ziv coding, information theory.

6.4.1. Robust rate-adaptive non-symmetric Slepian-Wolf coding

Participants: Cédric Herzet, Aline Roumy, Velotiaray Toto-Zaraso.

Most of Slepian-Wolf codecs proposed in the literature consider the asymmetric case where one source is assumed to be perfectly known at the decoder. In this setup, one source (also called the side information) is sent at its entropy rate, while the other one is compressed at its conditional entropy rate. However, a flexible rate allocation to the different sources is beneficial for some applications such as light-field multi-view compression or in a wireless sensor network, where the correlations may vary and where the sensors should operate at any rate in the Slepian-Wolf region in order to meet some power constraints. This problem is referred to as non-symmetric Slepian-Wolf coding.

In 2008, we have continued to work on the non-symmetric distributed source coding problem. In [42], we have developed a scheme that (1) allows the sources to operate at any rate in the Slepian-Wolf region, for a given correlation and (2) that can adapt also to a wide range of correlations. The scheme proposed in [42] uses LDPC codes. Then we considered the case of convolutional and turbo-codes. In all previous non-symmetric method based on channel codes, the decoder multiply the compressed data by an inverse submatrix of the code. This multiplication presents two drawbacks. First, if turbo codes are used, the submatrix has no periodic structure s.t. the whole inverse has to be stored and no fast implementation exists for the multiplication. Second, this multiplication may lead to error propagation. We have therefore proposed a fast and robust scheme: first the scheme limits the error propagation phenomena, second its complexity grows only linearly with the blocklength. For the case of turbo-codes, design rules have been derived in order for the decoder to recover the sources. This work has been submitted to the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 09) and to the IEEE International Conference on Communications (ICC'09).

6.4.2. Distributed video compression

Participants: Christine Guillemot, Denis Kubasov, Aline Roumy.

A distributed video coding and decoding algorithm has been developed in the context of the European IST-Discover project (see Section 5.2.1). Despite recent advances, distributed video compression rate-distortion performance is not yet at the level of predictive coding. Key questions remain to bring monoview and multi-view DVC to a level of maturity closer to predictive coding: 1)- finding the best SI at the decoder for data not - or only partially - known; 2)- estimating at encoder or decoder the *virtual* correlation channel from unknown - or only partially known - data.

In 2008, we have continued to work on improving the side information quality. Several state-of-the-art image and video denoising methods (e.g. wavelet thresholding with the oriented wavelet transform, “block-matching and 3D filtering” method) have been tested with the purpose to reduce the variance of the correlation noise. The problem of the correlation noise modeling has also been addressed. First, a method was developed for improving the basic estimation (using key frames) of the correlation noise variance at the decoder. The method is based on the Expectation-Minimization approach (EM) and allows to refine the correlation parameter estimation during the turbo-decoding (which iterations are regarded as minimization steps). Second, a method was proposed for the robust estimation of the correlation model parameter at the encoder. This method performs motion estimation at the encoder (in the same manner as it is done at the decoder) for only 1/5 (20%) randomly selected blocks of each frame, and constructs the side information (SI) for 20% of the frame area. This SI is the true side information (i.e. the same as used by the decoder), and the estimation of the correlation model parameter can be done precisely for those 20% of the frame. The law of large numbers guarantees that this locally estimated parameter is close to the globally optimal estimation with the high probability. This latter encoder-based robust estimation technique was developed to support another developed algorithm, namely the rate-distortion optimization of the number of quantization levels per frequency band. Previously, this number was defined before encoding (offline), and remained fixed for all frames throughout the video sequence. Therefore, changes of the correlation model with the time were not taken into account, thus making the bit plane allocation scheme sub-optimal for certain frames. The developed algorithm permits to make this allocation on-line, i.e. during the encoding and using the actual correlation model parameter, corresponding to the frame being encoded. Lastly, the source coding aspect of the distributed video coding (DVC) has been addressed by two means. First, the oriented wavelet transform (OWT) has been tested instead of the DCT, and has been shown to improve the rate-distortion performance for several video sequences (containing a lot of structured edge information, which the OWT is oriented to represent effectively). Second, we adopted a Huffman code instead of a fixed-length code (FLC) for the binarization of quantization indices. This allows to efficiently exploit their probability law, and significantly decrease the total number of bits in all bit planes. However, the problem of re-synchronization of bit planes (which has an arbitrary length in case of variable-length codes) in case of non-zero residual bit error rate is not fully solved.

6.5. Image and video watermarking and security

Keywords: *Robust watermarking, detection of weak signal, fingerprinting, steganography.*

In 2008, we mostly focused our efforts on two points: *fingerprinting* (also known as user forensics, traitor tracing, transactional watermarking, content serialization ...) has a long history in research but real applications have slowly emerged this year. *Reliability* of watermarking technique is a new concept we are bringing in this community. A watermarking technique is reliable if the probability of detection errors is very small and if one can assess this fact. Unfortunately, very poor estimators of this probability were used so far. This has some importance since watermarking is all about trade-off. A watermarking technique may be more robust than another just because it has a bigger probability of false alarm. Bad estimator of such probability prevents fair comparison and benchmarking of watermarking technique.

6.5.1. Fingerprinting

Participants: Ana Charpentier, Caroline Fontaine, Teddy Furon, Fuchun Xie.

Fingerprinting (also known as user forensics, traitor tracing, transactional watermarking, content serialization ...) aims at hiding in a robust and imperceptible way, an identifier of the consumer. The goal is to enable traceability of the content and to find back dishonest users who have illegally redistributed the content (for instance, posting it in a P2P network). Fingerprinting has a long history in research but real applications have slowly emerged this year: DRM systems over protect the content and thus they are not user-friendly. Content distribution could get rid off DRM thanks to fingerprinting used as a dissuasive weapon. This is a hot topic in the watermarking community. Fingerprinting is a difficult problem because it is a cross-design merging two layers: the fingerprinting code (the set of identifiers) and the watermarking technique hiding the identifiers in content.

A new trend in this domain is the probabilistic fingerprinting codes: The user identifiers are random binary strings with a secret statistic structure. G. Tardos introduced this concept in 2003, but it took time for the community to recognize his work as a major breakthrough. We have made four contributions with respect to probabilistic fingerprinting codes.

- Tardos seminal work proposes a code design and proves that his performances are optimal. However, no clue is given on why it works so good, and on the setting of the parameters. Our first work has been to understand where the magic of the code comes from.
- The second step proposes some improvements. We derive the optimal setting when the number of colluders is known, or when their collusion strategy is known.
- Tardos was only interested in showing a code design whose performances asymptotically achieves the theoretical bound. However, in a real-life application the length of the identifiers is fixed, and the asymptotic derivations do not hold any longer. This is especially critical for the probability of accusing an innocent. We have invented an algorithm assessing this risk.
- We also proposed an implementation where Tardos identifiers are embedded in image sequences. We use a “on-off” keying modulation with a zero-bit watermarking scheme. This enforces that the assumptions made about the impact of the collusion of contents on the fingerprinting code layer are really met. A new accusation process takes into account the watermarking layer and is thus more adapted to multimedia fingerprinting.

6.5.2. Reliability

Participant: Teddy Furon.

Watermark decoders are in essence stochastic processes. There are at least three sources of randomness: the unknown original content (for blind decoders), the unknown hidden message and the unknown attack the watermarked content has undergone. The output of the decoder is thus a random variable and this leads to a very disturbing fact: there will be errors in some decoded messages. This also holds for watermark detectors which have to take the decision whether the content under scrutiny has been watermarked or not. In order to be used in an application, a watermarking technique must be reliable. We introduce here the concept of reliability as the guarantee that not only these inherent errors very rarely happen, but also that their frequency or their probability is assessed to be below a given level.

A strong collaboration with the project-team ASPI within the framework of the national project ANR-Nebbiano brought some key results. The main element is an algorithm which estimates extremely low probabilities of rare event with a dramatic accuracy and with a very limited computer power. This fast algorithm has been developed with watermarking applications in mind, but its scope is a priori much broader. This algorithm has been applied to two very difficult problems: the estimation of error exponents. A watermarking technique is deemed sound when its probability of false alarm exponentially vanishes as the detector processes asymptotically more data. The rate of this exponential decay is thus a criterion to compare watermarking schemes. Unfortunately, there are very few schemes where a closed-form expression is known. Experimentally measuring the exponent is extremely challenging due to this exponential decay to be observed in the asymptotical regime. The other problem has been already mentioned above: In the Tardos code, identifiers have a length equal to $Kc^2 \log(\epsilon^{-1})$, where c is the number of colluders and ϵ the probability

of accusing an innocent. Some recent papers tried to refine Tardos first evaluation of the minimum constant K . In other words, they are looking for the minimal identifier length providing a probability of accusing an innocent smaller than ϵ . With our algorithm, we were able to give an experimental value of this minimum constant. This algorithm has been patented.

6.5.3. Watermarking Security

Participant: Teddy Furon.

The work about watermarking security only tackled the follow-up of the international challenge BOWS-2 which lasted from 17th of July 2007 to 17th of April 2008. 10 millions of images were submitted. We recorded connections from 450 different IP addresses, covering 33 countries. On average, the website received 90 unique visitors a day (not participating to the challenge, but consulting daily results). This was a major animation for the watermarking community. We are now in the process of extracting the scientific results and the compiling the lessons learnt with respect to watermarking security. Collaborations with the most successful participants are under way to write journal articles. The article describing the technique is accepted to EURASIP Journal on Information Security.

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, Luce Morin.

- Title : 3D video representation for 3DTV and Free viewpoint TV.
- Research axis : § 6.1.3.
- 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.

During 2008, we studied existing 3D video representations and started to test a new one. In the 3D video community, methods for scene representation are often classified along an axis showing which kind of data is used. The extreme left of the axis corresponds to the image-based representation. The extreme right corresponds to model-based representations. The main advantage of the first is that all the image information is stored, providing a potentially high quality of rendering. On the contrary, the main advantage of the second is that the redundancies are greatly reduced, at the expense of the overall quality. This shows that the choice of the representation depends on a trade-off between quality and compactness. The representation that we have tested is based on a global geometric model composed with a soup of polygons.

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

Participants: Jean-Jacques Fuchs, Christine Guillemot, Aurélie Martin.

- Title : Spectral deconvolution: application to compression
- Research axis : § 6.2.2.
- Partners : Thomson, Irisa/Inria-Rennes.
- Funding : Thomson, ANRT.
- Period : Oct.06- Sept.09.

This CIFRE contract concerns the Ph.D of Aurélie martin. The objective of the Ph.D. is to develop image spectral deconvolution methods for prediction in video compression schemes. Closed-loop spatial prediction has indeed been widely used in video compression standards (H.261/H.263, MPEG-1/2/4, H.264). In H.264 used for digital terrestrial TV, the prediction is done by simply “propagating” the pixel values along the specified direction. This approach is suitable in presence of contours, the directional mode chosen corresponds to the orientation of the contour. However, it fails in more complex textured areas. In 2008, we have continued the study of sparse approximations based on matching pursuit algorithms for spatial image prediction. The methods have been integrated in the JVT (ITU/MPEG joint video team) KTA software for validation.

7.1.3. Collaboration with Alcatel on robust video compression

Participants: Simon Bos, Christine Guillemot, Laurent Guillo, Aline Roumy.

- Title: Self adaptive video codec
- Research axis: 6.3.2
- Funding: Joint research laboratory between INRIA and Alcatel
- Period: Nov. 2008 - Nov. 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. This collaboration concerns the Ph.D. of Simon Bos. The goal is to design a video codec, which has the intrinsic knowledge of the dynamic video quality requirements, and which is able to self-adapt to the existing underlying transport network. In this approach, the video codec has to include:

- Means to dynamically "sense" the underlying transport channel (e.g BER, PER, Markov model)
- 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. enditemize

7.2. National contracts

7.2.1. DGE/Region-FUTURIMAGE

Participants: Christine Guillemot, Vincent Jantet, Claude Labit, Luce Morin, Gael Sourimant.

- 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-Sept.11.

The Futurim@ages projects studies the prospects of the future television video formats, in three different directions: 3DTV, high-dynamic range videos, and full-HD TV. In this context, Vincent Jantet will study during his Ph.D. the transmission 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). These works will tend to solve two major problems related to these videos:

- Compression: multi-view videos represent a huge amount of additional data with regards to standard videos, but consist of several views highly correlated. This correlation should be taken into account to store and transmit these videos.

- Restitution: stereoscopic or auto-stereoscopic devices display very specific camera points of view, which should be generated even if they do not correspond to acquisition points of view. Artifacts such as ghosting or bad modelled occlusions must be dealt with to render high quality 3D videos.

7.2.2. *RIAM-ESTIVALE*

Participant: Caroline Fontaine.

- Title : Secured exchanges for video transfer, in line with legislation and economy
- Research axis : § 6.5.1.
- Partners : LIS (INPG), ADIS (Univ. Paris XI), CERDI (Univ. Paris XI), LSS (Univ. Paris XI/Supélec), Basic-Lead, Nextamp, SACD.
- Funding : ANR.
- Period : 12/12/2005-12/06/2009

ESTIVALE is a project dealing with the diffusion of video on demand in several contexts: from personal use to professional use. People involved in the project are from different communities: signal processing and security, economists and jurist. The goal of the project is to design technical solutions for securing this delivery, through DRM and watermarking tools, and to remain consistent with the economical and juridical studies and demands. In 2008, the TEMICS project-team has contributed on the design of a practical efficient fingerprinting scheme for video (see Section 6.5.1).

7.2.3. *RIAM-MEDIEVALS*

Participants: Caroline Fontaine, Ana Charpentier, Teddy Furon.

- Title : Watermarking and Visual Encryption for Video and Audio on Demand legal diffusion
- Research axis : § 6.5.1.
- Partners : MEDIALIVE, LSS (Univ. Paris XI/Supélec), 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 2008, the TEMICS project-team has contributed on the state of the art of fingerprinting techniques and the choices concerning the new software architecture.

7.2.4. *ACI Nebbiano*

Participant: Teddy Furon.

- Title : Nebbiano
- Research axis : § 6.5.2, § 6.5.3.
- Partners : Laboratoire de Mathématiques de J.A. Dieudonné - Université de Nice, Laboratoire des Images et des Signaux - INPG Grenoble.
- Funding : ANR
- Period : Jan. 2007 - Dec. 2009.

The Nebbiano project studies the security and the reliability of watermarking techniques. The international challenge BOWS-2 was organized in this framework with researchers from INPG Grenoble. The patented algorithm estimating probabilities of rare event is also the result of a collaboration with team-project ASPI within Nebbiano.

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. It extends over a 3-year period and its aim is 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.

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 will run over a 3-year period and 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.

7.2.7. Alibaba

Participant: Jean-Jacques Fuchs.

This is a contract running over about one year between Irisa/Université de Rennes 1 and Thalès Communications. It started in 2007 and finished on March 15, 2008. It concerns the evaluation of the performances of a source localization algorithm developed by Irisa many years ago, when applied to potentially correlated sources in the high frequency band (3-30 MHz).

7.2.8. ANR-ESSOR

Participants: Olivier Crave, Christine Guillemot, Aline Roumy, Velotiaray Toto-Zaratosoa.

- Title : Distributed Video Coding
- Research axis : § 6.3.1; § 6.4.1.
- Partners : CNRS/LSS, ENST-Paris, CNRS/I3S;
- Funding : ANR.
- Period : 01/11/2006-31/10/2009

Compared with predictive coding, distributed video compression holds a number of promises for mobile applications: a more flexible coder/decoder complexity balancing, increased error resilience, and the capability to exploit inter-view correlation, with limited inter-camera communication, in multiview set-ups. However, despite the growing number of research contributions in the past, key questions remain to bring monoview and multi-view DVC to a level of maturity closer to predictive coding: estimating at encoder or decoder the *virtual* correlation channel from unknown - or only partially known - data; finding the best SI at the decoder for data not - or only partially - known. Solutions to the above questions have various implications on coder/decoder complexity balancing, on delay and communication topology, and rate-distortion performance. These questions are being addressed by the ANR-ESSOR project. The TEMICS project-team more specifically contributes on the design of Slepian-Wolf and Wyner-Ziv coding tools as well as on the design of robust and joint source-channel distributed coding strategies. More specifically, in 2008, the TEMICS project-team has developed a robust video coder and decoder based on multiple description coding with side information. We have also significantly contributed to the ESSOR distributed video codec by providing a number of components such as the Slepian-Wolf coder and decoder based on LDPC codes, an optimum mmse estimator,

7.2.9. ANR-ICOS-HD

Participants: Christine Guillemot, Joaquin Zepeda.

- Title : Scalable Indexing and Compression Scalable for High Definition TV;
- Research axis : § 6.2.6.
- Partners : Université de Bordeaux, CNRS/I3S;
- Funding : ANR.
- Period : 01/01/2007-31/12/2009

The objective of the project is 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 descriptors to the content. The TEMICS project-team contributed in particular on the study of new forms of signal representation amenable to both compression and feature extraction (see Section 6.2.6). A new concept called visual sentences which can be seen as a sparse extension of the concept of visual words has been introduced and assessed for image retrieval in large databases.

7.3. European contracts

7.3.1. NoE-ECRYPT

Participant: Caroline Fontaine.

- Title: ECRYPT
- Research axis:§ 6.5.1
- Funding: CEE.
- Period: Feb. 2004 - July 2008

ECRYPT aims at stimulating European research in the area of cryptography and data hiding. It is splitted into several virtual labs, the one concerning us being called WAVILA: Watermarking Virtual Lab. Our contribution is the design of a practical efficient fingerprinting scheme for video content.

7.3.2. FP6-IST NOE NEWCOM++

Participants: Christine Guillemot, Cédric Herzet, Aline Roumy.

- Title: NEWCOM: Network of Excellence in Wireless Communication.
- Research axis: 6.3.2, 6.4.1
- Funding: CEE.
- Period: Jan. 2008 - Dec. 2009.

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 leaded by P. Duhamel, and coordinate the task TR7.3 Tools for multi-terminal JSCC/D.

The WP Joint source and channel co-decoding 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 (v) This will lead us to work in a “cross layer mood” in the context of JSCC/D, and solve such problems as finding the best possible redundancy allocation at the various layers for a given global rate, even in a context of JSCC/D.

8. Other Grants and Activities

8.1. International initiatives

Participant: Aline Roumy.

- Title : Radio resource optimization in iterative receivers
- Research axis : § 6.3.3.
- Partners : CNRS-DGRSRT/Tunisian university.
- Funding : CNRS-DGRSRT/Tunisian university.
- Period : Jan. 07 - Dec. 09.

This is a collaboration with N. Sellami (ISECS, Sfax, Tunisia) and I. Fijalkow (ETIS, Cergy France). The goal of the proposed project is the analysis of turbo-like receiver in order to allocate the resources (power, training sequence length...) of the system. The grant supports travel and living expenses of investigators for short visits to partner institutions abroad.

9. Dissemination

9.1. Patents

Three patents have been filed:

- A. Martin, J-J. Fuchs, C. Guillemot, D. Thoreau, "Prédiction spatiale inter couche", Thomson patent, filed in Jan. 2008.
- C. Guillemot, L. Guillo, J-J. Fuchs, "Méthodes de prédiction spatiale et temporelle", INRIA patent, filed in Nov. 2008.
- F. Cérou, A. Guyader, T. Furon, "Outil de vérification informatique" patent application n° 08/04584 filed in 2008.

9.2. Invited talks and book chapters

- C. Fontaine gave a lecture on information hiding during the IRISA Researchers' School "Sécurité des logiciels et des contenus"
- C. Guillemot gave an invited keynote talk on distributed video compression, at the IEEE International Symposium on Consumer Electronics, Portugal, 14-16 April 2008.
- C. Guillemot gave an invited keynote talk on distributed source and video coding, at the International workshop on wireless sensor networks, Gent, 9 April 2008.

- C. Guillemot and A. Roumy have been invited to write a chapter entitled “Towards constructive Slepian-Wolf coding schemes” in the edited book on distributed source coding, Edited by M. Gaspar and P.L. Dragotti, Elsevier Inc., 2008.

9.3. Visits

- L. Morin has visited the Technical University of Cluj-Napoca, Romania, June 7-12, for an organising mission (OM) within the ERASMUS exchange program settled between IFSIC and UTC-Cluj.
- L. Morin has organized the visit of a Romanian delegation from UTC-Cluj to IFSIC and IRISA (Sept. 21-24), where they individually met 14 research teams. The members of the Romanian delegation were Prof. Sergiu Nedeveschi, Dean of the Automatics and Computer Science Faculty, Mrs Mihaela Dînsoreanu, Vice-Dean, Mr Alin Suciu, Assistant-professor, and Mr Octavian Cret, Assistant-Professor.

9.4. Leadership within the scientific community

- C. Fontaine is associate editor of the Journal in Computer Virology (Springer-Verlag);
- C. Fontaine was a member of the program committee of the conference SSTIC 2008 (Rennes, France, June);
- C. Fontaine was a member of the organizing committee of the conference SSTIC 2008 (Rennes, France, June);
- C. Fontaine is a member of the scientific advisory board of the Brittany competence center Diwall;
- J.J. Fuchs is a member of the technical program committees of the following conferences : Eusipco 2008 and GretsI 2008;
- J.J. Fuchs is a member of the committee that delivers the best thesis price in Signal and Image processing (prix de thèse en Signal-Image du club EEA).
- T. Furon is associate editor of the EURASIP journal on Information Security;
- T. Furon was a member of the technical program committees of the following conferences: SPIE 2008 Security, Steganography, and Watermarking of Multimedia Contents X, Information Hiding 2008, European Signal Processing Conference 2008, and International Workshop on Digital Watermarking 2008;
- T. Furon is the co-organiser of the international watermarking challenge BOWS-2;
- C. Guillemot is associate editor of the international journal “New Trends in Signal Processing”.
- C. Guillemot is associate editor of the journal IEEE Transactions on Signal Processing (2007-2009).
- C. Guillemot is an elected member of the IEEE MMSP (MultiMedia Signal Processing Technical Committee) international committee;
- C. Guillemot has been guest editor of a special issue on distributed video coding of the Euraspip journal on image communication, 2008.
- C. Guillemot is a member of the external scientific advisory board of the IST Network of Excellence VISNET2;
- 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.
- C. Guillemot was the general chair of the special sessions of the IEEE ICME conference, Hannover, May 2008.
- C. Guillemot was a member of the technical program committees of the following conferences: EUSIPCO 2008, IEEE-MMSP 2008, WIAMIS 2008, CORESA 2008;

- C. Guillemot has served as a member in the award committee of the Eurasip Image communication journal.
- C. Guillemot is member of the “Specif Thesis Award” committee.
- C. Guillemot is the coordinator of the ANR ICOS-HD project.
- C. Labit has served as 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.
- L. Morin was responsible for International Relationship for the IFSIC Engineer Degree.
- A. Roumy was a member of the technical program committee of EUSIPCO 2008 (European Conference on Signal Processing 2008).
- A. Roumy and C. Herzet coordinate the task TR7.3 Tools for multi-terminal JSCC/D of the Network of excellence Newcom++.

9.5. Exhibitions and demonstrations

- The TEMICS project-team presented demos at the exhibition held during the NEM summit held in Saint-Malo, Sept. 2008.

9.6. Teaching

- Enic, Villeneuve-d’Ascq, (C. Guillemot: Video communication) ;
- Esat, Rennes, (C. Guillemot: Image and video compression; T. Furon: Watermarking) ;
- Engineer degree Diic-INC, Ifsic-Spm, university of Rennes 1 (L. Morin, C. Guillemot, L. Guillo, T. Furon, G. Sourimant : image processing, 3D vision, motion, coding, compression, cryptography, communication) ;
- Engineer degree Diic-LSI, Ifsic-Spm, university of Rennes 1 (L. Morin, L. Guillo, G. Sourimant : compression, video streaming) ;
- Engineer degree DIIC, Ifsic-Spm, Université de Rennes 1: J-J. Fuchs teaches several courses on basic signal processing and control ;
- Supelec (T. Furon : steganography and watermarking).
- Master Research-2 SISEA: J-J. Fuchs teaches a course on optimization and C. Guillemot and C. Labit teach a course on image and video compression ;
- Master Research-2 Computer Science: C. Fontaine is in charge of the “Information and Computing Infrastructure Security” track, and teaches a course on information hiding ;
- Master, Security of Information Systems, Supelec-ENSTB (C. Fontaine: information hiding) ;
- Professional degree Tais-Cian, Breton Digital Campus (L. Morin, G. Sourimant : Digital Images -online course-) ;
- Master, Network Engineering, university of Rennes I (L. Guillo, Video streaming) ;
- Computer science and telecommunications magistère program, Ecole Normale Supérieure de Cachan, Ker Lann campus. (A. Roumy: Information theory and communication theory) ;
- Master SIC (Systèmes Intelligents et Communicants) at ENSEA, université de Cergy Pontoise. (A. Roumy: Information theory, Modern coding theory and Multiuser detection) ;

- DRT of Rennes1 university: C. Labit supervises a DRT cursus (K. Torres) with Thomson Grass Valley addressing the problem of rate-distortion control for video MPEG-like codecs.

10. 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", numéro spécial "Supplement on Secure Media III", vol. 53, n^o 10, 2005, p. 3976–3987.
- [3] V. CHAPPELIER, C. GUILLEMOT. *Oriented wavelet transform for image compression and denoising*, in "IEEE Transactions on Image Processing", vol. 15, n^o 10, Oct. 2006, p. 2892-2903.
- [4] J. FUCHS. *On the application of the global matched filter to DOA estimation with uniform circular arrays.*, in "IEEE Trans. on Signal Processing", n^o 49, Apr. 2001, p. 702-709.
- [5] T. FURON, P. DUHAMEL. *An Asymmetric Watermarking Method*, in "IEEE Trans. on Signal Processing", Special Issue on Signal Processing for Data Hiding in Digital Media and Secure Content Delivery, vol. 51, n^o 4, April 2003, p. 981-995.
- [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", vol. 12, Dec. 2003, p. 1599-1609.
- [7] H. JÉGOU, C. GUILLEMOT. *Robust multiplexed codes for compression of heterogeneous data*, in "IEEE Trans. on Information Theory", vol. 51, n^o 4, Apr. 2005, 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", vol. 12, n^o 9, Sept. 2003, p. 1107–1119.
- [9] S. PATEUX, G. LE GUELVOUT. *Practical watermarking scheme based on wide spread spectrum and game theory*, in "Signal Processing : Image Communication", n^o 18, Apr. 2003, 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", vol. 52, n^o 11, Nov. 2004, p. 3241-3252.

Year Publications

Doctoral Dissertations and Habilitation Theses

- [11] O. CRAVE. *Approches théoriques en codage vidéo robuste multi-terminal*, Ph. D. Thesis, Thèse de doctorat en Traitement du Signal, Télécom ParisTech, France, Dec. 2008.

- [12] D. KUBASOV. *Codage de sources distribuées: Nouveaux outils et application à la compression vidéo*, Ph. D. Thesis, Thèse de doctorat en Informatique, Université de Rennes 1, France, Dec. 2008.
- [13] S. MALINOWSKI. *Codes joints source-canal pour transmission robuste sur canaux radio-mobiles*, Ph. D. Thesis, Thèse de doctorat en Traitement du Signal et Télécommunications, Université de Rennes 1, France, Dec. 2008.

Articles in International Peer-Reviewed Journal

- [14] O. CRAVE, C. GUILLEMOT, B. PESQUET-POPESCU, C. TILLIER. *Distributed temporal multiple description coding for robust video transmission*, in "EURASIP Journal on Wireless Communications and Networking, special issue on multimedia over wireless networks", vol. 2008, Jan. 2008.
- [15] C. FONTAINE, C. DELPHA, P. DUHAMEL, A. BENJELLOUN TOUIMI, M. MILHAU, A. LE GUYADER, C. GIRAUD, D. MARTIN. *An end-to-end security architecture for multimedia content distribution on mobile phones*, in "ISAST Communication and Networking Journal", vol. 2, n^o 1, 2008, p. 81-91.
- [16] C. FONTAINE, F. GALAND. *How can Reed-Solomon codes improve steganographic schemes*, in "EURASIP Journal on Information Security", special issue "Secure Steganography in Multimedia Content"; to appear, 2008.
- [17] T. FURON, P. BAS. *Broken Arrows*, in "accepted for publication in EURASIP Journal on Information Security", 2008.
- [18] C. GUILLEMOT, F. PEREIRA, L. TORRES. *Editorial of the special issue on distributed video coding*, in "Eurasip Journal on Image Communication, Elsevier Science Inc.", vol. 23, June 2008, p. 337-338.
- [19] C. HERZET, V. RAMON, A. RENAUX, L. VANDENDORPE. *A Cramer-Rao Bound Characterization of the EM-algorithm mean speed of convergence*, in "IEEE Trans. Signal Processing", vol. 56, June 2008, p. 61-70.
- [20] I. H. KACEM, N. SELLAMI, A. ROUMY, I. FIJALKOW. *Training sequence length optimization for a turbo-detector using decision-directed channel estimation*, in "Research Letters in Communications", 2008, Article ID 860368.
- [21] S. MALINOWSKI, H. JÉGOU, C. GUILLEMOT. *Error recovery properties and soft decoding of quasi-arithmetic codes*, in "Eurasip Journal on Applied Signal Processing", Article ID 752840, 12 pages, vol. 2008, Jan. 2008.
- [22] S. MALINOWSKI, H. JÉGOU, C. GUILLEMOT. *Posterior marginals computation on aggregated state models for soft source decoding*, to appear, in "IEEE Trans. on communication", 2008.
- [23] F. PEREIRA, L. TORRES, C. GUILLEMOT, T. EBRAHIMI, R. LEONARDI, J. OSTERMAN. *Distributed Video Coding Selecting the Most Promising Application Scenarios*, in "Eurasip Journal on Image Communication, Elsevier Science Inc.", vol. 23, June 2008, p. 339-352.
- [24] G. RATH, C. GUILLEMOT. *Complementary matching pursuit algorithms for sparse approximation*, in "IEEE Trans. on Signal Processing", to appear, 2008.

- [25] G. RATH, W. YANG, C. GUILLEMOT. *Compression of Laplacian Pyramids through Orthogonal Transforms and Improved Predictions*, in "IEEE Trans. on Signal Processing", vol. 17, Sept. 2008, p. 1587-1604.
- [26] N. SELLAMI, A. ROUMY, I. FIJALKOW. *A proof of convergence of the MAP turbo-detector to the AWGN case*, in "IEEE Trans. on Signal Processing, accepted for publication", vol. 56, n^o 4, April 2008, p. 2716 - 2724.

International Peer-Reviewed Conference/Proceedings

- [27] X. ARTIGAS, S. MALINOWSKI, C. GUILLEMOT, L. TORRES. *Overlapped arithmetic codes with memory*, in "European Signal Processing Conference, EUSIPCO, Lausanne", Aug. 2008.
- [28] A. CHARPENTIER, F. XIE, T. FURON, C. FONTAINE. *Expectation Maximisation decoding of Tardos probabilistic fingerprinting code*, in "Security, steganography and watermarking of multimedia contents", N. MEMON (editor), 2008.
- [29] O. CRAVE, C. GUILLEMOT, B. PESQUET-POPESCU. *Multiple description source coding with side information*, in "European Signal Processing Conference, EUSIPCO, Lausanne", Aug. 2008.
- [30] F. CÉROU, P. DEL MORAL, T. FURON, A. GUYADER. *Rare event simulation for a static distribution*, in "Proc. 7th Int. Work. on Rare Event Simulation, Rennes, France", September 2008.
- [31] J. FUCHS. *A sparse representation criterion: recovery conditions and implementation issues*, in "17th World Congress IFAC, Seoul", July 2008, p. 10225–10229.
- [32] J. FUCHS. *On the use of sparse representations in the identification of line spectra*, in "17th World Congress IFAC, Seoul", July 2008, p. 12425–12429.
- [33] J. FUCHS. *On the use of the Global Matched Filter for DOA estimation in the presence of correlated waveforms*, in "42th ASILOMAR Conf. on Signals, Systems and Computers, Monterey", Oct. 2008.
- [34] J. FUCHS. *Sparse representations: recovery conditions and fast algorithm for a new criterion*, in "European Signal Processing Conference, EUSIPCO, Lausanne", Aug. 2008.
- [35] T. FURON, A. GUYADER, F. CÉROU. *On the design and optimisation of Tardos probabilistic fingerprinting codes*, in "Proc. of the 10th Information Hiding Workshop, Santa Barbara, Cal, USA", may 2008.
- [36] C. HERZET. *On the Convergence of the Iterative "Pseudo Likelihood" Maximization Algorithm*, in "IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP', Las Vegas, USA", April 2008, p. 3701-3704.
- [37] I. H. KACEM, N. SELLAMI, A. ROUMY, I. FIJALKOW. *Training sequence optimization for frequency selective channels with MAP equalization*, in "Proc. of IEEE International Symposium on Communications, Control and Signal Processing ISCCSP, Malta", March 2008, p. 532-537.
- [38] A. MARTIN, J. FUCHS, C. GUILLEMOT, D. THOREAU. *Atomic decompositions dedicated to AVC and spatial SVC prediction*, in "IEEE Intl. Conference on Image Processing, ICIP, San Diego, USA", Oct. 2008.

- [39] J. NAYAK, G. RATH, D. KUBASOV, C. GUILLEMOT. *Spatial prediction in distributed video coding using Wyner-Ziv DPCM*, in "IEEE Intl. Workshop on Multimedia Signal Processing, IEEE-MMSP, Cairns, Australia", Oct. 2008.
- [40] G. RATH, C. GUILLEMOT, J. FUCHS. *Sparse approximations for joint source-channel coding and its application to image transmission*, in "IEEE Intl. Workshop on Multimedia Signal Processing, IEEE-MMSP, Cairns, Australia", Oct. 2008.
- [41] G. RATH, C. GUILLEMOT. *A complementary matching pursuit algorithm for sparse approximation*, in "European Signal Processing Conference, EUSIPCO, Lausanne", Aug. 2008.
- [42] V. TOTO-ZARASOA, A. ROUMY, C. GUILLEMOT. *Rate-adaptive codes for the entire Slepian-Wolf region and arbitrarily correlated sources*, in "IEEE Intl. Conference on Acoustic, Speech and Signal Processing, Las Vegas, USA", Apr. 2008.
- [43] F. XIE, T. FURON, C. FONTAINE. *On-Off Keying modulation and Tardos fingerprinting*, in "ACM Multimedia and security 2008 Oxford, UK", Sept. 2008.

Scientific Books (or Scientific Book chapters)

- [44] C. GUILLEMOT, A. ROUMY. *in book "Distributed source coding"*, in "Chapter entitled "Towards constructive Slepian-Wolf coding schemes"", Elsevier Inc., 2008.

Research Reports

- [45] F. CÉROU, T. FURON, A. GUYADER. *Experimental assessment of the reliability for watermarking and fingerprinting schemes*, Technical report, n^o RR-6636, INRIA, 2008, <http://hal.inria.fr/inria-00319768/fr/>.