



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*

2007

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	7
4.1. Introduction	7
4.2. Compression with advanced functionalities	7
4.3. Multimedia communication	8
4.4. Copy protection, copyright enforcement and enriched content	8
5. Software	8
5.1. Libit	8
5.2. Soft VLC decoding library	9
5.3. PICS	9
5.4. Distributed video codec	9
5.5. Video communication platform	9
5.6. Oriented wavelet based image codec	10
5.7. Wavelet based video codec	10
5.8. 3D model-based video codec and player	10
5.9. Broken Arrows Software	11
5.10. Watermarking platform	11
6. New Results	11
6.1. Analysis and modelling of video sequences	11
6.1.1. 3d modelling of urban environments from video, GPS and GIS data	11
6.1.2. 3d/2d frame interpolation for distributed video coding	12
6.2. Sparse representation, compression and interaction with indexing	14
6.2.1. Sparse signal representation algorithms	14
6.2.2. Compression and prediction based on sparse signal representations	15
6.2.3. Scalable compression: prediction and perceptual coding	15
6.2.4. Feature extraction and subspace-based methods for compression and local description	16
6.3. Joint source-channel coding	16
6.3.1. Multiple description coding	16
6.3.2. Error-resilient source codes and soft decoding	17
6.3.3. Sparse representation based error decoding	17
6.3.4. Parameter estimation in convolutive and instantaneous mixtures	18
6.3.5. Robust estimation and detection	19
6.4. Distributed source and source-channel coding	19
6.4.1. Distributed source codes based on punctured and overlapped quasi arithmetic codes	19
6.4.2. Rate-adaptive asymmetric and symmetric Slepian-Wolf coding	20
6.4.3. Optimum Filters for Wyner-Ziv predictive coding	20
6.4.4. Correlated sources over multiple access channels	21
6.4.5. Distributed coding for sensor networks	21
6.4.6. Distributed video compression	22
6.5. Image and video watermarking and security	22
6.5.1. Optimal watermark detection	22
6.5.2. Watermarking Security	23

6.5.3. Fingerprinting	23
6.5.4. Steganography	24
6.5.5. Reliability	24
7. Contracts and Grants with Industry	25
7.1. Industrial contracts	25
7.1.1. CRE with France Telecom on 3D reconstruction of urban scenes	25
7.1.2. CIFRE contract with Thomson on the problem of spectral deconvolution for video compression and protection	26
7.1.3. Contract with Thalès Communications on localization of correlated sources in the high frequency band.	26
7.2. National contracts	26
7.2.1. RNRT-COSINUS	26
7.2.2. RIAM-ESTIVALE	27
7.2.3. ACI Nebbiano	27
7.2.4. ACI Masse de données - Codage	27
7.2.5. ANR-COHDEQ 40	28
7.2.6. ANR-ESSOR	28
7.2.7. ANR-ICOS-HD	28
7.3. European contracts	29
7.3.1. FP6-IST NoE SIMILAR	29
7.3.2. FP6-IST NOE NEWCOM	29
7.3.3. NoE-ECRYPT	30
7.3.4. FP6-IST STREP DISCOVER	30
8. Other Grants and Activities	30
9. Dissemination	31
9.1. Invited talks	31
9.2. Leadership within the scientific community	31
9.3. Awards	32
9.4. Exhibitions and demonstrations	32
9.5. Invitations and visits	32
9.6. Teaching	32
10. Bibliography	33

1. Team

Head of project-team

Christine Guillemot [DR Inria, HdR]

Administrative assistant

Huguette Béchu [TR Inria, shared with VISTA]

Inria research scientists

Teddy Furon [CR]

Cédric Herzet [CR, since the 4th of Dec.]

Gagan Rath [CR]

Aline Roumy [CR]

Cnrs research scientist

Caroline Fontaine [CR]

Faculty members

Jean-Jacques Fuchs [Professor, Ifsic, University of Rennes 1, HdR]

Luce Morin [Associate professor, Ifsic, University of Rennes 1, HdR]

CNRS technical staff

Laurent Guillo [IR]

Inria projects technical staff

Khaled Lajnef [IST-DISCOVER project, 1st Apr.-31 Aug., 2007]

Inria junior technical staff

Jérôme Allasia

Mathieu Desoubes [since 1st Nov. 2007]

Inria post-doctoral fellows

Zhenzhong Chen [funded by ERCIM, since Sept 10, 2007]

Fabien Galand [RIAM-ESTIVALE project, since Sept. 1st, 2007]

Yang Liu [until Aug. 31, 2007]

Ariane Herbulot [since Nov. 1, 2007, also with TexMex]

Ph.D Students

Thomas Collet [CIFRE contract with France Télécom, since Oct. 1st, 2007]

Olivier Crave [Inria grant, ANR-ESSOR]

Angélique Drémeau [Inria grant, since Oct. 15, 2007]

Denis Kubasov [Inria grant, IST-DISCOVER project]

Simon Malinowski [Ministry of research grant]

Aurélien Martin [CIFRE contract with Thomson]

Gael Sourimant [Inria grant, CRE France Télécom]

Velotiaray Toto-Zaraso [Ministry of research grant, since Oct. 1st, 2007]

Fuxun Xie [Inria grant, CORDIS]

Joaquin Zepeda [Inria grant, ANR-ICOS-HD project, also with TexMex]

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, ACI CODAGE, RNRT COSINUS, RIAM ESTIVALE), ANR ESSOR, ANR ICOS-HD) and European projects (IST-SIMILAR, IST-DISCOVER and 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. Let $E(\cdot)$ be the extraction function going from the content space \mathcal{C} to the components space, isomorphic to \mathbf{R}^N

$$\begin{aligned} E(\cdot) : \mathcal{C} &\mapsto \mathbf{R}^N \\ C &\rightarrow \mathbf{V} = E(C) \end{aligned}$$

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, also sketched in Fig. 1, 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 :

$$\begin{aligned} S(\cdot) : \mathcal{M} &\mapsto \mathbf{R}^N \\ m &\rightarrow \mathbf{S} = S(m) \end{aligned}$$

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, as sketched in Fig. 2, 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.

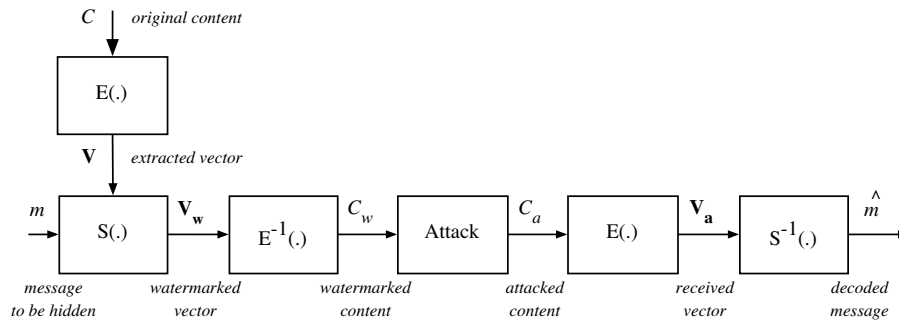


Figure 1. Classical watermarking scheme

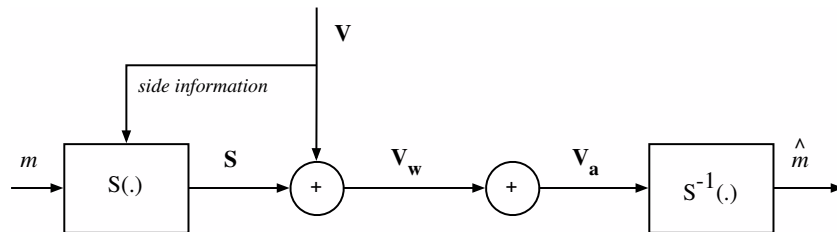


Figure 2. Watermarking as a problem of communication with side information.

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. Libit

Participant: Christine Guillemot [contact person].

Libit is a C library developed by Vivien Chappelier and Hervé Hégoü 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 developped, ported and modified simply. As examples and to ensure the correctness of the algorithms with respect to published results, some test programs are also provided. (URL: <http://libit.sourceforge.net>). The library is made available under the GNU Library General Public Licence.

5.2. 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.

5.3. PICS

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.

5.4. Distributed video codec

Participant: Denis Kubasov [contact person].

The TEMICS project-team contributed to the IST-Discover software which implements a distributed video coder and decoder. The executable files, along with sample configuration and test files, can be downloaded from <http://www.discoverdvc.org/>.

5.5. Video communication platform

Participants: Jérôme Allasia, 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 based 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.

5.6. 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.

5.7. 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.

5.8. 3D model-based video codec and player

Participants: Jérôme Allasia, Laurent Guillo, Luce Morin [contact person].

A 3D player supporting rendering of the 3D scene and navigation within the scene has been developed. It integrates as a plug-in in the 3D model-based video codec of the team. The Temics project-team has indeed 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. The first version of the software has been registered at the Agency for the Protection of Programmes (APP) under the number IDDN.FR.001.130017.000.S.P.2003.000.41200. The 3D videoc also supports scalable coding of both geometry and texture information. The 3D player integrated to the VISIUM communication platform allows remote access to a 3D sequence. The scalability of the coded representation of the 3D models and texture enables dynamic adaptation of the transmitted bitstream to both network and terminal capabilities.

5.9. 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 named "Broken Arrows". This technique is the main element of the international challenge BOWS-2 (Break Our Watermarking System - 2nd Edition). See URL: bows2.gipsa-lag.inpg.fr. The watermark embedder and the watermark detector have been coded in C, with optimization to lower as much as possible the detection time. For the first phase of the challenge (July 17th - October 17th), the technique is not disclosed and the contenders must remove the watermark from three images available on the website while maintaining a good quality. They submit their attacked images to the server which answers whether the watermark is still detectable and measures the PSNR. The number of submissions is limited to 30 per day. The winner succeeds to hack the three pictures with the highest mean PSNR. During this first episode, the watermark detector software run almost 20,000 times. In the second episode (17th October - 17th January, 2008), the watermark detector executable is distributed to the contenders and the number of submissions is no longer limited. This episode aims at studying the impact of oracle attacks where the pirate refers to a black sealed box detector as many times as needed to forge a pirated content. In the third phase (17th January, 2008 - 17th April, 2008), many images watermarked with the same secret key will be distributed to the contenders. This last phase aims at studying security attacks where the pirate gains some knowledge about the algorithm and the secret key by observing many watermarked contents.

5.10. Watermarking platform

Participants: Mathieu Desoubeaux, Teddy Furon [contact person].

This software platform based on Virtual Dub 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 will 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. The embedding and extraction algorithms are based on wide spread spectrum and dirty paper codes. Embedding and extracting parameters are optimized with the help of game theory taking into account optimal attack and potential de-synchronizations. The first version of the software has been 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 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.

6. New Results

6.1. Analysis and modelling of video sequences

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

6.1.1. 3d modelling of urban environments from video, GPS and GIS data

Participants: Thomas Collet, 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 stereovision 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. In addition, the model is limited to the areas captured by the video camera. In the case of urban environments, 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 2007, we focused on the automatic initialization of the registration procedure, and on texture extraction and fusion.

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.

Once camera poses are estimated for each video frame, each visible façade in each image provides a texture in the rectified façade plane space. All the extracted textures for a given façade are fused in a pixel-wise scheme to compute the final texture that will be used in the enhanced 3D model. The fusion consists of a weighted averaging function, depending on pixels visibility, luminance values spread and texture spatial resolution. The final textures are therefore cleaned of all occluding objects such as poles, cars, etc (see Fig. 3).

6.1.2. 3d/2d frame interpolation for distributed video coding

Participants: Christine Guillemot, Luce Morin.



(a)

(b)

(c)



(d)



(e)

Figure 3. Successive steps of the initial Gis-Video registration: (a)- rough pose estimate; (b)- image context (sky, façade and sky); (c)- lines extracted from the image; (d)- lines extracted from the projected 3D model; (e)- final estimated precise pose

Distributed Video Coding (DVC) has emerged as an alternative to classical motion-compensated codecs (e.g. MPEGx/H.26x). It offers improved robustness, scalability and low complexity at the coder. A drawback of such an algorithm is its reduced rate-distortion performances. This becomes clear by looking at the outline of the algorithm: a video sequence is split into a series of Group of Frames (GOF) whose first frames called “keyframes” are coded using intra coding (e.g. JPEG) and whose other frames are reconstructed at the decoder by first applying block-based motion compensation (BB-MC) between consecutive keyframes and correcting this prediction using some information called “parity bits” coming from the coder. BB-MC being a crude motion model, the keyframe frequency must be very high (usually every other frame) to maintain an acceptable PSNR, causing a major performance hit.

We have focused on the design of new motion models. First, motion models based on 3D mesh modeling have been investigated, assuming that the video contains a static scene acquired by a moving camera. Then, we have investigated several approaches with the aim of relaxing these constraints and considering generic 2D motions.

Assuming that the scene is static, the “epipolar geometry” can be estimated by detecting and matching corners on pairs or triplets of keyframes. The scene is then modeled by a mesh whose control point depths are estimated from the matched corners. This mesh defines a motion field over the intermediate frames, which allows their prediction from the keyframes. Experimental results have shown that such a 3D-DVC scheme allows keyframes to be separated by more than 10 frames. Moreover, the above results led us to consider techniques beyond strict DSC, which would keep its valuable characteristics. We have proposed two different schemes. The first scheme extracts points at the encoder and performs tracking at the decoder. Such a scheme enables to greatly improve side-information estimation by reducing misalignments. In the second scheme points are tracked between consecutive frames at the encoder thus introducing a limited temporal dependency between a keyframe and its following frames in the GOF. A better estimation of camera pose for intermediate frames provides better SI prediction and rate-distortion performances. The performance of the three proposed schemes has been assessed by comparing their rate-distortion performances with the standard H.264 codec (in Intra and IPPP modes), and with state-of-the art 2D-DVC codecs, on video sequences of static scenes. 3D-DVC-TD and 3D-qDVC-TE outperform H.264 Intra and classical 2D-DVC [25]. The 3D coders provide two important insights into frame interpolation for distributed video compression: first, motion fields between keyframes need to be close to the groundtruth, and second, small but precise motion adjustments are required at intermediate frame to align the intermediate motion fields with the groundtruth. Both of these properties are more difficult to attain with 2D motions than with 3D motions due to the absence of epipolar geometry and of global motion model at intermediate frames (the intermediate projection matrices).

Like the 3D codec, the 2D codec encodes the keyframes independently using H.264 intra. Sparse correspondences between keyframes are then detected at the decoder, and they serve as an initialization for dense motion estimation. They are then matched to obtain correspondences using constraints on motion size, normalized cross-correlation between blurred-image descriptors, sampling-independent sum of absolute differences, acceleration, and unicity. Correspondences are propagated between triplets of keyframes to make them denser. Dense motion estimation is cast into a maximum-a-posteriori problem where motion vectors are hidden variables forming a hidden Markov field with jointly-Gaussian probabilistic distributions. Motion adjustment at intermediate frames is obtained by detecting edges at the encoder and tracking these edges at the decoder. Edges are tracked using dense motion estimation between the decoded edge-images with the motion field between keyframes as prior. The proposed frame interpolation (OF) outperforms the classical Block-Based Motion Interpolation (BBMI) in terms of side-information correlation and rate-distortion. However, the PSNR improvements brought by edge-based alignment are not yet sufficient to overcome the bitrate overhead they generate.

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

Participant: Jean-Jacques Fuchs.

Sparse representations has become an important topic with numerous applications in signal or image processing. While we developed our research in this area, in an estimation-detection context [17], [41], [39], it is actually in a compression or coding context that it finds most of its applications. In order to further widen its applicability and open potentially new domains of applications we have considered a new criterion. Instead of the ubiquitous criterion $\min_x \|x\|_1$ under $\|Ax - b\|_2^2 \leq \rho_2$ or its equivalent forms $\min_x \frac{1}{2} \|Ax - b\|_2^2 + h\|x\|_1$ or $\min_x \|Ax\|_2^2$ under $\|A^T(Ax - b)\|_\infty \leq h$, we have considered the criterion $\min_x \|x\|_1$ under $\|Ax - b\|_\infty \leq \rho_\infty$ or $\min_x \|x\|_1$ under $\|Ax - b\|_1 \leq \rho_1$ and developed iterative [16] or fast algorithms to solve them [42]. While it appears that the corresponding representations are no longer parsimonious, they have different properties that might be of interest. Preliminary investigations in image denoising are presented in [42].

6.2.2. Compression and prediction based on sparse signal representations

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

Closed-loop spatial prediction has been widely used for image compression in transform (H.261/H.263, MPEG-1/2/4) or spatial (H.264) domains. In H.264, 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. We have addressed the problem of spatial image prediction in highly textured areas, and developed prediction methods based on sparse signal approximations. The problem of closed-loop spatial image prediction or extrapolation can indeed be seen as a problem of texture synthesis (close to 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. However, the stopping criterion (which is the energy of the residue) is computed on the region to predict. To compute it on the causal neighborhood would lead to a residue of small energy, however, this residue might take veryof potentially large values in the region to be predicted. The number of *atoms* selected in order to minimize the energy of the residue on the region to predict is transmitted. The decoder runs the algorithm with the *masked* basis functions and taking the previously decoded neighborhood as the known support. The number of atoms selected by the encoder is used by the decoder as a stopping criterion. Several algorithms have been considered for searching for the best lineat expansion: the Matching Pursuit (MP) algorithm and an alternative optimal sparse representation called Global Match Filter (GMF). The advantage of GMF, compared to MP, is that the best *atoms* are simultaneously selected instead of choosing them one by one. Significant prediction gains have been achieved in highly textured areas. The method has been extended to inter-layer prediction in a scalable video coding context.

6.2.3. Scalable compression: prediction and perceptual coding

Participants: Zhenzhong Chen, Christine Guillemot, Yang Liu, Gagan Rath.

Scalable representation of visual signals such as image and video is highly important for multimedia communications. Spatial scalability, as the name suggests, provides signal layers with different spatial resolutions. This is usually offered as a low resolution coarse signal together with several higher resolution enhancement layers. In the context of scalable video coding (SVC), the compression of these spatial layers is an important issue. For the intra coding, the current SVC standard uses two prediction options: the spatial intra prediction as used in H.264/AVC and the inter-layer texture prediction. The encoder selects the prediction mode with the minimal rate-distortion (RD) cost. These two prediction modes, however, can be considered jointly by simple modifications: The upsampled base layer can be used to predict the spatial intra mode in the enhancement layer. This helps reducing the number of candidate modes, thus the bits used to signal the mode index can be reduced. A filter across the boundary between the upsampled base layer and the enhancement layer neighboring pixels has also been introduced in the scalable coder in order to increase the quality of the reference for inter-layer prediction. A study on perceptual coding in order to optimize the rate allocation taking into account perceptual features has also been started.

6.2.4. Feature extraction and subspace-based methods for compression and local description

Participants: Christine Guillemot, Joaquin Zepeda.

In 2006, in collaboration with TEXMEX (Patrick Gros) through the thesis of François Tonin, we have worked on the design of a feature point extractor and of a local descriptor in signal representations given by the over-sampled steerable transforms. Although the steerable transforms, due to their properties of covariance under translations and rotations, and due to their angular selectivity, provide signal representations well-suited to feature point and descriptor extraction, the opposite constraints of image description and compression were not fully solved. In addition, the advent of multimedia applications on mobile and heterogeneous environments is triggering the need for scalable signal representation and description.

The objective of the study initiated in 2007, in collaboration with Ewa Kijak from TEXMEX, is to design scalable signal representation and approximation methods amenable to both compression (that is with sparseness properties) and image 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, which ideally is the set of perspective transformations. Reducing this set of transformations to the group of isometries, and adding the constraint of causality, the image representation is uniquely characterized by the Gaussian scale space. The Gaussian scale space is however not amenable to compression. We have thus started investigating subspace-based approaches and sparse representations for local image texture description. These aspects form the core of the ICOS-HD project started in 2007.

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

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. Several correlated coded representations of the signal are created and transmitted on different channels. The design goals are therefore to achieve the best average rate-distortion (RD) performance when all the channels work, subject to constraints on the average distortion when only a subset of the channels is received correctly. 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. Multiple description coding offers nice properties for robust video transmission in peer-to-peer networks.

Two-description coding schemes based on overcomplete temporal signal expansions and different frame splitting patterns, have been designed. Overlapping motion-compensated two-band or three-band Haar decomposition are applied on the frames of each description. These techniques result in good central RD performances, but in high PSNR quality variation at the side decoders. To enhance the quality of the signal reconstructed by the side decoders, extra Wyner-Ziv coded data are then transmitted. This amounts to having a systematic lossy Wyner-Ziv coding of every other frame of each description, or alternatively of the low frequency temporal information present in each description. This error control system can be used as an alternative to Automatic Repeat reQuest (ARQ) or Forward Error Correction (FEC). Satisfactory RD performances are achieved at the side decoders, the Wyner-Ziv data improving significantly the quality of the individual descriptions. However, when used as a FEC mechanism and when the two descriptions are received, the Wyner-Ziv data turns out to be completely redundant and does not contribute to improving the quality of the signal reconstructed by the central decoder.

To address the above limitation, we have started investigating the problem of multiple description coding with side information. The achievable rate-distortion region for multiple description coding has been recently established in the literature, when common side information about a correlated random process is known to each multiple description decoder. To approach the bounds of this rate-distortion region, we are working on the design of a coding algorithm based on multiple description quantization with side information. The approach will find 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. *Error-resilient source codes and soft decoding*

Participants: Christine Guillemot, Simon Malinowski.

In 2006, 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. So far, two types of trellises have been considered to estimate the sequence of emitted symbols from the received noisy bitstream: the bit-level trellis proposed by Balakirsky (initially proposed for variable length codes) and the bit-symbol trellis. The bit-level trellis leads to decoders of low complexity, however can not exploit symbol *a priori* information (e.g., termination constraint on the number of symbols sent), and hence suffers from some sub-optimality. In contrast, the bit-symbol trellis allows to exploit *a priori* information on the sequence of symbols and, coupled with the BCJR algorithm, it allows to obtain sequence estimates minimizing the Bit Error Rate (BER) and the Symbol Error Rate (SER). However, the number of states of the bit/symbol trellis is a quadratic function of the sequence length, leading to a complexity not tractable for realistic applications. We have thus developed a novel set of state models of lower complexity and the corresponding trellises for the estimation of the Hidden Markov chain. The state model is defined by both the internal state of the VLC decoder (i.e., the internal node of the VLC codetree) or the state of the decoding automaton of QA codes and the rest of the Euclidean division of the symbol clock by a fixed parameter T . Therefore, 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 to gracefully trade complexity against the estimation accuracy. The state aggregation leads to close-to-optimum estimations with significantly reduced complexity. For a given VLC or QA code, a method to estimate the value of the parameter T that yields close to optimum performance (i.e. decoding performance close to the one obtained on the bit-symbol trellis) has also been developed. This method is based on the analysis of the synchronisation properties of the VLC or QA codes. Minimum frame and bit error rates can be obtained by running respectively a Viterbi algorithm or a BCJR algorithm on the aggregated models.

In contrast with the optimal models, the aggregated state models do not keep track of the symbol clock values. Hence, symbol A Posteriori Probabilities (APP) do not come as a natural product of the estimation with aggregated state models. Yet, symbol APP are required to minimize the SER, as well as to use the estimation algorithm in an iterative source-channel decoding structure. In 2007, we have thus developed a low complexity method to compute symbol APP on aggregated state models used for soft decoding of sources encoded either with VLCs or with QA codes. Simulation results reveal that the SER obtained with the reduced complexity algorithm is the same as the one achieved with the optimal state model.

One critical aspect of VLC and QA is their very high sensitivity to transmission noise. The robustness of these codes can be increased by adding redundancy under the form of side information to the transmitted bitstream. This *a priori* information can be used to favor the selection of synchronous and correct paths in the soft decoding process run on the aggregated state model. 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 here does not lead to any modification of the compressed bit-stream. The side information can be transmitted separately. 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.

6.3.3. *Sparse representation based error decoding*

Participants: Jean-Jacques Fuchs, Gagan Rath.

Joint source-channel coding (JSCC) with real number BCH codes provides error resilience through overcomplete expansions of input signals. Such a framework, however, is associated with non-unique signal solutions when the expanded signals are corrupted by noise. Typically, under an additive noise model, the unknown noise parameters satisfy an under-determined system of linear equations which has infinite number of solutions. In this case, the usual procedure is to assume that most of the received signal components are uncorrupted, which, under the maximum likelihood decoding, results in a solution with the least number of non-zero elements. The decoding algorithm aims to exploit the imposed coding structure so as to arrive at this solution at reasonable complexity. The standard coding theoretic solutions, however, are not satisfactory due to the presence of the quantization noise. A possible alternative is to solve the under-determined system of equations without considering the underlying code structure.

The general problem of solving an under-determined system of equations is a well-known problem in the field of sparse approximation. The problem can be reformulated as approximating a known signal as a linear sum of a few elementary signals. These signals are drawn from a collection known as a dictionary, whose size is quite large as compared to the signal dimension and which is linearly dependent. Finding the solution with minimum number of non-zero elements can be termed as the sparsest solution with L_0 norm. This problem is known to be NP hard, and except for the exhaustive combinatorial approach, there is no known method for the exact solution. There are some well-known approaches such as matching pursuit (MP), orthogonal matching pursuit, etc., which find an approximate heuristic solution with tractable complexity. An alternative approach is to relax the L_0 norm condition by L_1 norm which helps in solving the problem through linear programming. Such solutions are categorized under basis pursuit (BP) algorithms. These algorithms can be applied in the above JSCC framework to decode the channel errors.

6.3.4. Parameter estimation in convolutive and instantaneous mixtures

Participant: Aline Roumy.

Many receivers need first to estimate some parameters before processing the received data. These parameters can be related to the sources (correlation between sources) or to the transmission (SNR, multipath coefficients). We have developed generic tools in order to evaluate the performance degradation due to channel imperfect knowledge at the receiver.

In wireless communications, the transmission channel introduces time-varying multipath fading to the transmitted signal and hence, a convolutive mixture of the transmit data is built where the convolution coefficients are the multipath coefficients of the channel. Therefore, an equalizer is needed to recover the transmitted data at the receiver. The optimal equalizer to be used is based on maximum *a posteriori* (MAP) detection and depends on the transmission channel, which is a priori unknown. Therefore, the receiver contains a channel estimation algorithm to estimate a proper channel parameter set. In this context, we have designed a new equalizer robust to channel estimation errors. This equalizer outperforms usual equalizers that estimates the channel and uses this estimate as the true channel. More surprisingly, our simulations show that this new equalizer get the same performance as the perfect channel knowledge equalizer [50].

Moreover, turbo-equalizers have been proposed to take into account the fact that the data are coded. They contain a MAP equalizer fed with *a priori* information on the transmitted data and provided by another module in the receiver, for instance the decoder. This has motivated our study of the impact of channel estimation and *a priori* information in a maximum a posteriori (MAP) equalizer. We have first considered the case where the MAP equalizer is fed with a priori information on the transmitted data and studied analytically their impact on the MAP equalizer performance. We have shown that the use of the a priori information is equivalent to a shift in terms of signal-to-noise ratio (SNR) for which we have provided an analytical expression. This study has been performed for MIMO equalizers [32] and for large size modulations (M-PSK) [33]. The behavior of the whole turbo-equalizer (with perfect channel knowledge) has also been studied. We have performed an analytical convergence analysis of turbo equalizers using MAP equalization and derived a condition for the turbo equalizer to converge to the Gaussian performance in the high SNR regime [28]. This work has been carried out in collaboration with N. Sellami (ISECS, Sfax, Tunisia), S. Chaabouni (SupCom, Tunis, Tunisia), I. Fijalkow (ENSEA, Cergy, France) and M. Siala (SupCom, Tunis, Tunisia), in the context of a CNRS-DGRSRT project.

6.3.5. Robust estimation and detection

Participant: Jean-Jacques Fuchs.

In some communication systems, the information is carried by a signature whose exact value needs to be exactly known to recover the transmitted information.

The basic model is the following, at each time instant t one observes

$$y_t = x - t + n_t, \quad \text{with: } x_t = a s_t,$$

where $a \in R^n$ is the signature, $\{s_t\}$ is the scalar sequence representing the information to be detected and estimated and e_t is the noise.

The case where the signature a is not precisely known has often been investigated. Depending upon the type of communications system that is considered, the uncertainty on the signature may be due to many different reasons such as perturbations in the transmission channel, bad calibrations, diffraction, deformations or measurement noise. We develop a detection scheme that is robust to a wide variety of such perturbations.

In addition to the previous relation, we assume that at the beginning of the transmission of the N snapshots $\{y_t\}$, one observes or knows b :

$$b = a + \alpha e,$$

a noisy version of the true signature a . in this relation e is a gaussian noise vector whose variance is known up to a multiplicative constant. The problem to be solved is then the following: knowing the $N+1$ vectors $\{y_t\}$, b decide if a signal $\{s_t\}$ has been transmitting and, if it is the case, estimate it. The previous model amounts to assuming that the true signature lies in a cone whose aperture depends upon α that thus allows to tune the degree of uncertainty. We have considered previously the case of a single snapshot ($N=1$) [15] and now extend it to the multi-snapshot case [40]. The techniques to be used are similar but the resulting algorithm quite different.

6.4. Distributed source and source-channel coding

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

6.4.1. Distributed source codes based on punctured and overlapped quasi arithmetic codes

Participants: Christine Guillemot, Simon Malinowski.

Distributed source coding (DSC) finds its foundation in the seminal Slepian-Wolf and Wyner-Ziv theorems. Most Slepian-Wolf and Wyner-Ziv coding systems are based on channel coding principles. The statistical dependence between the two sources is modeled as a virtual correlation channel analogous to binary symmetric channels or additive white Gaussian noise (AWGN) channels. The source Y (called the side information) is thus regarded as a noisy version of X (called the main signal). All the approaches based on channel codes assume memoryless sources. However, in practice the signals considered have memory. Here, we consider the design of Slepian-Wolf codes based on source codes and in particular on quasi-arithmetic codes, to try to capture the source memory in addition to the correlation between the different sources.

In conventional arithmetic coding, to encode a symbol, the current interval is partitioned into sub-intervals that do not overlap. The coded stream transmitted at a rate close to the entropy of the input symbol sequence is then uniquely decodable. Building upon our results on robust and soft decoding of arithmetic and quasi-arithmetic codes, we have investigated two new Slepian-Wolf coding strategies. 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 idea is to allow the intervals corresponding to each source symbol to overlap. This procedure can achieve arbitrary compression controlled by the amount of allowed overlap, but it leads to a code which is not uniquely decodable any more. 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 overlapping technique. The SER performance with respect to the cross-over probability of the side information has been assessed. The performance of the proposed distributed coding schemes is not to the level of the one obtained with turbo-codes. Turbo coding and decoding structures based on the punctured and overlapped quasi-arithmetic codes are currently being designed in order to reduce the gap with classical turbo-codes.

6.4.2. *Rate-adaptive asymmetric and symmetric Slepian-Wolf coding*

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

As said above, most Slepian-Wolf and Wyner-Ziv coding systems are based on channel coding principles. The statistical dependence between the two sources is modeled as a virtual correlation channel analogous to binary symmetric channels or additive white Gaussian noise (AWGN) channels. Capacity achieving channel codes can then be turned into optimal Slepian-Wolf codes. It is only recently with the advances in channel coding that practical Slepian-Wolf coding/decoding solutions have been proposed. Two main approaches can be found in the literature: the parity approach and the syndrome approach. The syndrome approach is optimal, but on the other hand the parity approach is more amenable to puncturing for rate adaptation purposes, in real scenarios where the correlation can vary in time. We have here addressed the problem of rate adaptive syndrome-based Slepian-Wolf coding with turbo codes. We have first considered the case of syndrome-punctured convolutional codes. A brute force optimal decoding algorithm would have a complexity growing exponentially with the number of punctured positions, since the search of the closest sequence has to be performed in a union of cosets. So, we have derived an optimal algorithm with a complexity that grows only linearly with the number of punctured positions. The idea has been generalized to the case of turbo-codes [49]. We are currently working on Wyner-Ziv and error resilient codecs based on our syndrome puncturing approach.

Most of Slepian-Wolf codecs proposed in the literature consider the asymmetric case where one source is assumed to be perfectly known to the decoder, that is transmitted at its 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 context of joint optimization of source rate and transmission power in networked sensors applications. We have thus also 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. Such a scheme is needed 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.

6.4.3. *Optimum Filters for Wyner-Ziv predictive coding*

Participants: Christine Guillemot, Gagan Rath.

This study has been done in collaboration with Jayanth Nayak from the University of California at Riverside. Predictive coding is an efficient technique for compressing sources with memory and it has been successfully employed in many domains of communications, compression, etc. The most common approach is to apply a linear prediction filter which aims to remove the temporal or spatial correlation from the input signal. This filter is typically designed by minimizing the variance of the residual signal, whose rate distortion function is equal to the rate distortion function of the source under Gaussianity assumption.

In a distributed coding scenario, the source with memory is available at the encoder, but the correlated source (with memory) is available at the decoder. Predictive coding in such a scenario is similar to the conventional predictive coding, but with an additional constraint that the residual be optimally correlated to the residual (from a predictive coding of the correlated source) at the decoder. So the problem basically consists of designing two linear filters, one for each source, so that the autocorrelation at the encoder is minimized but the cross-correlation between the two residuals is maximized. These two objectives are fulfilled by designing the filters by minimizing a judiciously selected objective function. It turns out that, under Gaussianity assumption, the solution filters are identical to the case when the two correlated sources are available at the encoder. The optimum filters are thus derived using the traditional linear prediction theory.

6.4.4. *Correlated sources over multiple access channels*

Participants: Christine Guillemot, Aline Roumy.

In collaboration with Onkar Dabeer (Tata Institute of Fundamental Research, India), the problem of transmitting correlated sources over the MAC channel has been studied. In this context, the source-channel separation theorem does not hold. We have therefore focused on the design of joint source-channel schemes. We have first restricted ourselves to the class of linear processes in order to propose low processing complexity algorithms. First, we have studied best linear joint source-channel transceivers for transmitting two correlated Gaussian memoryless sources over a Gaussian MAC. We have shown that, when the bandwidth expansion factor is one, uncoded transmission is the best linear code for *any* SNR. When the bandwidth expansion factor is two, we exhibited a good linear code that outperforms any TDMA based strategy [36]. Then, we addressed the problem of sources with memory. We considered a single colored source being observed by two sensors with independent noises (this problem is commonly referred to as the CEO problem) and derived closed-form solution to the optimal linear transceiver. More precisely, we have shown that when the source is white, uncoded transmission is the best linear code for any SNR. But for a colored source, whitening transmit filter is sub-optimal [37].

In a second approach, we have studied non-linear code design for the MAC channel. In this work, the two sources are assumed to be independent. Codes (and their related iterative decoding algorithms) that are able to get close to the boundary of the capacity region of the Gaussian multiple access channel, without the use of time sharing or rate splitting, have been designed. The approach is based on density evolution and/or its variant (mean, mutual information). LDPC codes have been designed for the 2-user Gaussian MAC. We have shown that it is possible to design good irregular LDPC codes with very simple techniques, the optimization problem being solved by linear programming [26]. We are currently working on the design of non-linear codes for correlated sources sent over a MAC channel.

6.4.5. *Distributed coding for sensor networks*

Participant: Aline Roumy.

In collaboration with M. Debbah (SUPELEC, Gif-sur-Yvette, France) A. Kherani and T. Banerjee (Dept. of Computer Science and Engineering, IIT Delhi, India), the problem of optimal wireless node density of a sensor network has been addressed. We have considered a distributed and separated access scheme (at each sensor separate source and channel coding is performed) and investigated the tradeoff between accuracy of the field investigated (taking into account the correlation of the sources) and the communication cost. A “good” operating point for this network, which maximizes the information gathered with a constraint on the network budget (including the number of nodes and their transmission powers) has been determined. A comparative study of the system performance obtained under various transmission/reception schemes and fading coefficient statistics [31] has been made. This work is supported in part by the Network of Excellence in Wireless Communications (NewCom), and by the INRIA-ARC project InFormAtioN theory (IFANY).

In collaboration with D. Gesbert (Eurecom Institute, Sophia-Antipolis, France) under the umbrella of the NewCom project, the problem of rate and power allocation in a wireless sensor network (WSN) has been studied in order to send without loss the data gathered by the nodes to a common sink. Correlation between the data and channel impairments dictate the constraints of the optimization problem. We assumed that the

WSN uses off-the-shelf compression and channel coding algorithms. More precisely source and channel coding are separated and distributed source coding (DSC) is performed by pairs of nodes. This raises the problem of optimal nodes matching. We have shown that under all these constraints the optimal design (including rate/power allocation and matching) has polynomial complexity (in the number of nodes in the network). A closed form solution is given for the rate/power allocation, and the matching solution is readily interpreted. For noiseless channels, the optimization matches close nodes whereas, for noisy channels, there is a tradeoff between matching close nodes and matching nodes with different distances to the sink. This fact is illustrated by simulations based on empirical measures. We have also shown that the matching technique provides substantial gains in either storage capacity or power consumption for the WSN wrt the case where the correlation between the nodes is not used [48], [27].

6.4.6. Distributed video compression

Participants: Christine Guillemot, Denis Kubasov, Gagan Rath.

A distributed video coding and decoding algorithm is being developed in the context of the European IST-Discover project. 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 2007, we have addressed the above issues in order to increase the codec rate-distortion performance. A new multihypotheses decoding approach has first been developed, which allows using multiple side information (SI). The advantage of using multiple SI is that the decoding becomes less sensitive to occasional side information defaults resulting from the limitations of motion-compensated interpolation in particular situations such as for example occluding regions. Another problem which has been addressed is the problem of rate control. Precise rate allocation requires an accurate modelling and estimation of the virtual correlation channel. The virtual channel estimation can be performed at the decoder from the previously received data. However, the rate control in this case requires a feedback channel to then control the rate of the Slepian-Wolf code. This is usually done by requesting syndrome or parity bits with some implications on latency and decoder complexity. We have instead developed a hybrid encoder-decoder rate control approach. A minimum rate of the Slepian-Wolf code is computed based on the entropy of the estimated bitplane crossover probability. This minimum rate value, although not fully accurate, is shown to be a relatively good estimate of the actual rate needed for the SW code. This approach drastically reduces the decoding time with only a small negative effect on the R-D performance. The last problem which has been addressed is the problem of optimal MMSE (Minimum Mean Square Error) signal reconstruction given some side information. Closed-form expressions of the optimal MMSE estimator have been derived for the particular case the correlation model is assumed to be Laplacian. These expressions have been derived both for single and multiple SI scenarios. This work has been done in collaboration with Jayanth Nayak, University of California at Riverside, USA. It has received the best student paper award at the IEEE MultiMedia Signal Processing Workshop, MMSP, Oct. 2007 (see section 9.3).

6.5. Image and video watermarking and security

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

In 2007, we have focused on the problems of *watermark detection*, of *fingerprinting*, that is, on the design of robust techniques that embed a different mark in a content, according to the users's identity. At last, we have proposed an improvement of *steganographic* schemes, which aim at hiding the communication itself, so with a constraint in terms of furtivity.

6.5.1. Optimal watermark detection

Participant: Teddy Furon.

Zero-bit watermarking (also known as watermark detection) is a useful framework for applications requiring high robustness. In 2006, we have developed a constructive framework for creating new watermarking schemes. At the detection side, the approach resorts to Locally Most Powerful (LMP) test. Therefore, for a given embedding function, the general expression of the LMP test gives us the best detection function. For a given detection function, we look for the embedding function which maximizes the asymptotic relative efficacy. We have derived closed form expressions for the optimum ends of the watermarking chain when the other end is given. This gives rise to a partial differential equation. The solutions are heavily dependent on the probability density function of the host signal.

In 2007, we have derived a practical watermarking scheme from the recent theoretical development of N. Merhav. This is a first step to so-called universal watermark detection whose performances do not degrade in presence of geometrical attacks. Some preliminary tests on images have been done where local invariant descriptors around salient points are watermarked. More extensive experiments are required to assess the global robustness of the approach. It is foreseen that while the watermark is robust against pure geometrical attacks, it is not the case when facing value-metric attacks.

6.5.2. *Watermarking Security*

Participants: Teddy Furon, Caroline Fontaine.

In 2004 and 2005, we have developed a theoretical framework for assessing the security level of watermarking schemes. This framework is based on the measure of information about the secret key that leaks from the watermarked contents observed by an opponent. The security levels of two well-known watermarking schemes, substitutive and additive spread spectrum techniques, have been analyzed. In 2006, we have analyzed the security levels of Quantized Index Modulation (QIM) based watermarking techniques. In QIM, a set of nested lattices is defined, each of them being associated to a symbol. The watermarked signal is (or is attracted towards) the quantized version of the host signal onto the lattice related to the hidden symbol to be transmitted. The problem is that the optimal performances are reached for well-known lattices. Therefore, no security is guaranteed. Nevertheless, it appeared that performances are not degraded when the lattices are dithered (ie. geometrically shifted) by a secret vector, shared by the embedder and the decoder, thus improving the level of security of the approach. However, the question of how long the secret dither will remain a real secret has never been investigated. We have first studied the mutual information between the dither and a set of watermarked contents (watermarked with the same secret dither). We have derived bounds which provides a lower (ie. pessimistic for the watermark designer) estimation of the security levels, i.e., of the number of watermarked contents needed to accurately estimate the secret dither. But, they do not give any clue concerning the estimation algorithm the opponent shall run, and especially its complexity.

Therefore, a practical algorithm with an affordable complexity has been designed. We have used a tool from the automatic, system identification community, so-called Set Membership Estimation (SME). Briefly, the watermarked signal can be regarded as a noisy observation of the dither. But, the noise has a bounded support given by the lattice Voronoi cell. Thus, one observation gives a bounded feasible set of values for the dither. Observing a group of watermarked contents, the opponent estimates the dither by finding the intersection of all the feasible sets. However, this is not an easy task as the description of this intersection requires an increasing number of parameters as the number of observations grows. Part of the work has been supported by ACI Nebbiano, and part of the work has been done in collaboration with the University of Vigo (Spain). In the area of security, we have also worked on the design of a dedicated architecture to enable a good cooperation between traditional DRM tools and digital watermarking, focusing on a new interaction between the SIM card and the watermarking schemes through cryptographic protocols.

6.5.3. *Fingerprinting*

Participants: Caroline Fontaine, Fabien Galand, Fuchun Xie.

Fingerprinting aims at hiding in a robust and imperceptible way, data which will differ for different legitimate users of the content. The goal is to enable traceability of the content and to resolve frauds. We have first focused on the choice of the embedded data that would enable best tracing capability. This choice is hard

because we have to face collusion attacks, in which colluders compare their contents in order to forge a new and untraceable one. So far, the problem has been addressed mostly by the error correcting codes' community, without any link with the signal processing community which is focusing on the embedding problem. This results in solutions that are not reliable in real life applications, for several reasons. First, the model of attack considered so far is not realistic at all. Second, the codes proposed in this context present too many constraints that make them unefficient: large alphabets, huge lengths, slow decoding algorithms. Hence, we have initialized a study of error correcting codes working with the Euclidean distance to see if they can be used directly with embedding techniques. At the same time, we are investigating models taking into account more realistic attacks.

6.5.4. Steganography

Participants: Caroline Fontaine, Fabien Galand.

Steganography aims at sending a message through a cover-medium, in an undetectable way: nobody, except the intended receiver of the message, should be able to tell if the medium is carrying a message or not. The matrix embedding approach relying on error correcting codes is often used in steganography since it modifies a small number of components. It provides an effective answer to the adaptive channel selection problem: the sender can embed the message in a way which depends on the cover-medium to minimize the distortion, and the receiver can extract the messages without being aware of the sender's choices.

Random error correcting codes may seem interesting for their asymptotic behavior, however they require solving hard problems: syndrome decoding and covering radius computation are NP-complete and Π_2 -complete respectively. Moreover, no efficient decoding algorithm is known, even for a small non trivial family of codes. From a practical point of view, this implies that the related steganographic schemes are too complex to be considered for real applications. Hence, it is of great interest to have a deeper look at other kinds of codes. We have shown that Reed-Solomon codes are good candidates for designing realistic steganographic schemes. If we compare them to BCH codes, Reed-Solomon codes improve the management of locked positions during embedding, hence ensuring a better control of the distortion: they are able to lock twice the number of positions. We proposed two methods based on these codes: the first one is based on a naive decoding process through Lagrange interpolation; the second one, more efficient, is based on the Guruswami-Sudan list decoding and allows us to control the trade-off between the number of locked positions and the embedding efficiency.

6.5.5. 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. Here are two application scenarii where a wrong estimation of the probability of error could give a disaster.

Copy protection. Assume commercial content are encrypted and watermarked and that future consumer electronics storage device have a watermark detector. These devices refuse to record a watermarked content which is not encrypted. The probability of false alarm is the probability that the detector considers an original piece of content (which has not been watermarked) as protected. The movie shot by a user during his holidays could be rejected by his storage device. This absolutely non user-friendly behavior really scared consumer electronics manufacturers. In the past, the Copy Protection Working Group of the DVD forum evaluated that, at most one false alarm should happen in 400 hours of video. As the detection rate was one decision per ten seconds, this implies a probability of false in the order of 10^{-5} . An accurate experimental assessment of such a

low probability of false alarm would demand to feed a real-time watermarking detector with non-watermarked content during 40,000 hours, ie. more than 4 years! Proposals in response of the CPTWG's call were, at that time, never able to guarantee this level of reliability.

Fingerprinting. In this application, users' identifiers are embedded in purchased content. When content is found in an illegal place (*e.g.* a P2P network), the right holders decode the hidden message, find a serial number, and thus they can trace the traitor, *i.e.* the client who has illegally broadcast their copy. However, the task is not that simple as dishonest users might collude. For security reason, anti-collusion codes have to be employed. Yet, these solutions have a non-zero probability of error defined as the probability of accusing an innocent. This probability should be, of course, extremely low, but it is also a very sensitive parameter: anti-collusion codes get longer (in terms of the number of bits to be hidden in content) as the probability of error decreases. Fingerprint designers have to strike a trade-off, which is hard to conceive when only rough estimation of the probability of error is known. The major issue for fingerprinting algorithms is the fact that embedding large sequences implies also assessing reliability on a huge amount of data which may be practically unachievable without using rare event analysis.

We are currently working with the project-team ASPI within the framework of the national project ANR-Nebbiano. We apply their rare event analysis tool, so-called Adaptive Multilevel Splitting, to these two scenarios: We have been able to estimate probabilities around 10^{-19} for watermark detection, and 10^{-8} for fingerprinting scheme. This work is not only an application of the tool, it also appeals for more theoretical development in order to set the optimal value of the parameters, which, for the moment, we tune 'manually'. Optimality here will be expressed in terms of estimation bias and variance, convergence and complexity.

7. Contracts and Grants with Industry

7.1. Industrial contracts

7.1.1. CRE with France Telecom on 3D reconstruction of urban scenes

Participants: Luce Morin, Gael Sourimant.

- Title : 3D reconstruction of urban scenes by fusion of GPS, GIS and video data.
- Research axis : § 6.1.1.
- Partners : France Télécom, Irisa/Inria-Rennes.
- Funding : France Télécom.
- Period : Oct.04-Sept.07.

This contract with France Telecom R&D (started in October 2004) aims at investigating the fusion of multi-modal data from video, GPS and GIS for 3D reconstruction of urban scenes. Video and GIS data 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. A GPS acquisition synchronized with video acquisition is added in order to provide a rough estimation of camera pose in a global coordinate system.

In 2007, we have focused on two issues. First, an automatic scheme for initial pose estimation has been studied, as GPS measures only provide camera position and no information on camera orientation. The proposed algorithm relies on identifying the GPS based translation and the translation estimated from self-calibration. Pose refinement is then performed through 2D/3D line correspondences embedded in a Ransac robust estimation scheme. Second, we have focused on texture extraction from the video data, and its mapping on the GIS 3D model. By using a robust selection algorithm for combining texture from different view-points, unmodeled objects such as trees or cars are detected and removed.

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 2007, we have addressed the problem of spatial image prediction in highly textured areas, and developed prediction methods based on sparse signal approximations. The method has been integrated in the JVT (ITU/MPEG joint video team) KTA software for validation.

7.1.3. Contract with Thalès Communications on localization of correlated sources in the high frequency band.

Participant: Jean-Jacques Fuchs.

- Title : Alibaba.
- Funding : Thalès Communications.
- Period : March.07- Dec.07.

This is a contract running over a 10 months period between Irisa/Université de Rennes 1 and Thalès Communications. It started on March 15, 2007 and concerns the evaluation of the performances of a source localization algorithm developed by Irisa several years ago, when applied to potentially correlated sources in the high frequency band (3-30 MHz).

7.2. National contracts

7.2.1. RNRT-COSINUS

Participant: Laurent Guillo.

- Title : COSINUS: Real-time IP service communication in a wireless network
- Research axis : § 5.5.
- Partners : Alcatel CIT, Institut EURECOM, IRISA /INIRIA Rennes, France Télécom, GET/ENST Bretagne, Thales Communications.
- Funding : Ministry of industry.
- Period : Dec. 04 - Sept. 07.

The main objective of the COSINUS project is to demonstrate the feasibility of real time services on IPv6 wireless networks (UMTS or WLAN). It addresses the following issues: Controlling the quality as perceived by the user, accounting for the specific nature and quality of the wireless link, managing the diversity of access networks (UMTS, WLAN). In this perspective, the project partners study the following technical aspects: header compression protocols (notably ROHC), unequal error protection (UEP) techniques, audio and video source encoding that is resilient to radio errors and self-adaptive for bit-rates, perceived quality assessment methods.

The TEMICS project-team contributes on the issue of video streaming with resilience and QoS support on UMTS links. In 2007, year of conclusion of the project, the effort has been dedicated to the assessment of the techniques in the UMTS platform of the project.

7.2.2. *RIAM-ESTIVALE*

Participants: Caroline Fontaine, Teddy Furon, Fabien Galand.

- Title : Secured exchanges for video transfer
- Research axis : § 6.5.3.
- Partners : LIS (INPG), ADIS (Univ. Paris XI), CERDI (Univ. Paris XI), LSS (Univ. Paris XI/Supelec), Basic-Lead, Nextamp, SACD.
- Funding : ANR.
- Period : 31/03/2006-31/03/2009

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.

TEMICS's role was, first, to contribute to the elaboration of scenarii for the delivery of videos over networks, and, second, to study their level of security to make these scenarii possible in real life applications, involving fingerprinting techniques (see § 6.5.3). Fingerprinting solutions have been studied including with real images to confront theoretical models with real life applications. As explained in § 6.5.3, we have shown that the main model proposed in the literature is not sufficiently secure to be used in realistic applications.

7.2.3. *ACI Nebbiano*

Participant: Teddy Furon.

- Title : Nebbiano
- Research axis : § 6.5.5, § 6.5.2.
- 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. It is decomposed into 3 axis of research. Two of them stem from the past national project ACI FABRIANO: Investigations of the role of the Independent Component Analysis in data hiding, and the design of new secure and robust watermarking techniques. A third topic has been introduced in the above section: the reliability of watermarking techniques, especially the assessment of extremely low probability of errors.

7.2.4. *ACI Masse de données - Codage*

Participants: Olivier Crave, Christine Guillemot, Simon Malinowski.

- Title : Coding of large volumes of data
- Research axis : § 6.3.1.
- Partners : ENST-Paris, INRIA (TEMICS), I3S Université de Nice-Sophia Antipolis.
- Funding : ANR.
- Period : Mid-Dec. 03 - Dec. 06.

The objective of this project is to federate research effort in the two following areas:

1. Motion-compensated spatio-temporal wavelet (MCSTW) scalable coding: Tools for scalability available in existing standards usually lack compression efficiency, and are not flexible enough to achieve combination of different scalability dimensions (e.g. spatial, temporal, SNR, object and complexity scalability) and sufficient fine granularity. MCSTW offers the ideal framework for scalable compression of video sequences. Precise research tasks include scalable motion estimation and coding methods, non-linear adaptive wavelet decompositions, more appropriate for representing temporal residuals, techniques for progressive transmission of information (embedded coding, multiple description coding, ...).
2. Distributed source video coding: Traditional predictive coding, exploiting temporal correlations in a sequence through computational-intensive motion estimation between successive frames leads to encoders with a complexity 5 to 10 times higher than the complexity of the decoders. This is well suited to streaming or broadcasting applications, but not to a transmission from a mobile terminal to a base station or for peer-to-peer mobile communications. The project is investigating multi-terminal and distributed source coding solutions building upon dualities with multiple description coding and with channel coding with side information.

7.2.5. ANR-COHDEQ 40

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 aims at establishing the feasibility of coherent detection in optical fibers transmission systems. As far as Irisa is concerned, the work will be done by ASPI and TEMICS.

7.2.6. ANR-ESSOR

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

- Title : Distributed Video Coding
- Research axis : § 6.3.1.
- 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.

7.2.7. ANR-ICOS-HD

Participants: Christine Guillemot, Joaquin Zepeda.

- Title : Scalable Indexing and Compression Scalable for High Definition TV;
- Research axis : § 6.2.4.
- Partners : Université de Bordeaux, CNRS/I3S;
- Funding : ANR.
- Period : 01/01/2007-31/12/2009

The objective of the project is to develop new solutions of scalable description for High Definition video content to facilitate their editing, their access 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. These descriptors must allow the collection of information related to actions, events or activities taking place in the video document and which can happen at different spatial and temporal resolutions. 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 6.2.4).

7.3. European contracts

7.3.1. FP6-IST NoE SIMILAR

Participants: Luce Morin, Gael Sourimant.

- Title: *European research taskforce creating human-machine interfaces SIMILAR to human-human communication.*
- Research axis: § 6.1.1
- Partners: around 40 partners from 16 countries.
- Funding: CEE.
- Period: Jan.04-Dec.07.

The TEMICS project-team is involved in the network of excellence SIMILAR federating European fundamental research on multimodal human-machine interfaces and contributes on the following aspects:

- In the context of 3D modelling of video sequences we have studied the fusion of multimodal data (e.g. real 2D video, synthetic 3d models) in the context of urban scenes automatic modelling. We have also developed an interface for interactive visualization of 3D videos, eg videos enhanced with 3D information. Depth information is given as a set of 3D triangular meshes, one model providing depth information for a sub-set of successive frames in the video. Streaming of such 3D videos for scalable transmission and distant visualization is being experimented.
- TEMICS has also participated to the SIMILAR workshop Interface'07, with a full-time participant to the project "Advanced multimedia interfaces for flexible communications".

7.3.2. FP6-IST NOE NEWCOM

Participant: Aline Roumy.

- Title: NEWCOM: Network of Excellence in Wireless Communication.
- Research axis: § 6.3.4
- Funding: CEE.
- Period: March 2004 - March 2007.

The NEWCOM project (Network of Excellence in Wireless COMmunication) addresses the design of systems "beyond 3G". This requires to successfully solve problems such as: the inter-technology mobility management between 3G and ad-hoc wireless LANs, the coexistence of a variety of traffic/services with different and sometimes conflicting Quality of Service (QoS) requirements, new multiple-access techniques in a hostile environment like a channel severely affected by frequency selective fading, the quest for higher data rates also in the overlay cellular system, scaling with those feasible in a wireless LAN environment, permitting seamless handover with the same degree of service to the user, the cross-layer optimisation of physical coding/modulation schemes with the medium access control (MAC) protocols to conform with fully packetised transmission as well as the TCP/IP rules of the core network, and the like. In 2007, we have, in collaboration with David Gesbert from Eurecom, Sophia-Antipolis, studied the optimal matching problem in a wireless sensor network. In collaboration with Merouane Debbah from Supélec, Gif-sur-Yvette, we have studied the optimal wireless node density of a sensor network. A new project called Newcom++ has been submitted in the FP7 framework. This network of excellence has been accepted but the contract is still under negotiation.

7.3.3. NoE-ECRYPT

Participant: Caroline Fontaine.

- Title : ECRYPT
- Research axis : § 6.5.2, § 6.5.3, § 6.5.4.
- Partners : 32 teams across Europe;
- Funding : CEE
- Period : Feb. 2004 - July 2008

ECRYPT aims at funding international meetings and events, to make European research in the area of cryptography and data hiding more dynamic. It is splitted into several virtual labs, the one concerning us being called WAVILA: WATERmarking VIRTual LAB. The TEMICS project-team has organized in June 2007 the annual WAVILA meeting WACHA in Saint-Malo, jointly with Information Hiding'07.

7.3.4. FP6-IST STREP DISCOVER

Participants: Christine Guillemot, Denis Kubasov, Khaled Lajnef, Gagan Rath.

- Title: Distributed Coding for Video Services
- Research axis: § 6.4.2, § 6.4.3, § 6.4.6
- Universitat Politècnica de Catalunya (UPC), Instituto Superior Técnico (IST), Ecole Polytechnique Fédérale de Lausanne (EPFL), Universität Hannover (UH), Institut National de Recherche en Informatique et en Automatique (INRIA-Rennes) Università di Brescia (UB).
- Funding: CEE.
- Period: Sept.05-Aug.07.

Video coding solutions so far have been adopting a paradigm where it is the task of the encoder to explore the source statistics, leading to a complexity balance where complex encoders interact with simpler decoders. This paradigm is strong. The objective of DISCOVER is to explore and to propose new video coding schemes and tools in the area of Distributed Video Coding with a strong potential for new applications, targeting new advances in coding efficiency, error resiliency, scalability, and model based-video coding.

The TEMICS project-team is coordinating - and contributing to - the workpackage dealing with the development of the theoretical framework and the development of Wyner-Ziv specific tools. In 2007, we have in particular designed optimum prediction filters for the Wyner-Ziv scenario (see Section 6.4.3), developed Slepian-Wolf codes based on quasi-arithmetic codes (see Section 6.4.1 and on channel codes (see Section 6.4.2). The TEMICS project-team also contributes to the development of algorithmic tools for the complete coding/decoding architecture and to the integration of the complete video codec. In that context, methods of side information extraction, of rate control and optimal MMSE signal reconstruction in presence of side information have been developed (see Section 6.4.6).

8. Other Grants and Activities

8.1. International initiatives

Participant: Aline Roumy.

- Title : Radio resource optimization in iterative receivers
- Research axis : § 6.3.4.
- 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. Invited talks

- C. Guillemot and A. Roumy gave an invited tutorial of 3 hours on joint distributed video compression, at the European conference on Signal Processing, Sept. 2007.
- A. Roumy was invited to present her joint work with M. Debbah on the optimal node density of a wireless sensor network, in the Newcom dissemination day conference, Paris, Feb. 2007.

9.2. Leadership within the scientific community

- C. Fontaine is associate editor of the Journal in Computer Virology (Springer-Verlag);
- C. Fontaine was a member of program committees of the following conferences: WCC 2007 (Rocquencourt, France, March), SSTIC 2007 (Rennes, France, June), Wacha 2007 (Saint-Malo, France, June), CORESA 2007 (Montpellier, France, October);
- C. Fontaine was a member of the organizing committees of the following conferences: WCC 2007 (Rocquencourt, France, March), SSTIC 2007 (Rennes, France, June), Wacha'07 (Saint-Malo, 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 : SAM2007 (Sensor Array and Multichannel Signal Processing Workshop), Eusipco 2007, GretsI 2007;
- 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 was the general co-chair of the 9th International Information Hiding (IH) conference, Saint-Malo, June 11-13, 2007;
- 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 2007 Security, Steganography, and Watermarking of Multimedia Contents IX, ACM Multimedia and Security 2007, Information Hiding 2007, European Signal Processing Conference 2007, and International Workshop on Digital Watermarking 2007;
- T. Furon was a reviewer for the 'Reproducible Research' campaign launched by IEEE Signal Processing Society;
- T. Furon is the co-organiser of the international watermarking challenge BOWS-2;
- C. Guillemot served as expert in the evaluation of research project proposals for the ministry of research of the region of Walonie in Belgium (2007) and for OSEO-ANVAR.
- 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 IMDSP (Image and MultiDimensional Signal Processing Technical Committee) and IEEE MMSP (MultiMedia Signal Processing Technical Committee) international committees;

- 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 (since Sept. 2007).
- C. Guillemot was the technical chair of the IST-Discover workshop, Lisboa, 6th Nov. 2007.
- C. Guillemot was a member of the technical program committees of the following conferences: IEEE-MMSP 2007, PCS 2007, Mobimedia 2007;
- A. Roumy participated in the IEEE Information Theory Winter School, March 2007, La Colle Sur Loup, France, where she chaired a session on code design;
- A. Roumy chaired the session Network source coding at the 2007 IEEE International Symposium on Information Theory (ISIT2007), Nice, France;

9.3. Awards

The paper "Optimal Inverse Quantisation in Wyner-Ziv Video Coding with Multiple Side Information" (D. Kubasov, J. Nayak, and C. Guillemot) has received the best student paper award at the IEEE MultiMedia Signal Processing Workshop, MMSP, Oct. 2007.

9.4. Exhibitions and demonstrations

- The TEMICS project-team has represented the INRIA research center of Rennes Bretagne Atlantique at the “fête de la science”, Rennes, Oct. 2007.
- The TEMICS project-team presented demos at the exhibition held during INRIA’s “40 years anniversary” forum held in Lille, Dec. 2007.

9.5. Invitations and visits

- M. Debbah (SUPELEC, Gif-sur-Yvette, France) visited the TEMICS project-team for one week in Jan. 2007.
- O. Dabeer (School of Technology and Computer Science, Tata Institute of Fundamental Research, Mumbai, India) visited the TEMICS project-team from May to July 2007.
- J. Garcia-Frias (Prof. University of Maryland, USA) visited the TEMICS project-team for one week in May 2007.
- K.P. Subbalakshmi (Prof. at the Stevens Institute of Technology, USA) visited the TEMICS project-team for one month in Oct. 2007.
- A. Roumy has visited Prof. K. Ramchandran at the Univ. of Berkeley, USA (Sept.-Dec. 2007).
- X. Artigas (UPC, Spain) has visited the TEMICS project-team for one week (Jan. 2007).

9.6. Teaching

- Enic, Villeneuve-d’Ascq, (C. Guillemot: Video communication) ;
- Esat, Rennes, (C. Guillemot: Image and video compression; T. Furon: Watermarking) ;
- INSA, Lyon, (C. Guillemot: Video communication) ;
- 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 STI: J-J. Fuchs teaches a course on optimization and C. Guillemot teaches a course on image and video compression ;
- Master, Security of Information Systems, Supelec-ENSTB (C. Fontaine) ;
- 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) ;
- Master of Science in Mobile Communications at Eurecom Institute, Sophia-Antipolis. (A. Roumy: Channel coding theory) ;
- A. Roumy organized a 20 hour course on “Random matrices and its application to communications” in collaboration with C. Tannoux (Human resources department, INRIA, Rennes). This course was given by M. Debbah (SUPELEC, Gif-sur-Yvette, France) in January 2007. This course was sponsored by the continued education program of INRIA.
- A. Roumy organized a 4 hour course on “Information theory: Gallagers’ proofs”. This course was given by O. Dabeer (School of Technology and Computer Science, Tata Institute of Fundamental Research, Mumbai, India) in July 2007, during his visit in the TEMICS project-team.

10. Bibliography

Major publications by the team in recent years

- [1] 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.
- [2] 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.
- [3] J.-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.
- [4] 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.
- [5] 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.
- [6] A. GUYADER, E. FABRE, C. GUILLEMOT, M. ROBERT. *Joint source-channel turbo decoding of entropy-coded sources*, in "IEEE Journal on Selected Areas in Communications, special issue on the turbo principle : from theory to practice", vol. 19, n^o 9, sept. 2001, p. 1680-1696.

- [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.

Year Publications

Doctoral dissertations and Habilitation theses

- [11] G. SOURIMANT. *3D reconstruction of video sequences for augmented reality : combining video, GIS and GPS datasets for city modeling*, Ph. D. Thesis, Thèse de doctorat en Informatique, Université de Rennes 1, France, Dec. 2007.

Articles in refereed journals and book chapters

- [12] F. CAYRE, C. FONTAINE, T. FURON. *Watermarking Security*, in "Digital Audio Watermarking Techniques and Technologies: Applications and Benchmarking", Idea Group Publishing, 2007.
- [13] 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", Nov. 2007.
- [14] C. FONTAINE, F. GALAND. *A survey of homomorphic encryption for non-specialists*, in "EURASIP Journal on Information Security, special issue on "Signal Processing in the Encrypted Domain"", to appear, 2007.
- [15] J. FUCHS. *A robust matched detector*, in "IEEE Transactions on Signal Processing", vol. 55, n^o 11, Nov. 2007, p. 5133-5142.
- [16] J. FUCHS. *Convergence of a sparse representations algorithm applicable to real or complex data*, in "IEEE Journal of Selected Topics in Signal Processing, special issue on convex optimization methods in signal processing", vol. 1, n^o 4, Dec. 2007.
- [17] J. FUCHS, S. MARIA. *A new approach to variable selection using the TLS approach*, in "IEEE Transactions on Signal Processing", vol. 55, n^o 1, Jan. 2007, p. 10-19.
- [18] T. FURON. *A constructive and unifying framework for zero-bit watermarking*, in "IEEE Trans. Information Forensics and Security", vol. 2, n^o 2, june 2007, p. 149-163.
- [19] C. GUILLEMOT, F. PEREIRA, L. TORRES, T. EBRAHIMI, R. LEONARDI, J. OSTERMANN. *Distributed Monoview and Multiview Video Coding: Basics, Problems and Recent Advances*, in "IEEE signal processing

magazine, Special Issue on Signal Processing for Multiterminal Communication Systems", Sept. 2007, p. 67-76.

- [20] X. JASPAR, C. GUILLEMOT, L. VANDENDORPE. *Joint source-channel turbo techniques: from theory to practice*, in "Proceedings of the IEEE, Special issue on the turbo principle", vol. 95, n^o 6, June 2007, p. 1345-1361.
- [21] M. JEANNE, C. GUILLEMOT, T. GUIONNET, F. PAUCHET. *Error-resilient decoding of context-based adaptive binary arithmetic codes*, in "Signal, Image and Video Processing Journal, Springer", vol. 1, n^o 1, Apr. 2007, p. 77-87.
- [22] H. JÉGOU, C. GUILLEMOT. *Entropy coding with variable length re-writing systems*, in "IEEE Trans. on Communication", vol. 55, n^o 3, Mar. 2007, p. 444-452.
- [23] S. MALINOWSKI, H. JÉGOU, C. GUILLEMOT. *Synchronization recovery and state space reduction for efficient soft decoding of Variable Length Codes*, in "IEEE Trans. on Inform. Theory", vol. 53, n^o 1, Jan. 2007, p. 368-377.
- [24] S. MALINOWSKI, H. JÉGOU, C. GUILLEMOT. *Error recovery and new state models for quasi-arithmetic codes*, in "Eurasip Journal on Applied Signal Processing", to appear, 2007.
- [25] M. MAÎTRE, C. GUILLEMOT, L. MORIN. *3D-model based frame interpolation for distributed video coding of static scenes*, in "IEEE Trans. on Image Processing", vol. 16, n^o 5, May 2007, p. 1246-1257.
- [26] A. ROUMY, D. DECLERCQ. *Characterization and Optimization of LDPC codes for the 2-user Multiple Access Channel*, in "EURASIP Journal on Wireless Communications and Networking", to appear, 2007.
- [27] A. ROUMY, D. GESBERT. *Optimal matching in wireless sensor networks*, in "IEEE Journal on Selected Topics in Signal Processing, special issue on convex optimization methods in signal processing", vol. 1, n^o 4, Dec. 2007.
- [28] 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", to appear, 2007.
- [29] W. YANG, G. RATH, C. GUILLEMOT. *Scalable video coding with novel decorrelation techniques*, in "Eurasip Journal on Applied Signal Processing, special issue on video adaptation for heterogeneous environments", Article ID 54342, 2007.

Publications in Conferences and Workshops

- [30] X. ARTIGAS, S. MALINOWSKI, C. GUILLEMOT, L. TORRES. *Overlapped Quasi-arithmetic Codes for Distributed Video Coding*, in "IEEE Intl. Conf. on Image Processing, ICIP, San Antonio, Texas, USA", Sept. 2007.
- [31] T. BANERJEE, A. KHERANI, M. DEBBAH, A. ROUMY. *Impact of Physical Layer Design on the Deployment of a Sensor Network with Correlated Observations*, in "IEEE International Workshop on Signal Processing Advances for Wireless Communications, Helsinki, Finland", June 2007.

-
- [32] S. CHAABOUNI, N. SELLAMI, A. ROUMY. *Lower bounds on the performance of the MAP equalizer with a priori over MIMO systems*, in "Proc.of ISSPA, International Symposium on Signal Processing and its Applications, Sharjah, UAE", Feb. 2007.
- [33] S. CHAABOUNI, N. SELLAMI, A. ROUMY. *The Impact of a priori information on the MAP equalizer performance with M-PSK modulation*, in "Proc.of 15th European Signal Processing conference, EUSIPCO, Poznan, Poland", Sept. 2007.
- [34] T. COLLEU, G. SOURIMANT, L. MORIN. *Une méthode d'initialisation automatique pour le recalage de données SIG et vidéo*, in "CCompression et REpresentation des Signaux Visuels (CORESA 2007), Montpellier, France", Nov. 2007.
- [35] O. CRAVE, C. GUILLEMOT, B. PESQUET-POPESCU, C. TILLIER. *Robust video transmission based on distributed multiple description coding*, in "15th European Signal Processing Conference, EUSIPCO, Poznan, Poland", Sept. 2007.
- [36] O. DABEER, A. ROUMY, C. GUILLEMOT. *Linear Transceivers for Sending Correlated Sources Over the Gaussian MAC*, in "National Communication Conference", Jan. 2007.
- [37] O. DABEER, A. ROUMY, C. GUILLEMOT. *Optimal Distributed Linear Transceivers for Sending Independently Corrupted Copies of a Colored Source Over the Gaussian MAC*, in "Proc.of ICASSP, IEEE Int. Conference on Acoustics, Speech and Signal Processing, Honolulu, USA", Apr. 2007.
- [38] C. FONTAINE, F. GALAND. *How can Reed-Solomon codes improve steganographic schemes*, in "9th Information Hiding – IH'07", Lecture Notes in Computer Science, Springer-Verlag, June 2007.
- [39] J. FUCHS. *Le filtre adapté global : un algorithme pour le cas complexe*, in "Actes du 20ème colloque GRETSI, Troyes", Sept. 2007.
- [40] J. FUCHS. *Matched detector and estimator with signature uncertainty*, in "Asilomar Conf. on Signals, Systems and Computers, Pacific Grove, CA", Nov. 2007.
- [41] J. FUCHS. *Sparse representations: Applications to adaptive array processing*, in "Adaptive Sensor Array Processing Workshop, Lexington, MA", June 2007.
- [42] J. FUCHS, C. GUILLEMOT. *Fast implementation of a $\ell_\infty - \ell_1$ penalized sparse representations algorithm : Applications in image coding*, in "Asilomar Conf. on Signals, Systems and Computers, Pacific Grove, CA", Nov. 2007.
- [43] D. KUBASOV, K. LAJNEF, C. GUILLEMOT. *Hybrid coder/decoder rate control for Wyner-Ziv video coding with a feedback channel*, in "IEEE Intl. MultiMedia Signal Processing Workshop, MMSP, Chania, Crete", Oct. 2007.
- [44] D. KUBASOV, J. NAYAK, C. GUILLEMOT. *Optimal Inverse Quantisation in Wyner-Ziv Video Coding with Multiple Side Information*, in "IEEE Intl. MultiMedia Signal Processing Workshop, MMSP, Chania, Crete", Oct. 2007.

- [45] G. LE GUELVOUT, T. FURON, F. CAYRE. *The good, the bad, and the ugly: three different approaches to break their watermarking system*, in "Security, Steganography, and Watermarking of Multimedia Contents IX", E. DELP, P. W. WONG (editors), vol. 6505, SPIE, Jan. 2007, 650517.
- [46] Y. LIU, G. RATH, C. GUILLEMOT. *Improved Intra prediction for H.264/AVC scalable extension*, in "IEEE Intl. MultiMedia Signal Processing Workshop, MMSP, Chania, Crete", Oct. 2007.
- [47] A. MARTIN, J. FUCHS, C. GUILLEMOT, D. THOREAU. *Sparse representation for Image Prediction*, in "15th European Signal Processing Conference, EUSIPCO 2007, Poznan, Poland", Sept. 2007.
- [48] A. ROUMY, D. GESBERT. *Optimal matching in wireless sensor networks*, in "IEEE International Symposium on Information Theory", June 2007.
- [49] A. ROUMY, K. LAJNEF, C. GUILLEMOT. *Rate-adaptive turbo-syndrome scheme for Slepian-Wolf Coding*, in "41st Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA", Nov. 2007.
- [50] N. SELLAMI, M. SIALA, A. ROUMY, I. KAMMOUN. *Generalized MAP: sequence detection for non ideal frequency selective channel knowledge*, in "Proc.of ICASSP, IEEE Int. Conference on Acoustics, Speech and Signal Processing, Honolulu, USA", April 2007.
- [51] G. SOURIMANT, L. MORIN, K. BOUATOUCH. *Fusion de données Gps, Sig et Vidéo pour la reconstruction d'environnements urbains*, in "Orasis 2007, onzième congrès francophone des jeunes chercheurs en vision par ordinateur, Obernai, France", June 2007.
- [52] G. SOURIMANT, L. MORIN, K. BOUATOUCH. *Gps, Gis and Video fusion for urban modeling*, in "CGI 2007, 25th Computer Graphics International Conference, Petropolis, RJ, Brazil", May 2007, p. 79-88.
- [53] G. SOURIMANT, L. MORIN, K. BOUATOUCH. *Gps, Gis and Video Registration for Building Reconstruction*, in "ICIP 2007, 14th IEEE International Conference on Image Processing, San Antonio, Texas, USA", The original publication is available at www.springerlink.com, Sept. 2007.
- [54] G. SOURIMANT, L. MORIN, K. BOUATOUCH. *Vers une reconstruction 3D de zones urbaines : mise en correspondance de données Gps, Sig et Vidéo*, in "GRETSI 2007, 21e Colloque en Traitement su Signal et des Images, Troyes, France", Sept. 2007.
- [55] C. TILLIER, O. CRAVE, B. PESQUET-POPESCU, C. GUILLEMOT. *A comparison of four video multiple description coding schemes*, in "15th European Signal Processing Conference, EUSIPCO 2007, Poznan, Poland", Sept. 2007.
- [56] W. YANG, G. RATH, C. GUILLEMOT. *Improved spatial inter-layer prediction for scalable video coding*, in "Proc.of IEEE Int. Conference on Acoustics, Speech and Signal Processing, ICASSP, Honolulu, USA", Apr. 2007.