

*Project-Team TEMICS**Traitement, Modélisation d'Images et
CommunicationS**Rennes*

THEME 3B

The logo consists of the word "Activity" in a white serif font, with a large, light grey, stylized letter "A" to its left. Below this, the word "Report" is written in a white serif font, with a large, light grey, stylized letter "R" to its left. The entire logo is set against a dark blue background.

Activity
Report

2003

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	4
3.4. Watermarking as a problem of communication with side information	5
4. Application Domains	6
4.1. Introduction	6
4.2. Compression with advanced functionalities	7
4.3. Multimedia communication	7
4.4. Copy protection, copyright enforcement and enriched content	7
5. Software	8
5.1. MOVIQS: Video communication platform with QoS support	8
5.2. WAVIX video codec	9
5.3. CHI-MARK2: Robust image watermarking tool	9
5.4. 3D Model-based video codec	9
6. New Results	9
6.1. Analysis and modelling of video sequences	9
6.1.1. 3d Scene modelling from monocular video sequences	9
6.1.2. Video object segmentation and representation for compression purposes	11
6.1.3. Shadow and motion analysis for video compression	12
6.2. Compression and joint source-channel coding	13
6.2.1. Fine grain scalable coding based on 2nd generation wavelets	13
6.2.2. Overcomplete frame expansions for joint source-channel coding	14
6.2.3. Error-resilient source codes	15
6.2.4. Joint source-channel decoding of variable length codes	15
6.3. Image and video watermarking	17
6.3.1. Side informed watermarking and game theory	17
6.3.2. Side-informed watermarking based on frame expansions	18
6.3.3. Interaction between compression, watermarking and indexation	18
7. Contracts and Grants with Industry	19
7.1. Industrial contracts	19
7.2. National contracts	19
7.2.1. RNRT-VIP	19
7.2.2. RNRT-DIPHONET	19
7.2.3. RNRT-DOMUS-VIDEUM	20
7.2.4. RNRT-EIRE	20
7.2.5. RNRT-COSOCATI	21
7.3. European contracts	21
7.3.1. IST-Busman	21
7.3.2. IST-Ozone	22
7.3.3. FP6-IST NoE SIMILAR	22
7.3.4. FP6-IST STREP DANAE	22
7.3.5. FP6-IST STREP ENTHRONE	22
8. Other Grants and Activities	22
8.1. Regional initiatives	22

8.1.1.	Britanny region contract on watermarking	22
8.2.	National initiatives	23
8.2.1.	ARC Télégéo	23
8.2.2.	ACI Fabriano	23
8.2.3.	CNRS specific action on watermarking	23
8.2.4.	CNRS specific action on scalable and robust video compression	24
8.3.	Bi-lateral international co-operation	24
8.3.1.	Project reviewing	24
9.	Dissemination	24
9.1.	Patents	24
9.2.	Standardization	24
9.3.	Leadership within the scientific community	25
9.4.	Invitations	25
9.5.	Teaching	25
10.	Bibliography	26

1. Team

Head of project-team

Christine Guillemot [DR Inria]

Administrative assistant

Stéphanie Lemaile [TR Inria, shared with VISTA]

Inria research scientists

Teddy Furon [CR]

Slavica Marinkovic [CR]

Henri Nicolas [CR]

Stéphane Pateux [CR, on secondment from France Télécom]

Gagan Rath [CR, from Oct. 1st 2003]

Pierre Siohan [DR, on secondment from France Télécom until Aug. 31, 2003]

Faculty members

Luce Morin [Associate professor, Ifsic, University of Rennes 1, on Inria secondment from Feb. 03]

CNRS technical staff

Laurent Guillo [IR]

Inria projects technical staff

Vincent Bottreau [Busman project, from Sept. 1st 2003]

Thomas Guionnet [VIP project, from Apr. 1st, 2003]

Anne Manoury [Domus videum project, until Sept. 31, 2003]

Inria junior technical staff

Brendan Catherine [until Sept. 30, 2003]

Cécile Marc [from Oct. 6, 2003]

Inria post-doctoral fellow

François Cayre [from Dec. 15, 2003]

Claudio Weidmann [until Jan. 31, 2003]

University of Rennes-1 post-doctoral fellow

Gagan Rath [until March 31, 2003]

Ph.D Students

Raphaële Balter [Cifre grant, France Télécom]

Guillaume Boisson [Cifre grant, Thomson Multimedia]

Fabien Catteau [Ministry of research grant]

Nathalie Cammas [Cifre grant, France Télécom]

Vivien Chappelier [Ministry of research grant]

Dubhe Chavira-Martinez [Conacyt grant]

Marc Chaumont [Ministry of research grant, until 30 Sept. 2003]

Robin Chatterjee [Inria grant, VIP project, from 1st Oct. 2003]

Jonathan Delhumeau [Inria grant, BUSMAN project]

Mireya Garcia-Vasquez [Conacyt grant]

Hervé Jégou [Ens-Cachan grant]

Khaled Lajnef [Inria grant, VIP project, from 1st Oct. 2003]

Gaetan Le Guelvouit [Inria- Brittany Region grant, until Sept. 30, 2003]

François Tonnin [Inria- Brittany Region grant, also with Texmex]

2. Overall Objectives

The main goal of the TEMICS project is the design and development of theoretical frameworks as well as al-

gorithms and practical solutions in the areas of analysis, modelling, coding, communication and watermarking of images and video signals. TEMICS 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, or navigation requires the development of video analysis and modelling algorithms. TEMICS focuses on the design of solutions for segmenting video objects and for extracting and coding their main attributes (shape, motion, illumination, ...). In order to support navigation within video scenes, the ability to construct a 3D model of the scene is a key issue. One specific problem addressed is the design of algorithms for 3D modelling from monocular video sequences with optimum tradeoff between model reliability and description cost (rate). Finally, the optimal support of the above functionalities in networked multimedia applications requires scalable, compact and transmission noise resilient representations of the models and of their attributes, making use of joint source-channel coding principles (see below).
- *Joint source-channel coding.* The advent of Internet and wireless communications, often characterised 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, stating that source and channel optimum performance bounds can be approached as close as desired by designing independently source and channel coding strategies, 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 in order to address the tradeoff performance-complexity.
- *Compression, scalable coding and distributed source coding.* Scalable video compression is essential to allow for optimal adaptation of compressed video streams to varying network characteristics (e.g. to bandwidth variations) in various applications (e.g. in unicast streaming applications with pre-encoded streams, and in multicast applications). 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. 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, light-weight video compression systems (extension of MMS-based still image transmission to video) 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 theorem established in 1973. Even though first theoretical foundations date back to early 70's, it is only recently that concrete solutions, motivated by the above applications, aiming at approaching the theoretic performance bounds have been introduced.

- *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. However, data hiding technology can also be used for enhancing a signal by embedding some meta-data.

The data hiding problem can be formalized as a communication problem : the aim of robust data hiding is indeed to embed the maximum 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. Our developments rely on this formalism, i.e., on scientific foundations in the areas of communication theory, such as channel coding with side information and joint source-channel coding concepts and algorithms.

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 (IETF, 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 (RNRT COSOCATI, DIPHONET, EIRE, VIP, RNTL DOMUS-VIDEUM) and European projects (IST-BUSMAN and IST-OZONE). 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

Key words: *3D reconstruction, computer vision, projective geometry, perspective projection, camera models, fundamental matrix, epipolar constraints.*

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 a 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

Key words: *Wavelet transforms, filter banks, oversampled frame expansions, error correcting codes, multiple description coding.*

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

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

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

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

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. Watermarking as a problem of communication with side information

Key words: *watermarking, side information, information theory, capacity, discrimination.*

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

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

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 :

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

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.

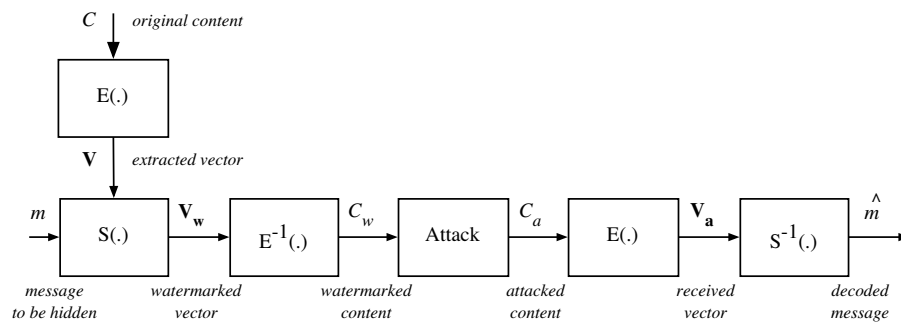


Figure 1. Classical watermarking scheme

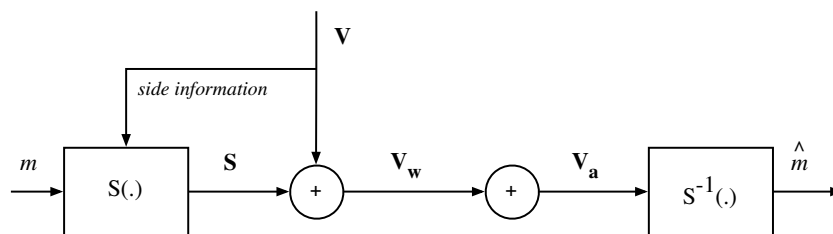


Figure 2. Watermarking as a problem of communication with side information.

4. Application Domains

4.1. Introduction

TEMICS addresses three main application domains : compression, including with advanced functionalities of scalability and content-based interaction, networked multimedia applications (on wired or wireless Internet)

with various requirements and needs in terms of scalable and compact representation, of resilience to channel noise, and content protection and enhancement.

4.2. Compression with advanced functionalities

The field of video compression has known, during the last decade, a significant evolution leading to the emergence of a large number of international standards (MPEG-4, H.264). Notwithstanding this 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, and is a natural application framework of many TEMICS developments. 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. In particular, fine grain scalable (FGS) coding solutions making use of mesh-representations and/or spatio-temporal frame expansions are developed in order to allow for rate adaptation to varying network bandwidth in streaming scenarios with pre-encoded streams.

Even though, for most multimedia applications, compression remains a key issue, this is not the only one that has to be taken into account. Emerging applications in the area of interactive audiovisual services show a growing interest for interactivity, content-based capabilities, (e.g. for 3-D scene navigation, for creating intermediate camera viewpoints) for integration of information of different nature, e.g. in augmented and virtual reality applications. These capabilities are not well supported by existing solutions. Interaction and navigation with the video content requires extracting appropriate models, such as regions, objects, 3-D models, mosaics, shots... These features are expected to be beneficial to multimedia applications requiring 3-D virtual scenes, such as video games or virtual visits of museums, virtual and augmented reality.

4.3. Multimedia communication

The emergence of networks such as 2.5G, 3G networks and ADSL but also of new terminal devices, e.g. handhelds, advanced mobile phones should create a propitious, yet challenging, ground for the development of advanced multimedia services. Networked multimedia is indeed expected to play a key role in the development of 3G and beyond 3G (i.e. all IP-based) networks, by leveraging higher available bandwidth, all-IP based ubiquitous and seamless service provisioning across heterogeneous infrastructures, and new capabilities of rich-featured terminal devices.

However, all-IP based ubiquitous and seamless multimedia service provisioning across heterogeneous infrastructures, presenting a number of challenges beyond existing networking and source coding capabilities, is only a vision so far. In particular, networked multimedia will have to face problems of transmission of large quantities of information with delay constraints on heterogeneous, time-varying communication environments with a non-guaranteed quality of service (QoS). End-to-end QoS provisioning, all the most challenging in a global mobility context is of the utmost importance for consumer acceptance and adoption. 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. These trends are exemplified within 3GPP and the IETF by the UDP-lite and ROHC initiatives. In that context, the joint source-channel coding paradigm sets the foundations for the design of efficient practical solutions to the above application challenges, that we address via different industrial (with Thomson Multimedia, France Telecom), national (RNRT-VIP, RNRT-COSOCATI) and European (IST-OZONE) partnerships.

4.4. Copy protection, copyright enforcement and enriched content

The problem of data hiding has gained considerable attention in the recent years as a potential solution for a wide range of applications encompassing copy protection, copyright enforcement, content enhancement by meta-data embedding, authentication, and steganography. TEMICS focuses, via its collaborations and contracts, on the three first applications.

Copy protection: The history of copy protection dates back from the analogue age. Yet, in the digital age, this issue is even more crucial. The biggest effort to build a digital right management system is the attempt of the copy protection technical meeting group for the DVD video format. The goal of copy protection systems is not to forbid copying but rather to enforce some usage rights (e.g. view now, view only for X days, copy once, copy locally).

Usually, conditional access to content as well as users rights management are offered via cryptographic functions. But, a dishonest user might record content in a decrypted form (at least from the analogue signals). The watermark is then just a flag warning the devices that pirated clear content is copyrighted and that it was protected. Basically, the watermark is used to distinguish copy free content from clear pirated contents. Therefore, the mark should be non perceptible and very robust to attacks. In this case, the capacity need not be large. The main issue is the security of the watermark primitive. TEMICS will address this application domain in the ACI FABRIANO.

Copyright enforcement: The availability of multimedia contents in digital forms has brought a number of security issues to the forefront. The "digital revolution" has made digital data very vulnerable to unauthorized use. Watermarking primitives offer technical solutions to these security problems by providing means to trace copies along the distribution chain (from the artist to the consumers), to spot illegal uses of copyrighted contents and to ultimately prove the ownership in case of copyright struggles. For this type of application, the watermark capacity need not be large, but the watermark must be non perceptible and very robust to attacks. The RNRT Diphonet project addresses this application of watermarking. The concept of security being in this context of utmost importance, as there may be usurpers hacking the copyright protection system, it is necessary to define a methodology for analyzing the security level of the watermarking system.

Content enhancement: Watermarking provides a way to embed meta-data into the multimedia content for enhanced services. The content becomes self-contained, the created meta-data transmission channel travelling with the content itself. With respect to traditional solutions where the data is transported beside the content, e.g. into a label (field, head of file, tags), data hiding based systems allow for seamless meta-data transport. When placed in separate channels, the data can be unintentionally removed when submitted to transformations such as D/A+A/D transformation, transcoding within heterogeneous networks. The data-hiding based solution should prevent the metadata from being lost. The embedded data is inside the content and no special steps need to be taken in storage media or transmission networks to keep the metadata and content together. The embedded data must be non perceptible, and possibly robust to some content processing (e.g. compression). This application requires high embedding capacity and possibly fast embedding and real-time decoding solutions. The IST-BUSMAN project addresses this application.

5. Software

5.1. MOVIQS: Video communication platform with QoS support

Participants: Laurent Guillo [correspondant], Cécile Marc.

With the support of several contracts (RNRT-VISI, IST-SONG), TEMICS has developed a video communication platform. This latter provides a test bed allowing to study and assess, in a realistic way, algorithms implementing joint source channel coding, video modelling or video coding. The software MOVIQS (*module pour de la vidéo sur Internet avec qualité de service*) is one of the platform component. It is a dynamic link library used by a video streaming server and the related clients. They can take advantage of three of its main mechanisms: video transport in both unicast and multicast mode, congestion control and rate regulation, and loss control. The release 1.0 of the software MOVIQS has been registered at the Agency for the Protection of Programmes (APP) under the number IDDN.FR.001.030031.000.S.P.2003.000.10200. The platform is still being developed. Its next main evolutions will rely on the integration of the fine grain scalable video coder WAVIX and of developments based on IPv6 and the ROHC framework resulting from the collaboration with the ARMOR project-team.

5.2. WAVIX video codec

Participants: Brendan Catherine, Christine Guillemot, Laurent Guillo [correspondant].

The software WAVIX (Wavelet based Video Coder with Scalability) is a low rate fine grain scalable video codec based on a motion compensated 2D+t wavelet analysis. In order to code the spatio-temporal sub-bands, the first release used the EBCOT algorithm provided by the Verification Model JPEG2000. That release 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 and then used by Thomson as part of a partnership. A new version, making use of the Jasper library and of error resilient tools, is still under development.

5.3. CHI-MARK2: Robust image watermarking tool

Participants: Gaëtan Le Guelvouit, Stéphane Pateux [correspondant].

This software implements several data-hiding techniques (embedding and extraction) for images and video. The algorithms implemented are based both on state of the art techniques (embedding and extraction based on wide spread spectrum) and on TEMICS theoretical results (a strategy based on the game theory and taking into account optimal attack and potential de-synchronizations, side-informed technique based on dirty paper codes). 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.

5.4. 3D Model-based video codec

Participants: Raphaële Balter, Luce Morin [correspondante].

From a video sequence of a static scene viewed by a monocular moving camera, this software allows 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. 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. This codec allows to compress at low and very low bit-rates (16 to 256 kb/s in 25Hz CIF format) with a satisfactory visual quality.

6. New Results

6.1. Analysis and modelling of video sequences

Key words: *motion, dense fields, disparity, segmentation, tracking, triangulation, 2D and 3D meshes, active meshes, hierarchical meshes, active contours, depth maps, projective geometry, mosaicking, augmented reality, illumination models.*

6.1.1. 3d Scene modelling from monocular video sequences

Participants: Raphaële Balter, Luce Morin.

A video representation scheme based on a set of 3D models has been developed. The approach originality resides in the construction of a set of independent 3D models linked by common view points (key images), instead of a unique 3D model as in classical approaches. The sequence of 3D models can be streamed for remote navigation in the scene. Several aspects have been optimized and enhanced this year.

The approach assumes that the camera undergoes non degenerated motion, i.e., the camera motion enables 3D reconstruction. Hence, the motion resulting from camera panning is assumed to be close to a pure rotation around the vertical axis, and the camera optical center is assumed to be static, leading to a unique viewpoint for all images. Therefore, in this case, 3D information cannot be retrieved. However, the video sequence can be efficiently described with 2D models, e.g., with mosaics. This observation led to the design of a hybrid modelling approach in which 3D models are used in presence of non degenerated motion, and in which 2D

mosaics are used for video segments where 3D reconstruction is not possible. The mosaic is obtained from a deformable mesh. The model type (2D or 3D) is chosen according to the magnitude of camera rotation and translation with respect to the apparent motion magnitude. A 3D textured cylinder is generated from the 2D panoramic image and is associated with camera positions. This allows a compatible 3D visualization scheme for both 2D and 3D models (see Fig. 3). 3D morphing has been added in order to cope with problems of discontinuities between the models in the sequence. The successive 3D models are mapped on a common parametric space ; the 2D parameters are merged into a 2D mesh containing the respective vertices and edges. Morphing is then achieved through vertices position interpolation, avoiding the need for re-meshing.



a



b



c

Figure 3. Hybrid 2D/3D modelling : (a) images from the original sequence. (b) 2D mosaic. (c) textured cylinder.

A technique to encode the 3D models has been designed. The 3D models are independent and produced by elevation from a uniform triangular mesh on each key image. The coding algorithm is based on a wavelet decomposition of the model geometry, associated with a unique topological model. This leads to a scalable

representation of the model geometry, preventing visual artifacts inherent to topological re-meshing. A patent has been filed (see subsection 9.1).

6.1.2. Video object segmentation and representation for compression purposes

Participants: Marc Chaumont, Dubhe Chavira-Martinez, Nathalie Cammas, Henri Nicolas, Stéphane Pateux. Object-based video coding approaches are often proposed for compression with advanced functionalities. Object-based video representation and coding allow for semantic interpretation and associated manipulation. Compression efficiency can also be improved by handling the occlusions, by selecting adapted coding techniques for each object, and by optimal allocation of bit rates to the different objects.

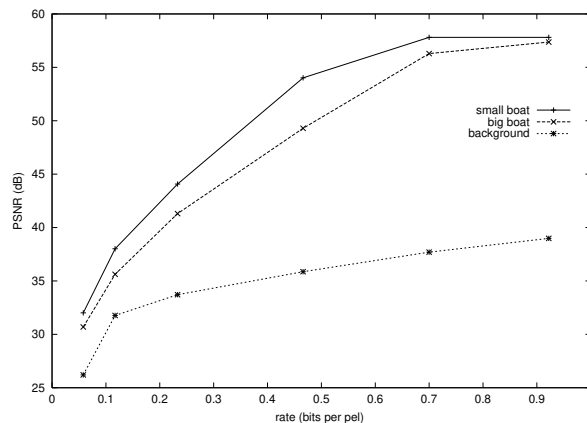
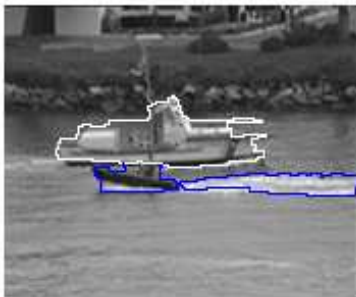


Figure 4. Tracking of video objects, and rate-distortion curves for each object using progressive coding.

Two object-based video coding algorithms making use of TEMICS segmentation tools with some temporal tracking refinements have been developed. The first algorithm relies on a predictive texture-based coding approach. The segmentation extracts a set of objects together with their mean texture. Video is then reconstructed

with a two-layer representation. In the first layer, the mean texture information of each object is robustly transmitted. In the second layer, segmentation, motion and texture refinement information is transmitted separately for each frame. Segmentation and motion information allow to warp the texture in order to obtain a coarse approximation of the image. This decomposition allows a progressive and robust transmission: any frame may be lost or dropped, refinement information can also be coded progressively (e.g., by bit-plane coding techniques). Early experiments show promising results at low bit-rate.

The second algorithm is based on an analysis-synthesis approach, allowing to de-correlate shape, motion and texture information, coupled with spatio-temporal wavelet decompositions. Notice that, in classical object-based coders, shape, motion and texture information are usually correlated, resulting with some limitations with respect to scalable representations. Motion is first estimated thanks to active meshes tracked over several frames. Using this information, texture and shape may be extracted and represented independently of the motion information. We then end up with three types of information to encode and transmit. Each information is then decomposed using spatio-temporal wavelet transforms and progressively encoded. The resulting bit-streams are fully scalable, i.e., spatially, temporally, in terms of SNR, and allow object-based scalabilities. The compression efficiency is further optimized by dynamically selecting different coding techniques for the different objects : a rate-distortion optimization procedure selects a coding technique among 3D model-based coding, mosaic-based representation, mesh-based or bloc-based coding.

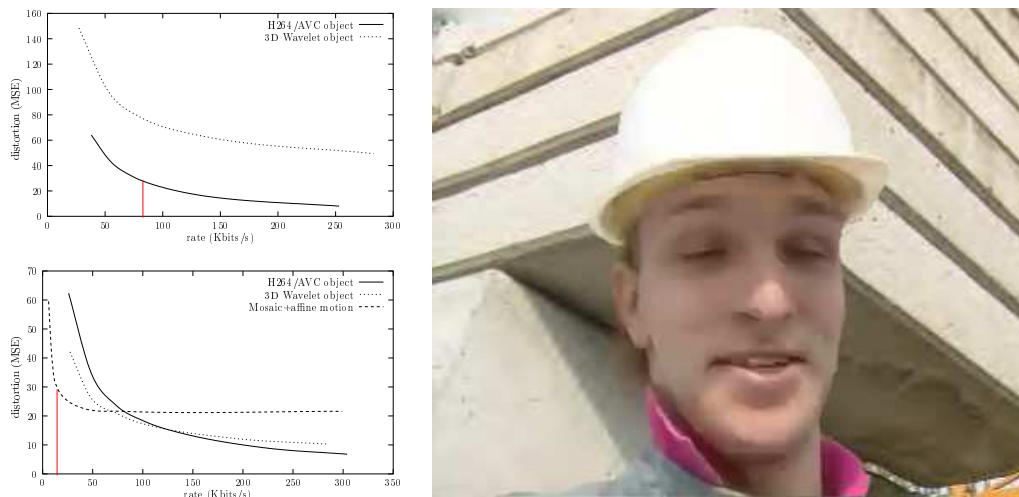


Figure 5. Optimal selection of the coding technique for object-based coding scheme. Targeted bit-rate : 100 kbit/s. 80 kbit/s is allocated to the foreground object (face) using H264 codec, 13 kbit/s is allocated to the background object using sprite model, and 6 kbit/s for shape information. Other targeted bit-rates would lead to different allocation and coding strategies.

6.1.3. Shadow and motion analysis for video compression

Participants: Fabien Catteau, Mireya Garcia-Vasquez, Henri Nicolas.

Motion compensation is a core technique for video compression. It allows to efficiently exploit the temporal redundancy existing between successive images. Moving shadows in a sequence create temporal activity which reduces the motion-compensated temporal prediction efficiency. We have thus focused on the optimization of motion analysis by taking into account the presence of shadows and augmented motion models beyond the classical translational model.

A global and local illumination analysis tool allowing to extract moving shadows has been developed in order to improve the temporal prediction. The representation of the shadow takes into account penumbra and

ambient light variations. The approach is based on a set of level lines defined in the luminance ratio space. Breaking nodes are detected on the level lines represented with B-spline functions. This provides smooth texture variations in the reconstructed shadows, allowing a relatively precise shadow prediction. Two data streams corresponding respectively to the shadow information (contour and texture) and to the sequence without the shadows are then coded separately.

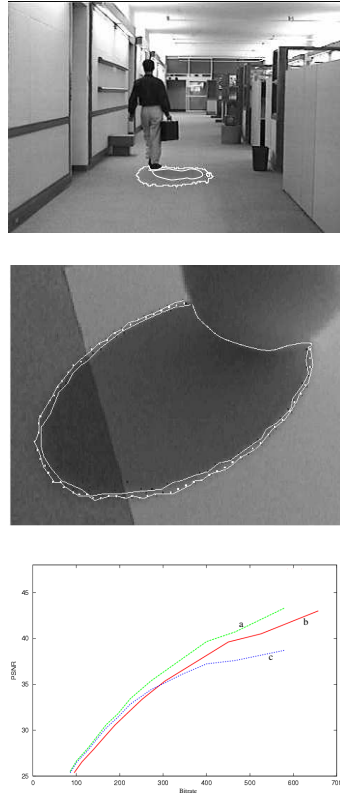


Figure 6. Left and middle: examples of shadow and penumbra segmentation (Hall and Desk sequences). Right: PSNR vs bitrate, (a) sequence after shadow removal, (b) original sequence, (c) sequence with separate coding of the shadow information (the prediction shadow error is not coded).

Block-based motion estimation approaches are often used with a translational model. Affine models have been developed in order to take into account zoom and 2-D rotations. However, non-planar rotations (i.e., rotations around an axis parallel to the image plane) are not taken into account. Here, non-planar rotations are estimated in order to reduce the prediction error. This model is defined by two parameters : the rotation angle and the angle defining the rotation axis. A prediction gain can be achieved on rotating blocks or on regions for which the translational motion model fails.

6.2. Compression and joint source-channel coding

Key words: *compression, error and erasure resilient coding and decoding, congestion control, multiple description coding, scalability, wavelet transforms, overcomplete frame expansions, rate-distortion theory, information theory, stochastic modelling, Bayesian estimation, probabilistic inference, turbo principle, Internet, wireless communication, quality of service.*

6.2.1. Fine grain scalable coding based on 2nd generation wavelets

Participants: Vivien Chappelier, Christine Guillemot, Slavica Marinkovic.

Wavelets are a well-known mathematical tool for representing 1-D signals with a finite number of discontinuities with a small number of coefficients. However, if we model images as homogeneous regions delimited by contours, discontinuities become curves instead of points and the separable wavelets become sub-optimal: high energy wavelets coefficients cluster around edges and all the bitrate is spent on edges. Thus, new approaches have been designed in order to take into account the geometry of the images.

Curvelets and contourlets focus on designing filter banks with directional selectivity in the high frequencies, so that coefficients represent oriented portions of edges instead of points. Their main advantage is that they avoid having a geometric model of the image. The counterpart is that discrete implementations of curvelet transforms are currently redundant, which limits their compression performance. The efficiency of contourlets for compression have been benchmarked, leading to the design of new directional critically-sampled filter banks based on quincunx wavelets.

The bandlets follow a different approach as they use a geometric model to describe the discontinuities of the image (parametrized curves or regularity flow) and wrap wavelets along these discontinuities. Though theoretically more efficient than curvelets, the main problem is to optimise the bitrate allocation to the geometric description and to the wavelet coefficients. We are currently addressing these issues by designing a second generation wavelet scheme based on wavelet lifting locally adapted to the multiresolution image geometry description.

6.2.2. *Overcomplete frame expansions for joint source-channel coding*

Participants: Christine Guillemot, Slavica Marinkovic, Gagan Rath.

Overcomplete frame expansions have been introduced recently as signal representations that would be resilient to erasures in communication networks. Unlike the traditional signal representations with orthogonal bases, here a signal is represented by an overcomplete set of vectors that has some desirable reconstruction properties. The redundancy inherent in the representation is exploited to protect the signal against unwanted channel degradations. Therefore the frame expansions can be looked upon as joint source-channel codes. Redundant block transforms such as those obtained from DFT, DCT, and DST matrices can be seen as producing discrete frame expansions in finite dimensional real or complex vector spaces whereas oversampled filter banks can be seen as providing frame expansions in $l_2(Z)$.

With discrete frame expansions, the associated redundant block transforms or, equivalently the frame operators, can be interpreted as the generator matrices of some real or complex block codes. Therefore such frame expansions can be characterized based on the properties of the parity check matrices of the associated codes, such as the BCH structure [22]. We observed that the frame expansions associated with low-pass DFT, DCT and DST codes possess this structure and thus can be utilized to correct coefficient errors and erasures [54]. Since the frame expansion coefficients are quantized and encoded before being transmitted over a digital network, the error and erasure correction efficiencies are affected by the quantization noise. Among the aforementioned frame expansions, the one associated with DFT codes was observed to be the least sensitive to quantization noise for a Gauss-Markov source. However, for image transmission with fixed-length coding they were observed to have similar signal-to-noise ratios.

The traditional BCH decoding or syndrome decoding approach is based on the concept of an error locator polynomial to localize the errors. However, the analogy between the DOA estimation in array signal processing and the error localization with quantized discrete frame expansions gives rise to newer decoding schemes. We have derived some subspace based approaches to error localization that are applicable to the discrete frame expansions characterized by the BCH structure [24]. The algorithms follow the eigendecomposition of syndrome covariance matrices, estimate the eigenvectors which span the error and the noise subspaces, and then estimate the error locations from the noise subspace eigenvectors. We observed that these decoding approaches improve the error localization efficiency over the syndrome decoding depending on the number of coefficient errors.

Oversampled filter bank (OFB) frame expansions can be viewed as a generalization of the overcomplete signal expansions by redundant block transforms. That is, block transforms can be seen as filter banks with a zero order polyphase description. Increasing the polyphase filter order adds memory to the code and an

OFB can be interpreted as a convolutional code over the real or complex field. We have studied OFB codes in light of the filter bank, frame and coding theory and examined erasure resilience of three OFB structures. The first OFB code is based on an iterated two-channel filter bank structure which presents the advantage of naturally supporting unequal loss protection and of being backward compatible to a critically-sampled wavelet transform. The power of erasure recovery of such transforms has been studied both theoretically (finding conditions for reconstruction) and experimentally [17]. The two other codes considered are based on the cosine modulated filter banks: an oversampled cosine modulated filter bank (CMFB) code and an OFB code similarly structured as a DFT code. The perfect reconstruction (PR) for the CMFB based OFBs with erasures is proven for the case of erasure patterns for which PR depends only on the general structure of the code and not on the prototype filters. For some of these erasure patterns the expression for the mean square reconstruction error is also independent of filter coefficients and can be expressed in terms of the number of erasures, and of parameters such as number of channels and oversampling ratio.

6.2.3. *Error-resilient source codes*

Participants: Christine Guillemot, Hervé Jégou.

Entropy coding, producing Variable Length Codes (VLCs), is a core component of any multimedia compression system (image, video, audio). The main drawback of VLCs is their high sensitivity to channel noise: bit errors may lead to dramatic decoder desynchronization problems. Most of the solutions known so far to address this problem consists in re-augmenting the redundancy of the compressed stream, e.g. using redundant source codes, synchronization markers, or channel codes. In 2002, we have designed a new family of codes, that we called *multiplexed codes*, which have the property of avoiding the dramatic desynchronization problem while still allowing to reach the entropy of the source. The idea underlying *multiplexed codes* builds upon the observation that most media compression systems generate sources of different priority. The design principle consists in creating fixed length codes (FLCs) for high priority information and in using the inherent redundancy to describe low priority data, hence the name “multiplexed codes”. The redundant set of FLCs is partitioned into equivalence classes according to high priority source statistics. The cardinality of each class, according to the high priority source statistics, is a key element so that the code leads to a description length as close as possible to the source entropy.

The first family of code has been designed taking into account the source stationary distributions only. In 2003, we have extended the *multiplexed codes* in order to take into account higher-order source statistics, focusing primarily on first-order statistics. The key aspect to limit error propagation inherent to the use of conditional probabilities relies on the choice of some appropriate partitions (or index assignment) of the set of codewords. Several index assignment methods have been developed. A *crossed-Index Assignment* method leads to the construction of a *kernel* made of codewords offering synchronization properties. The decoder re-synchronization capability can be further increased by periodic and adaptive use of stationary multiplexed codes. Soft decoding algorithms using MAP, MPM and MMSE criteria and exploiting both the channel characteristics and the residual redundancy (if any), have also been developed.

A second family of codes, called self-multiplexed codes, has been designed for error resilient and progressive transmission of variable length coded sources. These codes do not assume the existence of two or more sources. The design of these codes is based on the observation that some transitions in a codetree are more resilient, i.e., not subject to desynchronization. In order to optimize the SNR decoding performance of variable length codetrees, one has then to make sure that most of the source energy is localized on these synchronized transitions in the codetree. A code design procedure, based on the use of binary multiplexed codebooks applied to a single source multiplexed with itself has been developed. The resulting codes have the same compression efficiency as Huffman codes, the same computing cost, but lead to significantly higher SNR performances when they are decoded in presence of bit errors resulting from transmission noise.

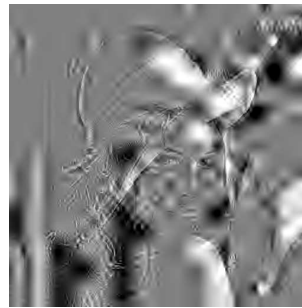
6.2.4. *Joint source-channel decoding of variable length codes*

Participants: Christine Guillemot, Thomas Guionnet, Pierre Siohan.

Arithmetic codes are now widely introduced in practical compressions systems (e.g. JPEG-2000, MPEG-4 and H.264 standards). When coupled with higher-order source statistical models, they allow to reach high compression factors. However, they are very sensitive to the presence of noise (transmission errors). It is then essential to design algorithms that would allow robust decoding of such codes, even in presence of bit errors. Two algorithms have been designed: the first algorithm follows sequential decoding principles, in the spirit of the Fano algorithm used for decoding channel codes, with an extra difficulty here residing in the fact that transitions on the coding decision tree are associated to a varying number of bits [16]. In order to control the decoder complexity, different pruning techniques have been designed. A validation of the approach in a JPEG-2000 decoder revealed a very significant gain with respect to standard decoding solutions (see Fig. 7). The algorithm allows in addition to flexibly control the trade-off between the complexity represented by the parameter W in Fig. 7 and the reliability of the estimation. This parameter corresponds to the maximum number of surviving paths : a higher value of W corresponds to a higher decoding complexity and a higher estimation reliability. A contribution to the standardization group has been made for a possible insertion in the standard document as informative annex.



(a) JPEG-2000 coded; PSNR = 37.41 dB; no channel error.



(b) JPEG-2000 coded; AWGN channel ($E_b/N_0 = 5$ dB), PSNR = 16.43 dB.



(c) JPEG-2000 coded, AWGN channel ($E_b/N_0 = 5$ dB), $W = 10$, PSNR = 25.15 dB.



(d) JPEG-2000 with sequential decoding, AWGN channel ($E_b/N_0 = 5$ dB), $W = 20$, PSNR = 31.91 dB.

Figure 7. Impact of transmission errors on the visual quality (0.5 bpp).

The sequential decoding approach remains however suboptimal due to the pruning needed to maintain the complexity within a tractable range. We have then considered quasi-arithmetic codes, which can be seen as a reduced precision implementation of arithmetic codes. A quasi-arithmetic coder/decoder can be modelled as finite-state automata. MAP estimation algorithms (e.g. the BCJR or the soft output Viterbi algorithm)

hence apply [15]. The dimension of the state-space can be traded against some approximation of the source distribution. The approach turns out to be well suited for extra soft-synchronization and to be used in a source-channel iterative structure in the spirit of serial turbo codes.

However, these estimation algorithms require *a priori* information on the statistics of the source assumed to be a first-order Markov chain. A method for estimating the source HMM parameters from the noisy sequence of bits that is received has been designed. The approach is inspired from the Baum-Welch algorithm with the computations made along the two recursions (forward and backward) of the BCJR algorithm used for the source decoding. The statistics estimation algorithm has also been incorporated in a joint source-channel decoding structure.

6.3. Image and video watermarking

Key words: *data hiding, signature, copyright, spread spectrum, channel coding, game theory.*

6.3.1. Side informed watermarking and game theory

Participants: Gaetan Le Guelvouit, Stéphane Pateux.

In 2002, we have developed a watermarking technique making use of both communication tools (wide spread spectrum modulation, modulation, error correcting codes), as well as game theory. A patent has been filed and the corresponding software has been registered at the APP. In 2003, we have derived models of attacks for different types of degradations, e.g., scaling, additive white Gaussian noise and de-synchronizations, and introduced the notion of informed attacks. The models lead to closed-form expressions for the different parameters of the embedding and extracting techniques to be used in a practical watermarking system, as well as to the performance bounds of the system. Error correcting codes, based on punctured convolutional codes, have also been introduced in the approach leading to a side-informed watermarking approach with good performance in terms of capacity as shown in Fig. 8. The approach has been validated with large professional data bases.

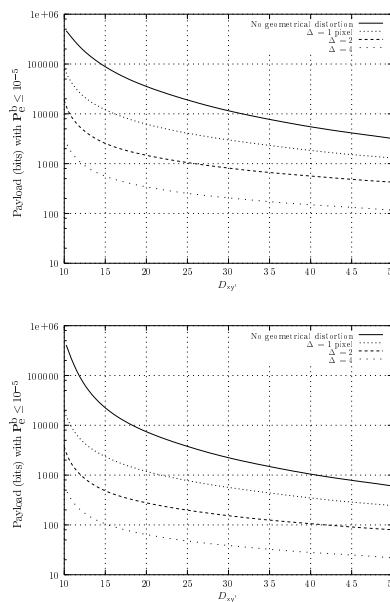


Figure 8. Impact of desynchronisation on the capacity of data-hiding for Lenna image. Left curves show theoretical capacity. Right curves show obtainable capacities considering a probability of error $P_e \leq 10^{-5}$.

6.3.2. Side-informed watermarking based on frame expansions

Participants: Jonathan Delhumeau, Teddy Furon, Christine Guillemot.

Costa¹ has shown that, for AWGN channels, the channel capacity of a side-informed emitter (i.e. in the case of watermarking, the embedding capacity) does not depend on the channel state power (i.e., on the host signal). In other words, the cover signal interference, inherent to spread spectrum techniques, is not present any more. However, the solution proposed, known as the *Ideal Costa Scheme* (ICS), and based on codebooks partitions, requires very large codebooks, hence is not realistic for practical applications. Suboptimal, but practical, approaches of watermarking with side information based on tractable codebooks, such as the SCS approach, have been proposed in the literature. They often rely on error correcting codes (e.g., block codes, convolutional codes, punctured convolutional codes, turbo codes) defined in finite fields.

The design of a side-informed watermarking system based on frame expansions that can also be seen as codes on the real or complex fields is under progress. These codes are actually used for constructing parallel Gaussian channels in the spirit of MIMO systems, channels that will then be used to convey the watermark message. The channels are being defined by quantizing complex-valued syndrome coefficients of the codes, with some distortion constraints in order to control the non perceptibility of the construction of the watermark channels within the content. These distortion constraints can be regarded as channel power constraints. Preliminary experimental results shown in Fig.9 turn out to be promising.

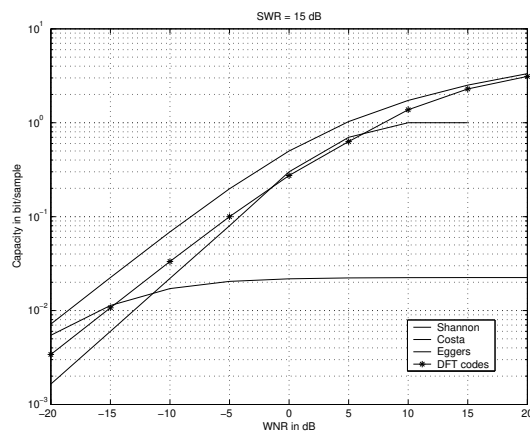


Figure 9. Capacity in bits hidden per content sample as a function of the watermark to noise ratio (ranging from -20 dB to 20 dB). The signal to watermark power ratio is set to 15 dB. The figure compares the approach with a)- the Shannon capacity of an AWGN channel with two sources of noise; b)- Costa's side-informed theoretical capacity; c)- the capacity achieved with the SCS system.

6.3.3. Interaction between compression, watermarking and indexation

Participants: Stéphane Pateux, François Tonnin.

In collaboration with the TexMex project team, we are studying the interactions between, and the mutual impact of, indexing, compression and watermarking. In particular we are investigating the impact of compression and watermarking on the descriptors extraction. We have also developed a multi-resolution salient point detector allowing to extract feature points in a wavelet image representation domain. The detector is hence inherently robust to wavelet based compression and provides extra information on the scale spread of the given feature points.

¹M.H.M. Costa, "Writing on dirty paper", IEEE Trans. on Information Theory, vol.29, No.3, May 1983.

7. Contracts and Grants with Industry

7.1. Industrial contracts

TEMICS has three Cifre contracts with industrial partners:

- Cifre contract with France Telecom RD in the context of the Ph.D of Nathalie Cammas in the area of video coding using active meshes and 3D wavelets. Active meshes are used in order to model the deformation of objects in a scene. The results achieved in 2003 are described in sub-section 6.1.2.
- Cifre contract with France Telecom RD in the context of the Ph.D of Raphaele Balter in the area of 3D-model based coding of video sequences. The results achieved in 2003 are described in sub-section 6.1.1.
- Cifre contract with Thomson Multimedia R&D in the context of the Ph.D of Guillaume Boisson in the area of scalable video coding based on motion-compensated spatio-temporal wavelets. The focus in 2003 has been on adaptive motion-compensated temporal filtering and on efficient and scalable coding of corresponding motion fields.

7.2. National contracts

7.2.1. RNRT-VIP

Participants: Christine Guillemot, Thomas Guionnet.

- Convention number : 2 01 A 0650 00 000MC 01 1
- Title : Video over wireless IP
- Research axis : § 6.2.2 and § 6.2.4.
- Partners : Comsys, ENSEA, France Télécom R&D, Irisa/Inria-Rennes, INRIA-Sophia, ENST-Paris, Université Paris-6, Thalès Communication.
- Funding : Ministry of industry.
- Period : Nov.01- Apr.04.

The project objective is to design error resilient video coding solutions and joint source-channel coding techniques for robust transmission of video signals over wireless IP networks. TEMICS contributes to VIP by designing estimation algorithms for robust decoding of arithmetic codes in presence of channel noise [16], [15], and by integrating these techniques in a fine grain scalable video coder and decoder. TEMICS also studies redundant wavelet transforms for erasure-resilient coding and decoding of video signals [17].

7.2.2. RNRT-DIPHONET

Participants: Gaëtan Le Guelvouit, Stéphane Pateux.

- Convention number: 2 01 A 0704 00 000MC 01 1
- Title: Picture broadcasting over the Internet
- Research area associated: § 6.3.1.
- Partners: Canon, CNRS (L2S), INRIA, Andia Press
- Funding: Ministry of industry.
- Period: Jan.02-Jun.04.

The aim of the Diphonet project is to develop protection and tracing tools for applications of professional images delivery over the Internet. The watermarking technique is used to insert copyright information as well as meta-data in order to trace the origin of a picture. The watermarking technique based on game theory reported in sub-section 6.3.1 has been evaluated and optimized against intentional attacks and potential desynchronizations.

7.2.3. *RNTL DOMUS-VIDEUM*

Participants: Anne Manoury, Henri Nicolas.

- Convention number: 2 02 C 0100 00 00 MPR 01 1
- Title: *Indexation, advanced visualisation and video content-based access.*
- Research axis: § 6.1.3
- Partners: Inria (METISS, TEMICS, VISTA project teams), Thomson Multimedia, Ecole polytechnique de Nantes, INA, SFRS.
- Funding: Ministry of industry.
- Period: Dec.01-July.04.

The aim of the project is to develop tools for indexing, content-based access and for advanced visualization of videos. We designed a technique for structuring a video sequence in a set of hyper-scenes, where each hyper-scene gathers similar scenes. This method is based on an initial scene shot decomposition assumed to be available. The criteria used to merge the initial shots are based on the use of 1-D mosaic representations (each initial shot is represented using two 1-D mosaic images). The similarity between two scenes is therefore evaluated by comparing their mosaic images. Such comparisons are done here using global statistical similarities, and a region-based matching criterion. These different criteria are embedded in a decision process. The tool has been applied to MPEG-2 compressed content. The 1-D mosaic images are thus computed using the MPEG-2 motion vectors. Then they are approximated by a polygonal representation in order to simplify the comparison process. The video structure is therefore obtained using a clustering algorithm. Experiments show that satisfactory hyper-scene structuration can be obtained.

7.2.4. *RNRT-EIRE*

Participants: Christine Guillemot, Henri Nicolas.

- Convention number: 2 02 C 0099 00 31324 01 1
- Title: *Optimization of image compression algorithms based on JPEG-2000.*
- Research axis: § 6.1.3 and § 6.2.1
- Partners: Inria, Thalès, I3S, CRIL Technology, ENSTA, IRCOM.
- Funding: Ministry of industry.
- Period: Nov.01-Feb.04.

The project objectives are to develop image and video compression algorithms with optimized rate-allocation algorithms and supporting fine grain scalability. TEMICS focuses on the scalable video compression aspects. The algorithm setting the basis for our contributions is based on a 3-D spatio-temporal decomposition of each Group Of Frames (GOF). A motion estimation based on a quadtree decomposition is incorporated in the temporal filtering in order to obtain a more efficient temporal de-correlation. The spatial decomposition is similar to the technique used in JPEG-2000. The obtained quantized coefficients are therefore coded using a 3-D EBCOT coding method. Furthermore, two GOF coding modes (Intra and Inter) are introduced to take into account the correlation between two successive GOF. The GOF length is chosen in function of the temporal variations between the successive GOF.

7.2.5. RNRT-COSOCATI

Participants: Pierre Siohan, Claudio Weidmann.

- Convention number : 2 02 C 0745 3132401.
- Title : *Joint source-channel coding for image transmission.*
- Research axis : § 6.2.4
- Partners : France Télécom R&D, Thalès (in replacement d'Alcatel Space Industries), Inria, CNES, GET, I3S.
- Funding : Ministry of industry.
- Period : Mar.00-Oct.03.

Cosocati is an exploratory RNRT project. Its goal is to study innovative joint source channel coding (JSCC) and decoding (JSCD) techniques for transmission of images over radio-mobile channels and for applications of earth observation from satellites. This year Temics has contributed to various improvements of JSCD techniques for Huffman codes and some of their variants (RVLC). Methods for online estimation at the receiver of the source statistics related to these VLCs have also been proposed. JSCD techniques, including estimation modules, have been tested with the H.263++ video standard using different models of transmission channels: Gaussian, Rayleigh and UMTS. Compared to classical hard decoding techniques of VLCs, significant improvements have been obtained.

7.3. European contracts

7.3.1. IST-Busman

Participants: Vincent Bottreau, Jonathan Delhumeau, Teddy Furon, Christine Guillemot.

- Convention number : 1 02 C 0186 00 00MPR 00 5.
- Title : Bringing user satisfaction to media access networks.
- Research axis : § 6.3.2.
- Partner : British Telecom, Framepool, HHI, INRIA, Motorola, QMUL, Telefonica, University of Munich.
- Period : Apr.02-Sept.04.
- Funding : CEE.

BUSMAN develops and integrates indexing and watermarking techniques to ease the search and use of video content. TEMICS contribution is focused on watermarking techniques to enrich video content by hiding meta-data. Different levels of robustness to a range of attacks such as transcoding required for transmission in heterogeneous networks, and compression at various rates must be provided. A first version of the software based on state-of-the-art techniques has been delivered for integration in the BUSMAN demonstrator comprising an authoring tool as well as a client-server delivery architecture over fixed and mobile networks. The technique should evolve to incorporate side-informed embedding solutions in order to improve the performance (see sub-section 6.3.2).

7.3.2. *IST-Ozone*

Participants: Christine Guillemot, Henri Nicolas.

- Convention number: 1 01 A0672 00 000MC 00 5
- Title: *New technologies and services for emerging nomadic societies.*
- Research axis: § 5.1
- Partners: Inria, Thomson Multimedia, Philips, IMEC, Epictoid.
- Funding: CEE.
- Period: Nov.01-May.04.

The goal of the OZONE project is to develop a pervasive computing and communication framework which will bring relevant information and services to the individual, anywhere and at anytime. The OZONE project can be viewed as the first step towards concrete ambient intelligence applications. Our contributions to the OZONE project is related to the transmission of the video data throughout the OZONE network. High quality performance of the video transmission system is essential to guarantee the quality of service required by user's needs in the context of such applications. Our contribution is related to the study and development of a video transmission platform incorporating mechanisms in support of end-to-end QoS such as congestion control and loss control. The software corresponding to the video transmission loss and network congestion control has been delivered to the project's partners.

7.3.3. *FP6-IST NoE SIMILAR*

Participants: Christine Guillemot, Luce Morin, Henri Nicolas.

The TEMICS team is involved in the network of excellence SIMILAR federating European fundamental research on multimodal human-machine interfaces, accepted in July 2003 as a result of the first call of the 6th framework programme and starting on the 1st of January 2004.

7.3.4. *FP6-IST STREP DANAE*

Participants: Robin Chatterjee, Christine Guillemot, Hervé Jégou, Gagan Rath.

The TEMICS team is involved in the STREP DANAE addressing issues of dynamic and distributed adaptation of scalable multimedia content in a context-aware environment. Its objectives are to specify, develop, integrate and validate in a testbed a complete framework able to provide end-to-end quality of (multimedia) service at a minimal cost to the end-user. The project has been accepted in July 2003 as a result of the first call of the 6th framework programme and starts on the 1st of January 2004.

7.3.5. *FP6-IST STREP ENTHRONE*

Participants: Christine Guillemot, Slavica Marinkovic.

The TEMICS team is involved in the Integrated project ENTHRONE aiming at an integrated management solution which would covers an entire audio-visual service distribution chain, including content generation and protection, distribution across networks and reception at user terminals. The project will address issues of scalable generation, protection, distribution and usage of multimedia content in heterogeneous environments. The project has been accepted in July 2003 as a result of the first call of the 6th framework programme and starts on the 1st of January 2004.

8. Other Grants and Activities

8.1. Regional initiatives

8.1.1. *Britanny region contract on watermarking*

Participants: Gaetan Le Guelvouit, Stéphane Pateux.

This contract supplies 50to Gaetan Le Guelvouit. The thesis started in Oct. 2000 and ended in Sept. 2003. The results achieved are described in sub-section 6.3.1.

8.2. National initiatives

8.2.1. ARC Télégeo

Participants: Raphaèle Balter, Luce Morin, Stéphane Pateux.

- Title : Télégeo: Geometry and Telecommunications.
- Research axis : § 6.3.1.
- Partners : Creatis-Insa de Lyon, ENST Paris, INRIA (ISA, TEMICS, PRISME).
- Funding : INRIA.
- Period : June 02 - June 04.

This ARC (Action de Recherche Coopérative) aims at creating a synergy in the area of geometric objects transmission over networks, and more specifically to study the representation of geometric objects for their transmission over heterogeneous networks. TEMICS contributes by providing compression algorithms for unstructured surface meshes, and techniques for progressive and scalable compression taking into account visual quality criteria.

8.2.2. ACI Fabriano

Participants: François Cayre, Teddy Furon.

- Title : Fabriano
- Research axis : § 6.3.1.
- Partners : CERDI, INRIA (TEMICS), LIS, LSS.
- Funding : Ministry of research, CNRS, INRIA.
- Period : Mid-Dec. 03 - Dec. 06.

Fabriano is an ACI (Action Concertée Incitative) dedicated to the study of technical solutions to the problem of security based on watermarking and steganography. In particular, this action aims at developing a theoretical framework for stegano-analysis to be applied for the design of algorithms that will allow to detect the presence of a message within a signal in the respect of rights and ethical issues.

8.2.3. CNRS specific action on watermarking

Participants: Teddy Furon, Stéphane Pateux.

- Title : Watermarking and data hiding for audiovisual communication.
- Research axis : § 6.3.1.
- Partners : UTC, INPG, ENST, Supelec, LIFL.
- Funding : CNRS.
- Period : Jan.03-Dec.03.

This collaboration aims at gathering national expertise in the field of watermarking. The goal is to draw prospective research directions addressing the following issues : watermarking of multi-component data (color images but also multi-modal signals), enriched media based on data-hiding, integrity control, steganography and steganalysis (in order to define security properties of data-hiding techniques).

8.2.4. CNRS specific action on scalable and robust video compression

Participant: Christine Guillemot.

- Title : Scalable and robust compression of video signals.
- Research axis : § 6.2.1, 6.2.2.
- Partners : INRIA, ENST, I3S-Université de Nice-Sophia Antipolis, Labri-University of Bordeaux.
- Funding : CNRS.
- Period : Jan.03-Dec.03.

The objective of the collaboration is to foster synergy among the main national actors in the field of video coding in order to prepare joint responses to calls for project proposals both at a national and European level. The precise goals are thus to identify new scientific and technical challenges and to define joint research directions in the areas of compression-oriented video analysis, spatio-temporal wavelet based compression, multiple description coding and distributed source coding.

8.3. Bi-lateral international co-operation

8.3.1. Project reviewing

- C. Guillemot has been appointed by the European commission to review project submissions within the IST programme of the 6th framework programme.
- C. Guillemot has been appointed by the Ministry of Research of the Wallonie region in Belgium to review research project submissions.
- C. Guillemot expertised two company projects for ANVAR.
- C. Guillemot is member of the commission 2 of the french RNRT national programme and as such serves as an evaluator of project proposal submissions.

9. Dissemination

9.1. Patents

- T. Guionnet and C. Guillemot, "Method for robust decoding of arithmetic codes", Inria patent No. 03 03288, 2003.
- N. Cammas, S. Pateux and N. Laurent, "Procédés et dispositifs de codage et décodage d'une séquence d'images par décomposition mouvement/texture et codage par ondelettes", France Télécom/Inria patent No. 0303449, 2003.
- J. Viéron, G. Marquant, G. Boisson, P. Robert, E. François and C. Guillemot, "Method for scalable encoding and decoding of an interlaced sequence of digital video data", Thomson/Inria patent No. 03290507, 2003.

9.2. Standardization

- T. Guionnet and C. Guillemot, "Soft decoding and synchronization of arithmetic and quasi-arithmetic codes", contribution to JPEG-2000/JWL, ISO/IEC JTC 1/SC 29/WG 1, N3013, July 2003;
- G. Marquant, J. Viéron, G. Boisson, P. Robert, E. François and C. Guillemot, "Response to the call for evidence on scalable video coding advances", joint Thomson/INRIA contribution to MPEG/SVC, ISO/IEC JTC1/SC29/WG M9784, Trondheim, July 2003;
- G. Marquant, J. Viéron, G. Boisson, P. Robert, E. François, C. Guillemot, "Demonstrations of scalable video coding on interlaced material", joint Thomson/INRIA contribution to MPEG/SVC, ISO/IEC JTC1/SC29/WG11 M9783, Trondheim, July 2003.
- R. Balter, P. Gioia, F. Galpin and L. Morin, "Comments on Model-based video streaming", MPEG/3DAV, ISO/JTC1/SC29/WG11 M10051, Brisbane, October 2003.

9.3. Leadership within the scientific community

- TEMICS has organized the Picture Coding Symposium 2003 in Saint-Malo, April 2003. C. Guillemot has served as general co-chair of the conference, L. Morin, H. Nicolas, S. Pateux, P. Siohan have served on the technical and organizing committees.
- C. Guillemot is associated editor of the journal IEEE Transactions on Image Processing;
- C. Guillemot is elected member of the international committee IEEE IMDSP (Image and MultiDimensional Signal Processing Technical Committee);
- C. Guillemot is member of program committees of the following conferences: SPIE-VCIP 2003, IEEE-ICIP 2003, ACM Multimedia 2003, IEE-VIE 2003, CORESA 2003;
- C. Guillemot is member of the steering committee of the CNRS network (RTP) on "ambient networks";
- H. Nicolas is member of the program committees of SPIE-VCIP 2003 and IEEE-ICIP 2003;
- H. Nicolas is member of the "commission de spécialistes" of the University of Rennes 1.
- S. Pateux is scientific coordinator for IRISA at the CNRT (Centre National de Recherche Technologique) TIM-Bretagne;
- S. Pateux is co-coordinator with F. Davoine from UTC for the publication of a book on watermarking of multimedia signals in the context of the IC2 collection of Hermès.

9.4. Invitations

- C. Guillemot was invited for a one-week visit at the University of Stuttgart, Germany;
- S. Pateux was invited for a one-week visit at the University of Dublin, Ireland.

9.5. Teaching

- Diic-INC, Ifsic, university of Rennes 1 (L. Morin, H. Nicolas, C. Guillemot : image processing, 3D vision, motion, coding, compression, cryptography, communication) ;
- Diic-INC, Ifsic, Dess-Mitic, university of Rennes 1 (S. Pateux : Watermarking multimedia documents) ;
- Diic-LSI, Ifsic, university of Rennes 1 (H. Nicolas : compression) ;
- DEA Stir, option Image, university of Rennes 1 (C. Labit, H. Nicolas : compression) ;
- DEA Computer Science, university of Rennes 1 (C. Guillemot, S. Pateux : Video compression and communication) ;
- Enic, Villeneuve-d'Ascq, ENSTBR (C. Guillemot: Video communication) ;
- Ensar Rennes (L. Morin : Basics of image processing, and mathematical morphology) ;
- DEA ELECTRICAL ENGINEERING, INSA, university of Rennes 1 (P. Siohan : Digital TV) ;
- ESIGETEL Fontainebleau, (S. Pateux : Watermarking multimedia signals; C. Guillemot : Video compression and communication) ;
- Project Cian, Breton Digital Campus (L. Morin : Digital Images) ;

10. Bibliography

Major publications by the team in recent years

- [1] C. GUILLEMOT, T. TURLETTI. *Systèmes multimédia communicants*. Éditions Hermès, 2000, chapter Algorithmes de codage et de contrôle de transmission sur réseaux hétérogènes, pages 25-82.
- [2] 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 », number 9, volume 19, sept., 2001, pages 1680-1696.
- [3] J. P. LEDUC, J. M. ODOBEZ, C. LABIT. *Adaptive motion-compensated wavelet filtering for image sequence coding*. in « IEEE Transactions on Image Processing », number 6, volume 6, June, 1997, pages 862-879.
- [4] R. MOHR, L. MORIN, E. GROSSO. *Relative positioning with uncalibrated cameras*. in « in : *Geometric Invariance in Computer Vision*, Editors: A. Zisserman J.L. Mundy, MIT Press », 1992, pages 440-460.
- [5] H. NICOLAS, C. LABIT. *Motion and illumination variation estimation using a hierarchy of models: application to image sequence coding*. in « Journal of Visual Communication and Image Representation », number 4, volume 6, December, 1995, pages 303-316.
- [6] P. RAULT, C. GUILLEMOT. *Symmetric delay factorization: generalized framework for paraunitary filter banks*. in « IEEE Trans. on signal processing », number 12, volume 47, Dec., 1999, pages 3315-3325.

Doctoral dissertations and “Habilitation” theses

- [7] M. CHAUMONT. *Représentation en objets vidéo pour un codage progressif et concurrentiel des séquences d’images*. Ph. D. Thesis, Université de Rennes 1, Nov., 2003.
- [8] T. GUIONNET. *Codage robuste par descriptions multiples pour transmission sans fil d’information multimédia*. Ph. D. Thesis, Université de Rennes 1, March, 2003.
- [9] J. VIÉRON. *Codage scalable et contrôle de congestion pour transmission vidéo sur réseaux hétérogènes*. Ph. D. Thesis, Université de Rennes 1, April, 2003.
- [10] G. LE GUELVOUT. *Tatouage robuste par étalement de spectre avec prise en compte de l’information adjacente*. Ph. D. Thesis, Université de Rennes 1, Nov., 2003.

Articles in referred journals and book chapters

- [11] J. ANDREAUX, T. FURON, A. DURAND, E. DIEHL. *SmartRight: A Copy Protection System for Digital Home Networks*. in « accepted for publication in IEEE Signal Processing Magazine, special issue on digital right management. ».
- [12] M. BARNI, F. BARTOLINI, T. FURON. *A general framework for robust watermarking security*. in « Signal Processing », number 10, volume 83, October, 2003, pages 2069-2084, Special issue on Security of Data Hiding Technologies, invited paper.

- [13] F. GALPIN, L. MORIN, K. DEGUCHI. *Compression performance of computer vision based coding*. in « IEICE Transactions on Information and Systems », number 3, volume E86-D, March, 2003, pages 10–15.
- [14] C. GUILLEMOT, P. CHRIST. *Joint source-channel coding as an element of a QoS framework for 4G wireless multimedia*. in « accepted for publication in Elsevier Journal on Computer Communications, special issue on research directions for 4G wireless networks ».
- [15] T. GUIONNET, C. GUILLEMOT. *Soft and joint source-channel decoding of quasi-arithmetic codes*. in « accepted for publication in Eurasip Journal on Applied Signal Processing », 2003.
- [16] T. GUIONNET, C. GUILLEMOT. *Soft decoding and synchronization of arithmetic codes: Application to image transmission over noisy channels*. in « IEEE Transactions on Image Processing », volume 12, Dec., 2003.
- [17] R. MOTWANI, C. GUILLEMOT. *Tree-structured oversampled filter banks as joint source-channel codes: Application to image transmission over erasure channels*. in « accepted for publication in IEEE Transactions on Signal Processing », 2003.
- [18] L. OISEL, É. MÉMIN, L. MORIN, F. GALPIN. *One-dimensional dense disparity estimation for three-dimensional reconstruction*. in « IEEE Transactions on Image Processing », number 9, volume 12, September, 2003, pages 1107–1119.
- [19] S. PATEUX, G. LE GUELVUIT. *Practical watermarking scheme based on wide spread spectrum and game theory*. in « Signal Processing : Image Communication », number 18, apr, 2003, pages 283–296.
- [20] D. PINCHON, P. SIOHAN, C. SICLET. *Design techniques for orthogonal modulated filter banks based on a compact representation*. in « accepted for publication in IEEE Transactions on Signal Processing », 2003.
- [21] L. QUAN, L. MORIN, L. OISEL. *Reconstruction de scènes 3D à partir de vues multiples*. series Chapitre dans le traité IC2 sur perception visuelle par imagerie vidéo, de la série Traitement du Signal et de l'Image, Hermès, 2003, pages 179–211.
- [22] G. RATH, C. GUILLEMOT. *Frame-theoretic analysis of DFT codes with erasures*. in « Accepted for publication in IEEE Transactions on Signal Processing », 2003.
- [23] G. RATH, C. GUILLEMOT. *Performance analysis and recursive syndrome decoding of DFT codes for bursty erasure recovery*. in « IEEE Transactions on Signal Processing », number 5, volume 51, May, 2003, pages 1335–1350.
- [24] G. RATH, C. GUILLEMOT. *Subspace-based method for error and erasure recovery with DFT codes for wireless channels*. in « Accepted for publication in IEEE Transactions on Signal Processing », 2003.
- [25] P. SIOHAN, C. ROCHE. *Derivation of extended Gaussian functions based on the Zak transform*. in « accepted for publication in IEEE Signal Processing Letters », 2003.
- [26] J. VIÉRON, C. GUILLEMOT. *Real-time constrained TCP-compatible rate control for video over the Internet*. in « accepted for publication in IEEE Transactions on Multimedia », 2003.

- [27] J. VIÉRON, T. TURLETTI, K. SALAMATIAN, C. GUILLEMOT. *Source and channel adaptive rate control for multicast layered video transmission based on a clustering algorithm*. in « accepted for publication in Eurasip Journal on Applied Signal Processing, special issue on Multimedia over wireless IP », 2003.

Publications in Conferences and Workshops

- [28] R. BALTER, F. GALPIN, L. MORIN. *Morphing 3D automatique de cartes de profondeur maillées*. in « Journées d'études et d'échanges CORESA'2003, Lyon, France », pages 223–226, Jan., 2003.
- [29] R. BALTER, F. GALPIN, L. MORIN. *Very Low bitrate compression of video sequence for virtual navigation*. in « Picture Coding Symposium, PCS'2003, Saint-Malo, France », pages 305–308, April, 2003.
- [30] R. BALTER, L. MORIN. *Morphing 3D bidirectionnel entièrement automatique*. in « Journées Francophones des Jeunes Chercheurs en Vision par Ordinateur, ORASIS 2003 », pages 363–371, March, 2003.
- [31] G. BOISSON, E. FRANÇOIS, D. THOREAU, C. GUILLEMOT. *Motion-Compensated Spatio-temporal context-based arithmetic coding for full scalable video compression*. in « Picture Coding Symposium, PCS'2003 », Apr., 2003.
- [32] N. CAMMAS, S. PATEUX. *Codage vidéo scalable par maillage et ondelettes 3D*. in « CORESA'03 », pages 151-154, Lyon, France, January, 2003.
- [33] N. CAMMAS, S. PATEUX. *Fine-grain scalable video coding using 3D wavelets and active meshes*. in « IVCP'03. Proc. of SPIE », volume 5022, Santa Clara, California, USA, January, 2003.
- [34] V. CHAPPELIER, C. GUILLEMOT, S. MARINKOVIC. *Turbo Trellis Coded Quantization*. in « Proc. of the Intl. symp. on turbo codes », Sept., 2003.
- [35] M. CHAUMONT, N. CAMMAS, S. PATEUX. *Fully scalable object based video coder based on analysis-synthesis scheme*. in « International Conference on Image Processing, ICIP'2003 », September, 2003.
- [36] M. CHAUMONT, S. PATEUX, H. NICOLAS. *Codage efficace de contour avec perte utilisant la consistance spatio-temporelle*. in « GRETSI'2003 », September, 2003.
- [37] M. CHAUMONT, S. PATEUX, H. NICOLAS. *Efficient lossy contour coding using spatio-temporal consistency*. in « Picture Coding Symposium, PCS'2003 », pages 289-294, April, 2003.
- [38] D. CHAVIRA-MARTINEZ, S. PATEUX. *Codage vidéo basé objet à granularité fine*. in « GRETSI'03 », volume 2, pages 303–306, Paris, Sept., 2003.
- [39] D. CHAVIRA-MARTINEZ, S. PATEUX. *Non-rigid object tracking using a likelihood spatio-temporal model*. in « Visual Communications and Image Processing », volume 5150, pages 334-345, July, 2003.
- [40] J. DELHUMEAU, T. FURON, N. HURLEY, G. SILVESTRE. *Improved Polynomial Detectors for Side-Informed Watermarking*. in « Security and Watermarking of Multimedia Contents IV », SPIE Electronic Imaging, pages 311-321, Santa Clara, Cal., USA, January, 2003.

- [41] J. DELHUMEAU, L. MORIN, S. PATEUX. *Segmentation basée mouvement 3D pour la détection d'objets indépendants*. in « Journées d'études et d'échanges CORESA'2003, Lyon, France », pages 209–212, Jan., 2003.
- [42] T. GUIONNET, C. GUILLEMOT. *Robust decoding of arithmetic codes for image transmission over error-prone channels*. in « Proc. of the IEEE Intl. Conference on Image Processing, ICIP'03 », Sept., 2003.
- [43] M. JEANNE, P. SIOHAN, J.-C. CARLACH. *Comparaison de deux approches du décodage conjoint source-canal*. in « Grets'i'03 », Paris, France, September, 2003.
- [44] M. JEANNE, R. ZHANG, B. GIROD, P. SIOHAN. *Codage de sources distribuées : comparaison de 2 systèmes approchant la borne de Wyner-Ziv*. in « CORESA'03 », Lyon, France, January, 2003.
- [45] M. JEANNE, R. ZHANG, B. GIROD, P. SIOHAN. *Distributed source coding : comparison of two methods close to the Wyner-Ziv bound*. in « ISIT'03 », Yokohama, Japan, July, 2003.
- [46] H. JÉGOU, C. GUILLEMOT. *Error-resilient binary multiplexed source codes*. in « Proc. of the IEEE Intl. Conference on Accoustic, Speech and Signal Processing, ICASSP'03 », April, 2003.
- [47] H. JÉGOU, C. GUILLEMOT. *Source multiplexed codes for error-prone channels*. in « Proc. of the IEEE Intl. Conference on Communications, ICC'03 », April, 2003.
- [48] R. MOTWANI, C. GUILLEMOT. *Filter bank frame expansions as joint source-channel codes for robustness to erasures*. in « Proc. of the IEEE Intl.symposium on information theory, ISIT'03 », July, 2003.
- [49] R. MOTWANI, C. GUILLEMOT. *Quantized frame expansions based on tree-structured filter banks for erasure recovery*. in « Proc. of the IEEE Intl. Conference on Accoustic, Speech and Signal Processing, ICASSP'03 », April, 2003.
- [50] R. MOTWANI, C. GUILLEMOT. *Two-channels oversampled filter banks as joint source-channel codes for erasure channels*. in « Proc. of the IEEE Intl. Conference on Image Processing, ICIP'03 », Sept., 2003.
- [51] M. MULLARKEY, N. HURLEY, G. SILVESTRE, T. FURON. *Application of Side-Informed Embedding and Polynomial Detection to Audio Watermarking*. in « Int. Conf. on Multimedia and Expo », IEEE, Baltimore, MY, USA, July, 2003.
- [52] D. PINCHON, C. SICLET, P. SIOHAN. *Synthèse de bancs de filtres modulées comprenant un nombre de sous-bandes arbitrairement grand*. in « Grets'i'03 », Paris, France, September, 2003.
- [53] D. PINCHON, P. SIOHAN, C. SICLET. *A fast design method for biorthogonal modulated filter banks*. in « Proc. of 11th European Signal Processing Conference », Toulouse, France, September, 2002.
- [54] G. RATH, C. GUILLEMOT. *A subspace algorithm for error localization with overcomplete frame expansions and its application to image transmission*. in « IEEE International Conference on Image Processing, ICIP'03 », Sept., 2003.

- [55] G. RATH, C. GUILLEMOT. *A subspace based method for error correction with DFT codes*. in « IEEE International Conference on Communications ICC, 2003, Anchorage, Alaska », May, 2003.
- [56] G. RATH, C. GUILLEMOT. *Characterization of a class of error correcting frames and their application to image transmission*. in « Picture Coding Symposium, PCS'2003, Saint-Malo, France », April, 2003.
- [57] G. RATH, C. GUILLEMOT. *Error and erasure correction using overcomplete frame expansions*. in « IEEE Signal Processing Workshop on Signal Processing Advances in Wireless Communications, Rome, Italy », June, 2003.
- [58] G. RATH, C. GUILLEMOT. *ESPRIT-like Error Localization Algorithm for a Class of Real Number Codes*. in « IEEE Global Communications Conference, GlobeCom'03 », Dec., 2003.
- [59] G. RATH, C. GUILLEMOT. *Subspace Algorithms for Error Localization with DFT Codes*. in « IEEE ICASSP 2003, Hong Kong, China », April, 2003.
- [60] C. SICLET, P. SIOHAN. *Weyl-Heisenberg signal expansions over R in $L_2(Z)$ and duality relations involving MDFT filter banks*. in « Proc. International Conference on Acoust., Speech and Signal Processing », Hong Kong, China, April, 2003.
- [61] J. VIÉRON, C. GUILLEMOT. *Low rate FGS video compression based on motion-compensated spatio-temporal wavelet analysis*. in « Proc. of the SPIE Intl Conference on Visual Communication and Image Processing, VCIP'03 », July, 2003.
- [62] C. WEIDMANN, P. SIOHAN. *Décodage conjoint source-canal avec estimation en ligne de la source*. in « CORESA'03 », Lyon, France, January, 2003.
- [63] C. WEIDMANN. *Reduced complexity soft-in soft-out decoding of variable-length codes*. in « ISIT'03 », Yokohama, Japan, July, 2003.
- [64] A. MANOURY AND H. NICOLAS. *Segmentation temporelle de vidéos numériques en vue de la définition de structure vidéo et utilisation de mosaïques 1D pour la classification*. in « Proceedings of GRETSI, Paris, France », September, 2003.
- [65] J. VIÉRON AND C. GUILLEMOT AND H. NICOLAS. *Compression vidéo FGS à bas débit basée sur une décomposition ondelettes 2D+t*. in « Proceedings of CORESA, Lyon, France », January, 2003.
- [66] J.M. PINEL AND H. NICOLAS. *Moving shadows segmentation on Lambertian surfaces in video sequences*. in « Proceedings of VCIP, Lugano, Switzerland », July, 2003.
- [67] G. LE GUELVOUIT, S. PATEUX, J. DELHUMMEAU. *Construction de codes pour tatouage avec prise en compte de l'information adjacente*. in « GRETSI Symp. on Image and Signal Processing », Paris, France, Sept., 2003.
- [68] G. LE GUELVOUIT, S. PATEUX. *Wide spread spectrum watermarking with side information and interference cancellation*. in « Proc. SPIE », Santa Clara, CA, January, 2003.

-
- [69] M. CHAUMONT, S. PATEUX AND H. NICOLAS. *Efficient lossy contours coding using spatio-temporal consistency*. in « Proceedings of PCS, St Malo, France », April, 2003.
- [70] M. GARCIA AND H. NICOLAS. *An efficient and direct non-planar rotation estimation algorithm for video applications*. in « Proceedings of VCIP, Lugano, Switzerland », July, 2003.
- [71] M. GARCIA AND H. NICOLAS. *Estimation of Non-planar Rotation for Video Coding Application*. in « Proceedings of PCS, St Malo, France », April, 2003.
- [72] M. GARCIA AND H. NICOLAS. *Non-planar rotation estimation for video object trajectory analysis*. in « Proceedings of ICME, Baltimore, USA », July, 2003.