INRIA

Project-Team TEMICS

TraitEment, Modélisation d'Images et CommunicationS

Rennes

THEME COG

Activity Report

2004

Table of contents

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]
Gagan Rath [CR]

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
Thomas Guionnet [VIP project, until Sept. 30th, 2004]
Sophie Le Delliou [DIPHONET project, 50% TEMICS - 50% TexMex, until Sept. 30th, 2004]

Inria junior technical staff
Cécile Marc

University teaching and research assistant
Marc Chaumont [ATER, until Aug. 30th, 2004]

Inria post-doctoral fellows
François Cayre
Ceilidh Hoffman [50% TEMICS, 50% PLANETE Sophia-Antipolis, from Sept. 1st, 2004]
Marion Jeanne [COPARO Project, from Oct. 1st, 2004]

Ph.D Students
Raphaèle Balter [Cifre grant, France Télécom]
Guillaume Boisson [Cifre grant, Thomson Multimedia, until 30th of Sept. 2004]
Nathalie Cammas [Cifre grant, France Télécom]
Vivien Chappelier [Ministry of research grant]
Dubhe Chavira-Martinez [Conacyt grant, until June 30th, 2004]
Robin Chatterjee [Inria grant, DANAE project]
Jonathan Delhumeau [Inria grant, BUSMAN project]
Mireya Garcia-Vasquez [Conacyt grant, until June 30th, 2004]
Hervé Jégou [Ens-Cachan grant]
Khaled Lajnef [Inria grant, VIP project]
Gael Sourimant [Inria grant, CRE France Télécom, from Oct. 1st, 2004]
François Tonnin [Inria- Brittany Region grant, also with Texmex]

2. Overall Objectives

2.1. introduction

The goal of the TEMICS project is the design and development of theoretical frameworks as well as algorithms and practical solutions in the areas of analysis, modelling, coding, communication and watermarking

of images and video signals. 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 characterized by narrow-band, error and/or loss prone, heterogeneous and time-varying channels, is creating challenging problems in the area of source and channel coding. Design principles prevailing so far and stemming from Shannon's source and channel separation theorem must be re-considered. The separation theorem, 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

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

3D reconstruction is the process of estimating the shape and position of 3D objects from views of these objects. TEMICS deals more specifically with the modelling of large scenes from monocular video sequences. 3D reconstruction using projective geometry is by definition an inverse problem. Some key issues which do not have yet satisfactory solutions are the estimation of camera parameters, especially in the case of a moving camera. Specific problems to be addressed are e.g. the matching of features between images, and the modelling of hidden areas and depth discontinuities.

3D reconstruction uses theory and methods from the areas of computer vision and projective geometry. When the camera $\mathcal{C}_i$ is modelled as a *perspective projection*, the *projection equations* are :

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

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

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

Signal representation using orthogonal basis functions (e.g., DCT, wavelet transforms) is at the heart of source coding. The key to signal compression lies in selecting a set of basis functions that compacts the signal energy over a few coefficients. Frames are generalizations of a basis for an overcomplete system, or in other words, frames represent sets of vectors that span a Hilbert space but contain more numbers of vectors than a basis. Therefore signal representations using frames are known as overcomplete frame expansions. Because of their inbuilt redundancies, such representations can be useful for providing robustness to signal transmission over error-prone communication media.

Consider a signal $\mathbf{x}$. An overcomplete frame expansion of $\mathbf{x}$ can be given as $F\mathbf{x}$ where F is the frame operator associated with a frame $\Phi_F \equiv \{\phi_i\}_{i \in I}$, ϕ's are the frame vectors and I is the index set. The ith frame expansion coefficient of $\mathbf{x}$ is defined as $(F\mathbf{x})_i \equiv \langle \phi_i, \mathbf{x} \rangle$, 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 $\tilde{\Phi}_F \equiv \{(F^h F)^{-1}\phi_i\}_{i \in I}$. Tight frame expansions, where the frames are self-dual, are analogous to orthogonal expansions with basis functions.

Frames in finite-dimensional Hilbert spaces such as $\mathbf{R}^K$ and $\mathbf{C}^K$, known as discrete frames, can be used to expand signal vectors of finite lengths. In this case, the frame operators can be looked upon as redundant block transforms whose rows are conjugate transposes of frame vectors. For a K-dimensional vector space, any set of N, $N > K$, vectors that spans the space constitutes a frame. Discrete tight frames can be obtained from existing orthogonal transforms such as DFT, DCT, DST, etc by selecting a subset of columns from the respective transform matrices. Oversampled filter banks can provide frame expansions in the Hilbert space of square summable sequences, i.e., $l_2(\mathbf{Z})$. In this case, the time-reversed and shifted versions of the impulse responses of the analysis and synthesis filter banks constitute the frame and its dual.

Since overcomplete frame expansions provide redundant information, they can be used as joint source-channel codes to fight against channel degradations. In this context, the recovery of a message signal from the corrupted frame expansion coefficients can be linked to the error correction in infinite fields. For example, for discrete frame expansions, the frame operator can be looked upon as the generator matrix of a block code in the real or complex field. A parity check matrix for this code can be obtained from the singular value decomposition of the frame operator, and therefore the standard syndrome decoding algorithms can be utilized to correct coefficient errors. The structure of the parity check matrix, for example the BCH structure, can be used to characterize discrete frames. In the case of oversampled filter banks, the frame expansions can be looked upon as convolutional codes.

3.3. Rate-distortion theory

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

Coding and joint source channel coding rely on fundamental concepts of information theory, such as notions of entropy, memoryless or correlated sources, of channel capacity, or on rate-distortion performance bounds. Compression algorithms are defined to be as close as possible to the optimal rate-distortion bound, $R(D)$, for a given signal.

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

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

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

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

Digital watermarking aims at hiding discrete messages into multimedia content. The watermark must not spoil the regular use of the content, i.e., the watermark should be non perceptible. Hence, the embedding is usually done in a transformed domain where a human perception model is exploited to assess the non perceptibility criterion. The watermarking problem can be regarded as a problem of creating a communication channel within the content. This channel must be secure and robust to usual content manipulations like lossy compression, filtering, geometrical transformations for images and video.

When designing a watermarking system, the first issue to be addressed is the choice of the transform domain, i.e., the choice of the signal components that will *host* the watermark data. Let $E(.)$ be the extraction function going from the content space $\mathcal{C}$ to the components space, isomorphic to $\mathbf{R}^N$

$$E(.): \quad \begin{matrix} \mathcal{C} & \mapsto & \mathbf{R}^N \\ C & \to & \mathbf{V} = E(C) \end{matrix}$$

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

$$S(.): \quad \begin{matrix} \mathcal{M} & \mapsto & \mathbf{R}^N \\ m & \to & \mathbf{S} = S(m) \end{matrix}$$

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

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

A new paradigm stating that the original host signal $\mathbf{V}$ shall be considered as a *channel state* only known at the embedding side rather than a source of noise, 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.
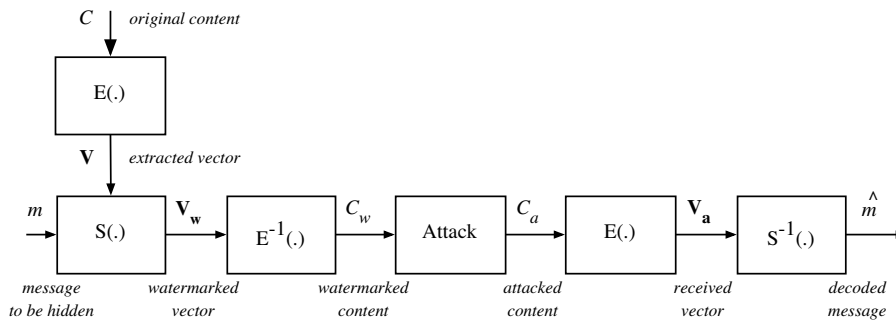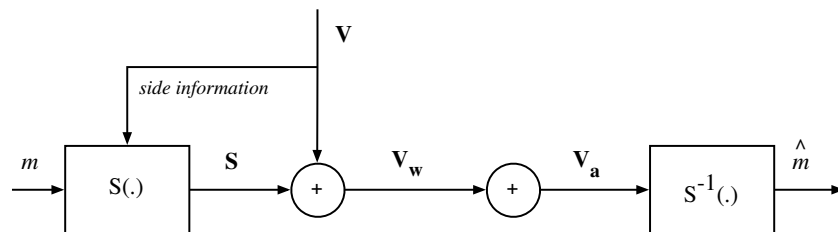

Figure 1. Classical watermarking scheme

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

4. Application Domains

4.1. Panorama

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 addresses 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 traveling 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. Video communication platform

Participants: Christine Guillemot, Laurent Guillo [correspondant], Thomas Guionnet, Cécile Marc.

With the support of several contracts (RNRT-VISI, IST-SONG), TEMICS had started the development of a video communication platform. This platform provides a test bed allowing to study and to assess, in a realistic way, new algorithms implementing joint source channel coding, video modelling or video coding. The platform is still under development. In 2004, a collaborative effort between the ARMOR and the TEMICS project-teams has led to the development and integration of new or improved components, which are described below.

5.1.1. MOVIQS: Video communication with QoS support

Participant: Laurent Guillo [correspondant].

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.

5.1.2. WAVIX: Wavelet based video codec

Participants: Chrisitine Guillemot, Laurent Guillo [correspondant], Thomas Guionnet, Cécile Marc.

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 subbands, the first release used the EBCOT algorithm provided by the Verification Model JPEG2000. That release 1.0 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. Wavix now takes advantage of three main improvements. First, the JPEG2000 library is replaced by JasPer, which performs better. Then, the movement and texture information have been embedded in a single bitstream. That makes transmission easier. Finally, header protections and soft decoding techniques for quasi-arithmetic codes are now included. These techniques improve the quality of the decoded video when the video is transmitted over a noisy channel. Moreover, the decoding time can be reduced if WAVIX knows what parts of the bitstream are corrupted or not. This last information can be provided by a specific transport layer such as UDP-Lite.

5.1.3. WULL: A Windows UDP-Lite library

Participants: Laurent Guillo [correspondant], Cécile Marc.

WULL is Windows library that implements the UDP-Lite protocol according to the very new RFC 3828. UDP-Lite is a new transport protocol that is used by application, which would rather receive damaged data than having corrupted data discarded by the network. UDP-Lite is similar to UDP. However, UDP-Lite allows applications to specify the length of checksum coverage. Checksum coverage is the number of bytes, counting from the first byte of the UDP-Lite header, that are covered by the checksum. So, a packet is divided into an error sensitive part (covered by the checksum) and an error insensitive part (not covered by the checksum). Only errors in the sensitive part cause the packet to be discarded by the transport layer. UDP-Lite is especially useful for application that are error tolerant, such video codec and WAVIX in particular. Using UDP-Lite as transport protocol in our video communication platform allows the transport layer to signal to WAVIX that it is going to process corrupted data. Taking into account this information can significantly improve the decoding speed. Wull provides developers with specific UDP-Lite sockets they can use in the same way they use classic UDP socket. It can be used with both IPv4 and IPv6. To get a more mature software, the ARMOR and TEMICS project-team run interoperability tests with an other UDP-Lite implementation carried out by the team headed by Paul Christ of the University of Stuttgart. As a partner in the DANAE project, France Telecom also uses WULL. WULL 1.0 has been registered at the Agency for the Protection of Programmes (APP) under the number IDDN.FR.001.270018.000.S.P.2004.000.10000

5.1.4. ROHC: Robust header compression

Participants: Laurent Guillo [correspondant], Cécile Marc.

The IETF has standardized the ROHC (RObust Header Compression) protocol. This mechanism drastically reduces the header size of the network packets and consequently increases performances on slow and noisy links. First, a collaborative work with the ARMOR project-team and the ENST-Bretagne school lead to an implementation of the RFC 3095 for both IPv4 and IPv6. Then, as WULL was available, we have extended this first implementation by adding two new profiles IP/UDP-Lite/RTP and IP/UDP-Lite according to the IETF draft draft-ietf-rohc-udp-lite-02. Tests are still being run on the video communication platform.

5.2. CHI-MARK2: Robust image watermarking tool

Participants: Sophie Le Delliou, Teddy Furon [contact person].

This software implements several data-hiding techniques (embedding and extraction) for images and video. It has been created last year and reported in the last annual rapport. Bugs correction and improvements in 2004 were mainly driven by the requirements of Canon, prime of the DIPHONET national project. This concerns the creation of usage profiles, procedures for a zero-knowledge watermarking decoding, and removal of black frames. These modifications were grouped in one update version (v3.1, january 2004) of the CHI-MARK2 software (IDDN.FR.001.480027.001.S.A.2002.000.41100. at the Agency for the Protection of Programmes (APP)).

5.3. Independent Component Analysis module

Participants: François Cayre, Teddy Furon [contact person].

In the framework of watermark security analysis, we develop simulations tools to hack and disclose secret keys of Spread Spectrum based watermarking techniques. The core process is an independent component analysis of the watermarked signals. As we have not found any satisfying implementation of such a functionality, we develop our own C++ software based on the FastICA algorithm of A. Hyvarinen. This module has been submitted to IT++ sourceforge project, a well known C++ library of mathematical, signal processing, speech processing and communications classes and functions. It will be included in the next release of the library.

5.4. 3D Model-based video codec

Participants: Raphaèle Balter, Luce Morin [contact person].

>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

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

6.1.1. 3d Scene modelling from monocular video sequences

Participants: Raphaèle Balter, Luce Morin.

Video coding based on a 3D representation of the scene is a new and promising area of research. This type of representation turns out to be more compact than classical pixel-based representations. Over the past three years, we have developed a video coding approach based on a set of 3D models. 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 representation of the scene with a unique 3D model, as done in computer vision, is indeed not appropriate for coding purposes. The sequence of 3D models can be streamed for remote navigation in the scene. Although promising, the approach so far was not enforcing the extracted 3D information to be consistent along the entire sequence. Possible discontinuities between the different models were resulting in annoying artifacts in the reconstructed images.

In 2004, effort has been first dedicated to the design of solutions to handle this problem of discontinuities. The approaches designed rely on techniques of morphing and on so-called evolutive 3D models. The first approach designed rely on "a posteriori" 3D morphing over regularly meshed 3D models. A joint 2D

parameterization of the surfaces of pairs of adjacent models gives a geometric correspondence between the two models. A common connectivity mesh including all vertexes and faces of the two models is then created by 2D mesh fusion. Linear interpolation is then applied on the re-meshed 3D models. This scheme allows a smooth evolution of the geometry (shape) of the two models. However, it does not avoid ruptures in the models connectivity. In addition, the re-meshing done for each 3D model leads to significantly increased decoder complexity.
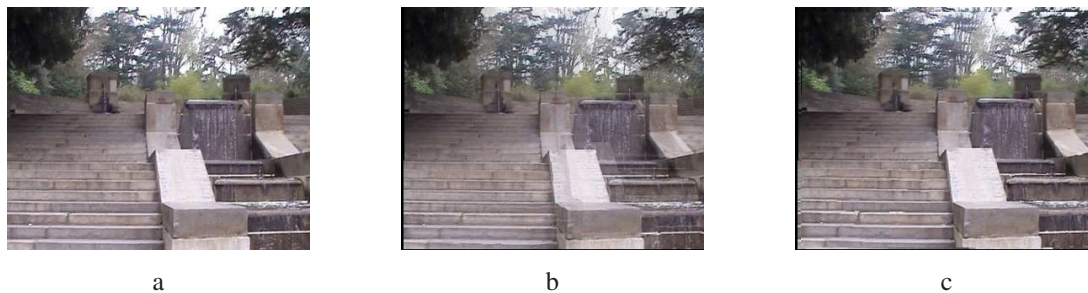

| a | b | c |

Figure 3. (a)-Image from the original sequence; Images reconstructed without 3D morphing (b) and with 3D morphing (c)

One can alternatively constrain the 3D models, when being extracted from the video, so that the visible subsets that are common to the two models have the same connectivity (i.e., same vertices and faces). The 3D models are first constructed independently by elevation from a uniform triangular mesh (i.e., from a depth map which has been meshed) on each key image. It is then tracked and updated to account for appearing areas, while preserving the existing connectivity. This common connectivity provides natural correspondences between the models. Hence, 3D morphing can then be performed using classical interpolation techniques. The set of resulting 3D models turns out to have good properties in terms of geometry, texture and connectivity continuity.

A scalable video coding scheme based on the 3D model representation of the scene described above has also been developed. The information on the geometry, connectivity, and texture of the 3D models is encoded and transmitted, as well as the camera position for each frame. Both the texture and the geometry of the models are encoded using a wavelet-based representation. The consistent connectivity of the models enables a consistent wavelet decomposition of the sequence of 3D models. This in turn allows for efficient and progressive coding and decoding of the models geometry.

6.1.2. *Video object segmentation and representation for compression purposes*
Participants: Marc Chaumont, Dubhe Chavira-Martinez, Nathalie Cammas, Henri Nicolas.

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. Two object-based video coding algorihms 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). 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. Motion is first estimated using

active meshes tracked over several frames as the support for the estimation. The use of hierarchical meshes allows for long term motion tracking and accurate motion estimation. It also allows precise registration of images, resulting in efficient texture and motion decorrelation. Using this information, texture and shape may be extracted and represented independently of the motion information. 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 scalability.

Video coding efficiency depends on the accuracy of the motion fields and on the way these motion fields are being exploited in the motion-compensated temporal transform. In occlusion areas, most motion estimation methods fail, due to spatial and temporal motion discontinuities. A new algorithm based on non-manifold motion has been developed for handling occlusions in motion estimation and representation. In an occlusion area, the mesh is constrained along the frontier between two objects having different displacements, so that different meshes can be constructed on both sides of the frontier. The corresponding triangles are thus overlapping. A non-manifold motion field is thus produced, enabling occluding objects to move independently from each other. The hierarchical mesh representation and estimation ensures the consistency of the motion field. The use of active meshes together with this approach for handling the frontier between objects (called *cracklines*) allows to improve the temporal prediction (see Fig. 4).

(a)   (b) (c)   (d)
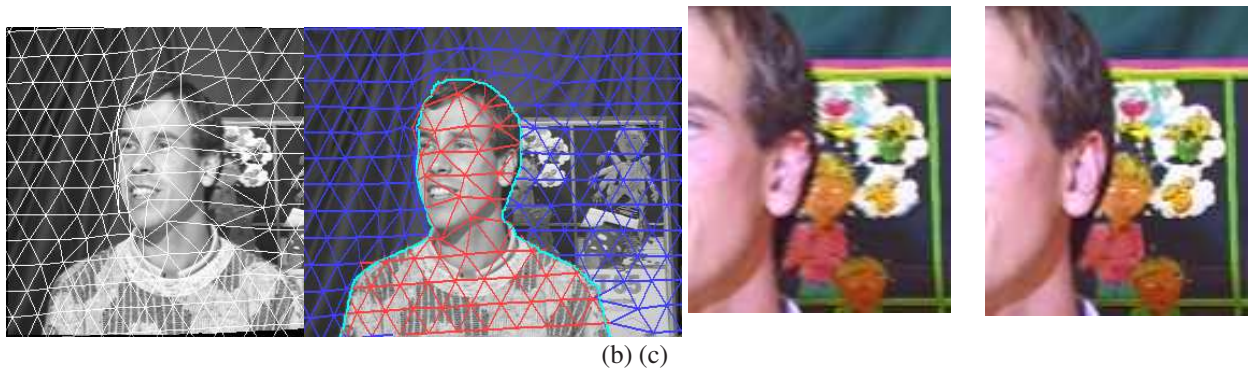
Figure 4. (a)- Reconstructed image with the corresponding estimated manifold mesh : the occlusion areas (right side of the character) generate elongated triangles, hence reconstruction artifacts (c). (b)- Reconstructed image with the corresponding estimated non-manifold mesh : the motion estimation done independently for the object and for the background improves the reconstruction in occlusion areas (d).

6.1.3. *Shadow and motion analysis for video compression*
Participants: 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.

In order to be able to correctly compensate the moving cast shadows, a realistic cast shadow model has been defined. This model takes into account the penumbra effect and the modifications of the ambient light. It has been incorporated in the joint cast shadow and light source position estimation previously developed. The shadow segmentation method has also been improved. It is now based on the minimization of an energy term using a clustering method. If the contours of the object which creates a shadow are known, the projection of the light source position on the image plane is determined.

Once the shadow contours have been determined, they are represented using 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. The moving cast shadows are removed from the original images. Two data streams corresponding respectively to the shadow information (contour and texture) and to the sequence without the shadows are then coded separately. The images without the shadows are coded using the scalable video coding scheme based on a 3-D subband decomposition developed by TEMICS. The approach can be beneficial for very low bitrate video surveillance systems where the shadows are useless information which can be only roughly coded or even not coded at all, and where only the moving objects represent relevant information.

6.2. Compression and distributed source coding

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

6.2.1. *Fine grain scalable video coding based on motion-compensated temporal filtering*

Participants: Vincent Bottreau, Christine Guillemot, Thomas Guionnet.

In 2002 and 2003, TEMICS has developed a scalable video coder/decoder, called WAVIX, based on motion-compensated spatio-temporal wavelet transforms. The algorithm and the software have been the starting point in the preparation of a joint Thomson/INRIA proposal to ISO in reply to the call for proposal initiating the specification of a scalable video coding standard (in MPEG-21/SVC). The objectives of the call were very challenging since the coder was supposed to generate a unique bitstream embedding resolutions and rates going from QCIF at 7.5 Hz and 48 Kbits/s to standard definition at 30Hz and a few Mbits/s, with a number of intermediate operation points in terms of spatial and temporal resolutions and in terms of bit rates. Spatial wavelet transforms used in the proposal as well as in competing solutions, designed for compression purposes only, were leading to aliasing artifacts in the embedded lower resolution signals. We have worked on the design and validation of a spatial transform based on three lifting steps. This transform has been the object of a joint proposal INRIA/UIC/Thomson to ISO. Another critical aspect in scalable video compression, especially at low rate and low spatial resolutions, is the bit rate needed to encode the motion fields which can potentially be high. TEMICS has worked on a scalable motion representation procedure.

6.2.2. *Oriented wavelet transforms for efficient contour representation in scalable image and video coding*

Participants: Vivien Chappelier, Christine Guillemot.

Wavelets are well-known mathematical tools for representing 1-D signals with a finite number of discontinuities with a small number of coefficients. However, for images modelled as homogeneous regions delimited by contours, curve discontinuities are not fully captured by separable wavelets. In image compression applications, high energy coefficients cluster around the edges and most of the bitrate is spent to code the contours. Thus, new transforms (e.g., curvelets and contourlets) have been designed to better take into account - and capture - geometrical patterns present in images. Curvelets and contourlets are implemented with filter banks with directional selectivity in the high frequencies, so that the resulting coefficients represent oriented portions of edges instead of points. Their main advantage is that they do not require a geometric model of the image. The counterpart is that discrete implementations of curvelet transforms are currently highly redundant, which limits their interest for compression applications. The bandlets follow a different approach as they use a geometric model to describe the discontinuities of the image (parametrized curves or regularity flows) and wrap wavelets along these discontinuities. Though theoretically more efficient than curvelets for compression purposes, this approach is computationally intensive and its main problem lies in the optimization of the bitrate allocation between the image geometry description and the wavelet coefficients.

A hybrid transform based on both contourlets and wavelets has first been designed. A projection technique (based on the POCS - projection on convex sets - technique) has also been coupled with the quantization applied in the redundant transform domain in order to minimize the distortion introduced by quantization. This modified contourlet transform gives better non-linear approximation results compared to wavelets when used for compressing images with directional features. This transform being redundant, we have then designed a new critically sampled transform based on wavelet lifting locally oriented according to multiresolution image geometry information. The orientation is restricted to a binary information per wavelet coefficient (horizontal/vertical for even levels, diagonal/antidiagonal for odd levels) so as to minimize the orientation map coding cost. In a first approach Markov random fields have been used to regularize the map and further reduce its coding cost (Fig. 5). However, this model is very dense, hence its coding cost remains too high. A quad-tree structure has then been used to describe the geometry of the image leading to an efficient representation and a simpler rate-distortion optimization (see Fig. 6). This transform exposes comparable energy compaction properties as bandlets for a significantly reduced complexity. For example, using stationnary entropy of the subbands as a measure of the bit rate, the oriented wavelet transform achieves a 1.25dB gain in PSNR at 0.3bpp for the reconstruction of the Barbara image compared to wavelets (Fig. 7). In order to assess the efficiency of this new transform in a real compression system, we are currently adapting a context-based arithmetic encoder with contexts adapted to the transform properties as well as considering other entropy coding techniques.

Figure 5. Maps of filtering orientations regularized with Markov random fields. The image on the left represents the horizontal (red) or vertical (green) filtering directions for the first decomposition level. The image on the right represents the diagonal '/' (blue) or antidiagonal '\' (yellow) filtering directions for the second decomposition level.

6.2.3. Distributed Source Coding

Participants: Christine Guillemot, Khaled Lajnef.

Distributed source coding (DSC) is a general framework which applies to highly correlated signals that are coded separately and decoded jointly. This framework applies to sensor networks but also to video compression. In the latter application, the motivation is to reduce the complexity of the encoder, at the expense of an increase of complexity of the decoder. From a theoretical point of view, DSC finds its foundations in the Slepian-Wolf theorem established in 1973. The Slepian-Wolf theorem states that for dependent binary sources Y and Z, the error decoding probability is close to zero for rates such that $R_Y \geq H(Y|Z), R_Z \geq H(Z|Y), R_Y + R_Z \geq H(Y, Z)$. This theorem has been extended to continuous-valued Gaussian sources by Wyner and Ziv in 1976. They have shown that for two correlated Gaussian sources Y and Z, if Z is available

Figure 6. Maps of filtering orientations obtained with quad trees. The image on the left represents the horizontal (red) or vertical (green) filtering directions for the first decomposition level. The image on the right represents the diagonal '/' (blue) or antidiagonal '\' (yellow) filtering directions for the second decomposition level.

Figure 7. Images coded/decoded with a separable wavelet transform (left, PSNR=27.25dB) and with an oriented wavelet transform (right, PSNR=28.50dB) at 0.3 bpp.

at the decoder, the rate-distortion performance obtained for the encoding of Y is the same whether the encoder knows the realization of Z or not. The results of Slepian and Wolf on one hand and of Wyner and Ziv on the other hand provide asymptotic bounds to lossless and lossy distributed coding of two correlated sources. They do not induce practical solutions for coder/decoder design.

Most practical DSC solutions are derived from channel coding. The statistical dependence between the two sources is modelled as a virtual correlation channel analogous to binary symmetric channels or additive white Gaussian noise (AWGN) channels. The source Y (called the side information) is thus regarded as a noisy version of X (called the main signal). Practical solutions based on channel codes like block and convolutional codes, turbo codes and LDPC have been designed mostly for the encoding of two sources.

Practical video compression schemes applying the DSC paradigm in a pixel or transform-domain (making use of classical de-correlating transform such as the DCT) have been reported. Key frames (in general every two frames) are coded using intraframe coding. The remaining frames (the Wyner-Ziv coded frames) are then coded separately but decoded conditionally to the side information which can be generated by interpolation of previously decoded frames. In other words, the Wyner-Ziv coded frames are intraframe coded and interframe decoded. One can then understand easily that the encoding is orders of magnitude less complex than the motion-compensated hybrid predictive coding. The first results show rate-distortion performance superior to that of intraframe coding, but there is still a gap with respect to conventional motion-compensated interframe coding.

In an attempt to reduce the performance gap with respect to conventional motion-compensated interframe coding, we have worked on the derivation of performance bounds and on the design of a DSC system with three correlated sources. The structure of dependencies considered and the resulting coding system can be regarded as the DSC counterpart of bidirectional or multiple reference predictive coding which brought a significant performance gain in classical video compression systems. We have extended the rate performances bounds for both binary and Gaussian correlation models. We have designed a coding/decoding system based on punctured turbo codes [53]. The performances with theoretical sources evidence the benefits in terms of compression efficiency. The next step is the validation in a real distributed video coding system.

6.3. Joint source-channel coding

6.3.1. *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 $l_2(Z)$. 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. With discrete frame expansions, the associated redundant 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. 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 [25].

The traditional BCH decoding or syndrome decoding approach is based on the concept of an error locator polynomial to localize the errors. However, 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. This leads to particular difficulties beyond traditional syndrome decoding techniques. Observing some analogy between the DOA estimation problem in array signal processing and the error localization with quantized discrete frame expansions, we have developed new decoding schemes based on subspace projection methods [30], [29]. The subspace based approaches to error localization developed are applicable to the discrete frame expansions characterized by the BCH structure. 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.

The above approaches apply to block transforms with a BCH structure, however cannot be easily applied to oversampled filter bank (OFB) based frame expansions. The problem of decoding OFB codes can be viewed as a problem of decoding real-number convolutional codes in presence of impulse noise errors and background noise. In addition, in contrast to finite-field convolutional codes, real numbered convolutional codes have infinite state-space size and therefore Viterbi algorithm can not be applied. We have thus developed an error localization procedure relying on hypothesis testing. The syndrome decoding algorithm consists of two steps: error localization and error correction. The error localization problem is treated as an M-ary hypothesis testing problem. The tests are derived from the joint probability density function of the syndromes under various hypothesis of impulse noise positions and in a number of consecutive windows of the received samples (to account for the encoding memory of the convolutional code). The error amplitudes are then estimated from the syndrome equations by solving them in the least square sense. The message signal is reconstructed from the corrected received signal by a pseudoinverse receiver. The performance of this algorithm has been first tested for a Bernoulli Gaussian impulse noise model. We have then considered an image compression system with a complete encoding chain consisting of an OFB based signal decomposition, scalar quantization and a variable length entropy code (VLC) or a fixed length code (FLC). The noise due to errors at the output of the VLC/FLC decoder (input of the OFB syndrome decoder) has been modelled as a Bernoulli-Gaussian or a quantizer-dependent impulse noise. The error localization procedure for the quantizer-dependent impulse noise model has been developed. We have further shown how the soft information at the output of the soft-input-soft-output (SISO) VLC/FLC decoder can be used in the localization procedure of the OFB syndrome decoder. The localization procedure has been further improved by introducing in the localization procedure of the OFB decoder per symbol reliability information produced by the SISO VLC/FLC decoder. The a posteriori probabilities of the source symbols produced by y the SISO VLC/FLC decoders are used in the calculation of the hypothesis a priori probabilities. The performance of these algorithms has been tested in the system with a tree structured subband decomposition by a wavelet filter bank and a Huffman or an FLC. The results show that introducing the soft information in the localization procedure of the syndrome decoding algorithm significantly improves the probability of detection and decreases the mean square error. We have further proposed an algorithm for the iterative decoding of the OFB-FLC chain. In this algorithm, the trellis for the decoding of the FLC encoded source coefficients modeled by the first order Markov source is iteratively pruned with the help of the hypothesis a posteriori probabilities. This is done based on the information on the symbols for which errors have been detected in the OFB syndrome decoder. The performance of these algorithms has been tested in the image compression system based on the subband decomposition by a wavelet filter bank.

6.3.2. *Joint Source-Channel Coding for MIMO-based Wireless Systems*

Participants: Robin Chatterjee, Christine Guillemot, Gagan Rath.

Theoretical studies show that the capacity of a wireless channel can be increased by using multiple antennas at both the transmitter and the receiver. Using such a multiple-input multiple-output (MIMO) antenna system with multiplexing of data however may not result in high reliability. One of the ways to improve the reliability is to use diversity in space and time where redundant data is transmitted from multiple antennas over time. The design of redundant data streams to be transmitted from different transmitting antennas over time is the subject of so-called space-time coding. Space-time coding as such is applied over the symbol stream where

the symbols are signal constellation points for a given modulation scheme. The design of a space-time code is primarily based on the minimization of the probability of transmitting a codeword and decoding a different codeword at the receiver.

A space-time block codeword typically consists of original symbols and their complex conjugates. Such a block of symbols can be generated by incorporating redundancy at the higher levels of the communication system as well. For example, one could use a multiple description coding where different descriptions of the source signal can be transmitted from different transmitting antennas which would promise similar or better performance as using a space-time code. This problem thus can be looked upon as a joint source-channel coding problem aimed for MIMO-based communication system. Consider a MIMO wireless communication system with n_t transmitting antennas and n_r receiving antennas. In order to achieve high reliability, the source signal needs to be separated into n_t redundant streams. Consider a source vector $\mathbf{x}$ consisting of K real-valued components. It can be expanded to Kn_t components using an overcomplete frame expansion as $\mathbf{y} = F\mathbf{x}$, where F is a frame operator associated with a frame having Kn_t frame vectors in the K-dimensional space. The components of $\mathbf{y}$ can be split into n_t vectors each having K components, and these vectors can be transmitted from n_t transmitting antennas after being quantized and modulated. In this case, frame expansion builds redundancy into the system which is exploited as spatial diversity. It is known that certain frame expansions provide error correcting capabilities as well. Therefore, such a frame expansion can increase the reconstruction reliability further by correcting residual errors in the decoded coefficients. This design however brings up several questions such as

- Is this system better than the system with a space-time code in the sense of lower bit error rate or lower reconstruction error for the same SNR?;
- Are the encoding and decoding algorithms less complex than those with a space-time code?;
- What are the criteria for finding the optimal frame expansion?;
- What are the characteristics of the frame operator which leads to the optimal performance?;
- Is the optimal frame expansion related to the optimal space-time code for a given bit rate, and if so, how?

To answer these questions, we need to investigate the proposed system both in theory and in simulation which is presently underway.

6.3.3. *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 [48]. 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. This class of codes has been extended so that higher-order source statistics are exploited [49]. The error-resilience, for the high priority source, is expressed analytically as a function of the indexes assigned to the different codewords. The formulation is in turn used as a criterion in index assignment algorithms based on the binary switching algorithm or making use of simulated annealing.

Strategies for error resilient and progressive transmission of classical VLCs (in particular of Huffman codes) have also been designed [48], [50]. These *bitstream construction* (BC) approaches allow to improve the error-resilience of the code, even when using simple hard decoding techniques. The performance can be further improved by using MAP, MPM or MMSE estimators. In contrast with solutions proposed so far in the litterature, the solutions designed have a linear complexity. The resulting bitstream structure is amenable to progressive decoding. The design of a progressive expectation-based decoding approach led to the introduction of code properties and design criteria for increased resilient and progressive decoding criteria. The VLC code has to be designed so that the symbol *energy* is mainly concentrated on the first bits of the symbol representation (i.e. on the first transitions of the corresponding codetree). This energy distribution criterion is used in the design of codes and in the corresponding index assignment.

Finally, another family of codes, called *self-multiplexed codes* and introduced in 2003, has been generalized. The generalization is based on the fact that a VLC based on a binary codetree can be seen as a set of re-writing rules of the form

$$a_i \rightarrow b_1..b_n,$$

where $a_i \in \mathcal{A}$ is a symbol of the source alphabet $\mathcal{A}$ and $(b_1..b_n) \in \{0,1\}^*$. Using this formalism, self-multiplexed codes can be regarded as the set of codes defined by a set of re-writing rules of the form

$$a_i b_1..b_m \rightarrow b_1..b_n,$$

with $m < n$. *Self-multiplexed* codes can be further generalized to the set of codes defined by a set of re-writing rules of the form

$$a_i s_1..s_m \rightarrow b_1..b_n,$$

where $(s_1..s_m) \in \{0,1\}^*$. This class of codes naturally extends codes based on binary codetrees and encompass codes which can be regarded as finite state automata including quasi-arithmetic codes. This extension introduces additional degrees of freedom in the index assignment, allowing to improve the decoder resynchronization properties and the code performance in a context of soft decoding.

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

Participants: Christine Guillemot, Thomas Guionnet, Marion Jeanne.

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 algorihms that would allow robust decoding of such codes, even in presence of bit errors. In the last two years, 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. In order to control the decoder complexity, different pruning techniques have been designed. The algorithm allows in addition to flexibly control the trade-off between the complexity and the reliability of the estimation. This parameter corresponds to the maximum number of surviving paths : a higher value of surviving paths corresponds to a higher decoding complexity and a higher estimation reliability.

A second decoding algorithm has been developed considering quasi-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 [17]. 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. In 2003, the approach has been first validated in a JPEG-2000 decoder. The integrated solution revealed a very significant gain with respect

to standard decoding solutions. The approach (patented) has been promoted within the JPEG-2000 standardization group and adopted within the part 11 of the standard dealing with JPEG wireless (JWL).

In 2004, effort has been dedicated to the validation of the solution in the context of context-based adaptive arithmetic coding used in the H.264 video coding standard and very likely to be used in the emerging MPEG-21 SVC (Scalable Video Coding) standard. The extra difficulty comes from the fact that both the encoder and the decoder learn the source statistics as long as the sequence of symbols is being encoded or decoded. A method for estimating the source HMM parameters from the noisy sequence of bits that is received in a robust manner needs to be designed. This is still under progress. The next step will be to promote the solution in the context of a core experiment on error resilience defined by the MPEG-21/SVC standardization group.

6.4. Image and video watermarking

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

In the framework of the french research network 'ACI sécurité', we have developed a cryptographic approach of the security of watermarking schemes (also called steganalysis). This analysis is based on the Kerckhoffs principle, Shannon's study of crypto-system, and Fisher's information measurement.

Basically, we estimate the amount of information about the secret key leaking from the observations made by the opponent. Although this is very classical in cryptanalysis, it has never been done in watermarking. For instance, cryptography deals with discrete variable, whereas watermarking usually plays with real signals. This is a typical problem because Shannon's equivocation or uncertainty of random discrete variables has no physical signification: it cannot be interpreted as an information measurement when applied on real signals. A different tool has to be used. We chose Fisher's information measurement.

The approach aims at assessing the number of watermarked contents that allows an accurate estimation of the secret signal. It is very well known in the theory of estimation that the Fisher Information Matrix yields a minimal bound of the mean square error whatever the estimator used by the opponent. This bound is a decreasing function of the number of observations. Such an approach is truly related to watermarking security since an opponent may actually access the watermarking communication channel hidden in host content. The disclosure of the secret allows the opponent to erase, modify, or embed watermarks.

Our work is also inspired by Diffie and Hellman cryptanalysis classification as the observations made by the opponent might be not only watermarked content. Depending on the application framework, we distinguish several attacks:

- Known Message Attack: the opponent knows which messages have been embedded into the watermarked content he observes,

- Known Original Attack: the opponent has access to pairs of original and watermarked contents,

- Watermark Only Attack: the opponent only observes watermarked contents.

Notice that in the Known Original Attack, the opponent doesn't need to hack the observed watermarked contents as he also has in hands the original versions. He first deduce some knowledge about the secret key from these pairs of content. Then, later on, he uses this knowledge to forge pirated content whose original version are not available.

Theoretical security levels have been estimated and assessed with experiments on a huge watermarked images database. The results obtained include the number of contents that have to be taken into account to gain an order of magnitude on the estimation of the secret signals, for every kind of attacks (KMA, KOA, WOA). Such a work has never been carried out in watermarking, although it is very common in classical cryptanalysis. Our theoretical results show that the vast majority of watermarking schemes (*ie*, techniques based on a spread spectrum modulation) are actually not secure: a relatively low number of contents available to an opponent may easily lead to disclosure of the secret signals.

Based on this analysis, we have implemented a security attack of an actual robust watermarking scheme for still images. Under the assumption of 1000 images available to the opponent, we have shown that in the

worst case (Watermarked Only Attack), an opponent may gain sufficient knowledge of the secret signals to perform a watermark removal attack at a low distortion: attacked images look almost perfect. However, not only watermark removal is achievable, since the opponent gains full access to the secret channel, he can also try to read/write the watermark (given additionnal high-level knowledge - for example watermark content structure), which was not possible with previous attacks.

The goal of this work is to warn the watermarking community that security is a crucial issue, so far underestimated. People are usually very concerned by the robustness of the watermark. Huge improvements have been achieved in this domain in the last few years. However, a robust watermarking technique may not be a secure primitive. This matter is extremely important when the watermarking technique is deployed on a large scale bank of contents.

(a) Attack by compression. Best quality for a successful attack: PSNR=21.8 dB.

(b) Steganalysis attack. Best quality for a successful hack: PSNR=35.8 dB.

Figure 8. Comparison between two strategies for attacking the Le Guelvouit and Pateux robust watermarking technique for still images. In the second strategy, the pirate has observed 300 watermarked pictures with different hidden messages to estimate the secret key.

6.4.1. *Interaction between watermarking and indexation*

Participants: François Tonnin, Sophie Le Delliou, Teddy Furon.

In collaboration with the TexMex project, we are studying the interactions between, and the mutual impact of indexing and watermarking. In the framework of the copyright protection of digital images, a system crawls the Internet and analyses the images found in suspicious websites. The system must recognize copyrighted pictures belonging to its database. In this case, an alarm signal is sent to the copyright holder who checks wether the website has cleared the corresponding copyright fees. This is done with a collaboration between indexing and watermarking as follows. Basically, the indexing process first finds the nearest picture in the large database from a suspicious image. If the distance score is weak, this constitutes a first element of proof. The indexing process also sends side information to the watermark decoder: the secret key and hidden message used at the embedding of the original picture, and also an estimated geometric distortion (angle of rotation, scale factor of a stretch...) between the suspicious and the would-be original images (*ie.* the nearest image found in the database). The detection of this message with this secret key is a second element of proof. This collaboration decreases the probability of false alarm while increasing the robustness of the watermarking detection test.

6.4.2. *Video watermarking*

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

The watermarking scheme developed for still images (see above subsection) is indeed one of the most robust techniques we are aware of. The work now is to transpose this technique to the video domain. The core process being unchanged, the challenge is to find a suitable domain of insertion in video.

A first idea is to insert the watermark signal in the wavelet domain, as we did for still images. The video is first packed in group of frame. A wavelet transform is applied along the temporal axis. Tests have shown that motion compensation, although usually done in video compression, is not robust enough. Then, a 2D wavelet transform is performed in the spatial domain. High scale coefficients are too sensitive to noise compression to reliably support some watermark information. The watermark signal is added to the remaining coefficients with a suitable energy allocation (depending on the perceptual impact and on an optimization of the robustness against the worse attack). The inverse transform completes the embedding. This video watermarking technique successfully embeds 64 bits in 5 seconds of a QCIF video (small size) encoded with MPEG4 at 64kbits/s.

A second track is currently explored in order to be robust against some geometric distortions such as a rotation of ± 5 degrees, a scaling factor from 0.5 to 2. Images are decomposed into small blocks whose dimension is proportional to the total size. This decomposition is thus invariant to a stretch. The energy of middle frequency coefficients is measured in each block. The watermark signal slightly changes this energy distribution along the video. A rotation of a small angle has a light impact on these descriptors when the number of blocks per image is low. Preliminary tests assess the robustness of the descriptors.

6.4.3. *Interaction between compression and indexation*

Participants: Christine Guillemot, 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.

7. Contracts and Grants with Industry

7.1. Industrial contracts

TEMICS has five 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 2004 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 2004 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 2004 has been on adaptive motion-compensated temporal filtering and on efficient and scalable coding of corresponding motion fields.

- CRE contract with France Telecom R&D (starting in October 2004) in the context of the Ph.D of Gaël Sourimant on the area of 3D reconstruction of urban scenes by fusion of GPS, GIS and video data.

- CRE contract with France Telecom R&D (starting in November 2004) on the problem of distributed source coding. The objective is to investigate this new coding paradigm and assess its potential for compression with mobile light-weight encoding systems.

TEMICS also supervises DRT projects in collaboration with industrial partners:

- H. Nicolas supervises the Degree of Technological Research ("Diplôme de Recherche Technologique" (DRT)) of Fabrice Templon realized at NEXTREAM on the subject "Development of graphic representations of MPEG-4 bitstreams and application to the optimization of the choice of the coding modes and parameters".

- H. Nicolas supervises the Degree of Technological Research ("Diplôme de Recherche Technologique" (DRT)) of Lila Huguenel realized at Thomson on the subject "MPEG-4 AVC video compression based on regions of interest".

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 : § 6.3.1 and § 6.3.4.
- 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 and by integrating these techniques in a fine grain scalable video coder and decoder. A procedure for scalable coding of mmotion vector fields has been developed. The encoding procedure is based on context-based arithmetic coder, state of the art entropy coding procedure. This entropy coder being sensitive to noise, the approach for robust decoding of arithmetic codes relying on Bayesian estimation tools has been integrated in the decoder. The resulting video codec is being described in section 5.1.2. TEMICS also studies redundant wavelet transforms for erasure-resilient coding and decoding of video signals.

7.2.2. RNRT-DIPHONET

Participants: Sophie Le Delliou, Teddy Furon.

- Convention number: 2 01 A 0704 00 000MC 01 1
- Title: Picture broadcasting over the Internet
- Research area associated: § 6.4.1, § 5.2.
- 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. The watermarking technique based on game theory reported in sub-section 6.4.1 has been evaluated and optimized against intentional attacks and potential de-synchronizations. A final review has successfully concluded the project in September 2004.

7.2.3. *RNTL DOMUS-VIDEUM*

Participant: Henri Nicolas.

- Convention number: 2 02 C 0100 00 00 MPR 01 1
- Title: *Indexation, advanced visualization 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. This approach has been tested on a corpus of documentary content and compared with two manual hyper-scenes representations based on semantic and color criteria respectively. The automatic clustering is globally close to the manual color-based hyper-scenes decomposition. The conclusion that we can derive from this comparison is that the proposed automatic method is able to provide with coherent hyper-scene decompositions.

7.2.4. *RNRT-EIRE*

Participant: 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
- 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 coded using a 3-D EBCOT coding method. In order to optimize the compression performance, an adaptive version of the coding scheme has been implemented. It allows to automatically valid or not the motion estimation, the use of a reference image corresponding to the background area (if the camera is fixed), and the selection of the inter GOF coding mode. Furthermore, it is possible to choose the subband coding modes and the GOF length according to the intensity of the temporal variations between the successive original images. This adaptativity allows to significantly improve the compression performance.

7.3. European contracts

7.3.1. IST-Busman

Participants: Vincent Bottreau, Jonathan Delhumeau, Teddy Furon.

- Convention number : 1 02 C 0186 00 00MPR 00 5.
- Title : Bringing user satisfaction to media access networks.
- Research axis : § 6.4.2.
- Partner : Bristish 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 second version of the software based on side information and game theory 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 final review is scheduled for end of November 2004.

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, Khaled Lajnef, Luce Morin, Henri Nicolas, Gael Sourimant.

- Convention number: 104C05310031324005
- Title: *European research taskforce creating human-machine interfaces SIMILAR to human-human communication.*
- Research axis: § 6.2.3, § 6.1.1
- Partners: around 40 partners from 16 countries.
- Funding: CEE.
- Period: Jan.04-Dec.07.

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

- In the context of 3D modelling of video sequences we have focused on an hybrid representation mixing 2D and 3D representations of video data. Cylindrical and spherical mosaics are used for unified coding and visualization of 2D and 3D data. Such an approach allows to make no assumption on the camera acquisition path, and still provides the benefits of 3D functionalities for virtual reality applications.

- Shadow and light source detection and analysis techniques have been developed. They are used to artificially create cast shadows of natural objects inserted in a video sequence. These methods will be applied to the mixing of 2-D (original and mosaic images, video objects) and 3-D (synthetic objects and 3-D model) video data.

- TEMICS is contributing on a distributed coding framework in a context of multimodality and coordinating with EPFL a working group dealing with the development of an information theoretic for analysis and representation of multimodalities.

7.3.4. FP6-IST STREP DANAE

Participants: Vincent Bottreau, Christine Guillemot, Hervé Jégou, Gagan Rath.

- Convention number: 104C045731324005
- Title: *Dynamic and distributed Adaptation of scalable multimedia content in a context-Aware Environment.*
- Research axis: § 6.2.1, § 6.3.3
- Partners: ENST, France Télécom, Imperial College London (ICL), Inria, Museon, Siemens, T-systems, University of Aachen, University of Geneva, University of Klagenfurt.
- Funding: CEE.
- Period: Jan.04-June.06.

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. TEMICS contributes on the aspects of fine grain scalable video coding and on the study of new source codes for increasing the error resiliency of the scalable video coder while preserving its compression and scalable properties. In collaboration with other DANAE partners, TEMICS contributes to different core experiments defined in the context of MPEG-21/SVC: a core experiment on spatial transforms, on error resilience and on coding with multi-rate adaptability.

8. Other Grants and Activities

8.1. National initiatives

8.1.1. ARC Télégéo
Participants: Raphaèle Balter, Luce Morin.

- Title : Télégéo: Geometry and Telecommunications.
- Research axis : § 6.1.1.
- 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.

8.1.2. ACI Fabriano
Participants: François Cayre, Teddy Furon.

- Title : Fabriano
- 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 to 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. TEMICS proposed a theoretical framework for security level assessment of watermarking technique. It has been applied to substitution and spread spectrum based schemes.

8.2. Bi-lateral international co-operation

8.2.1. Project reviewing

- C. Guillemot has been appointed by the European commission to review project submissions within the IST programme of the 6th framework programme.
- H. Nicolas has been appointed by the Rhônes-Alpes region to evaluate a research project.
- H. Nicolas expertised one company project for ANVAR.

9. Dissemination

9.1. Standardization

- R. Balter, P. Gioia and L. Morin, "Preliminary test and evidences for EE on 3D model based movie streaming" MPEG/3DAV, ISO/JTC1/SC29/WG11 M10707, Munich, Germany, March 2004.
- J.Vieron, E. François, V. Bottreau, C.Guillemot, G.Marquant, G.Boisson, "Fully scalable video coding based on 2D+t wavelet technology ", common contribution Thomson/INRIA to MPEG/SVC, ISO/IEC JTC1/SC29/WG M10569, March 2004.
- V. Bottreau, C. Guillemot, R. Ansari, E. François, " Spatial transform using three lifting steps ", common contribution Thomson/INRIA to MPEG/SVC, ISO/IEC JTC1/SC29/WG M10904, July 2004.
- E. François, G. Boisson, J. Viéron, C. Guillemot, V. Bottreau, " Evaluation of motion accuracy scalability ", common contribution Thomson/INRIA à MPEG/SVC, ISO/IEC JTC1/SC29/WG M10902, July 2004.

9.2. Leadership within the scientific community

- T. Furon is member of the program committees of Int. Workshop on Digital Watermarking 2004;
- T. Furon partly organized the special issue on multimedia content protection of the 'Matinales of Rennes Atalantes';
- C. Guillemot is associate editor of the journal IEEE Transactions on Circuit and System for Video Technology;
- C. Guillemot is elected member of the international committee IEEE IMDSP (Image and MultiDimensional Signal Processing Technical Committee);
- C. Guillemot is member of the external scientific advisory board of the IST-FP6 Network of Excellence VISNET;
- C. Guillemot has been nominated by the ministry of research as a french representative within the management committee of COST (Action COST 292 "Sematic and multimodal analysis of digital media" ;
- C. Guillemot is member of program committees of the following conferences: IEEE-ICIP 2004, WIAMIS 2004, IWSSIP'04, ACM Multimedia 2004, IEE-VIE 2004, CORESA 2004, PCS 2004;
- C. Guillemot is member of the steering committee of the FP6-IST Network of Excellence SIMILAR;
- 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 IEEE-ICIP 2004 and ISIVC 2004;
- H. Nicolas is member of the "commission de spécialistes" of the University of Rennes 1.

9.3. Invitations

- T. Furon made a talk at YACC'04 (Yet Another Conference in Cryptography);
- C. Guillemot gave an invited seminar at the Technical University of Munich;
- C. Guillemot gave an invited seminar at the Nanyang Technological University, Singapore;
- L. Morin will give an invited talk in PCS'2004 Special Session on *Convergence of Computer Vision and Visual Communication*", in San Francisco, USA.
- G. Rath has spent three months as an invited researcher at the National Technical University of Singapour.

9.4. Teaching

- Master of Multimedia Network Security, Telecom Paris (F. Cayre: Watermarking techniques for still images);
- Master of Multimedia Network Security, Telecom Paris (F. Cayre: Watermarking attacks: robustness and security);
- Diic-INC, Ifsic, university of Rennes 1 (L. Morin, H. Nicolas, C. Guillemot : image processing, 3D vision, motion, coding, compression, cryptography, communication) ;
- Diic-LSI, Ifsic, university of Rennes 1 (H. Nicolas : compression) ;
- MASTER Master research STI, university of Rennes 1 (C. Labit, H. Nicolas : compression) ;
- ESIGETEL Fontainebleau, (H. Nicolas : Video compression and communication) ;
- Enic, Villeneuve-d'Ascq, ENSTBR (C. Guillemot: Video communication) ;
- Ensar Rennes (L. Morin : Basics of image processing, and mathematical morphology) ;
- 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*, chap. Algorithmes de codage et de contrôle de transmission sur réseaux hétérogènes, Éditions Hermès, 2000, p. 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", vol. 19, nº 9, sept. 2001, p. 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", vol. 6, nº 6, June 1997, p. 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, p. 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", vol. 6, nº 4, December 1995, p. 303-316.

[6] G. RATH, C. GUILLEMOT. *Performance analysis and recursive syndrome decoding of DFT codes for bursty erasure recovery*, in "IEEE Transactions on Signal Processing", vol. 51, nº 5, May 2003, p. 1335–1350.

[7] P. RAULT, C. GUILLEMOT. *Symmetric delay factorization: generalized framework for paraunitary filter banks*, in "IEEE Trans. on signal processing", vol. 47, nº 12, Dec. 1999, p. 3315-3325.

Doctoral dissertations and Habilitation theses

[8] N. CAMMAS. *Codage vidéo scalable par maillages et ondelettes t+2D*, Ph. D. Thesis, Thèse de doctorat en Informatique, Université de Rennes 1, France, November 2004, http://www.irisa.fr/bibli/publi/theses/2004/cammas/cammas.html.

[9] D. CHAVIRA-MARTINEZ. *Schéma de codage orienté-texture: une approche robuste et scalable pour la diffusion sur réseaux*, Ph. D. Thesis, Thèse de doctorat en Informatique, Université de Rennes 1, France, July 2004.

[10] M. GARCIA-VAZQUEZ. *Analyse des mouvements non-plans des objets vidéo: Application à la manipulation et à la compression basée objet des séquences vidéo*, Ph. D. Thesis, Thèse de doctorat en Informatique, Université de Rennes 1, France, June 2004.

Articles in referred journals and book chapters

[11] J. P. ANDREAUX, A. DURAND, T. FURON, E. DIEHL. *SmartRight: A Copy Protection System for Digital Home Networks*, in "IEEE Signal Processing Magazine", Special Issue on Digital Right Management, vol. 21, n° 2, March 2004, p. 100-108.

[12] R. ANSARI, C. GUILLEMOT, N. MEMON. *The JPEG lossy standard*, Chapter in the book The Handbook in Image and Video Processing, Eds Prentice-Hall, 2004.

[13] F. CAYRE, C. FONTAINE, T. FURON. *Watermarking security: Theory and Practice*, in "accepted in IEEE Trans. on Signal Processing", Special Supplement on Secure Media, oct 2004.

[14] T. FURON. *Tatouage asymétrique*, F. DAVOINE, S. PATEUX (editors)., Traité IC2 Information-Commande-Communication, série Traitement du Signal et de l'Image, Hermes Science, March 2004, p. 215–227.

[15] C. GUILLEMOT, P. CHRIST. *Joint source-channel coding as an element of a QoS framework for 4G wireless multimedia*, in "Elsevier Journal on Computer Communications, special issue on research directions for 4G wireless networks", vol. 27, n° 8, May 2004, p. 762-779.

[16] C. GUILLEMOT, P. SIOHAN. *Joint source-channel decoding with soft information: a survey*, in "accepted for publication in Elsevier Journal on Applied Signal Processing, special issue on the turbo principle", 2004.

[17] T. GUIONNET, C. GUILLEMOT. *Soft and joint source-channel decoding of quasi-arithmetic codes*, in "Eurasip Journal on Applied Signal Processing", n° 3, March 2004, p. 393-411.

[18] H. JÉGOU, C. GUILLEMOT. *Error-resilient first-order multiplexed codes : performance bounds, design and decoding algorithms*, in "accepted for publication in IEEE Trans. on Signal Processing", 2004.

[19] H. JÉGOU, C. GUILLEMOT. *Robust multiplexed codes for compression of heterogeneous data*, in "accepted for publication in IEEE Trans. on Information Theory", 2004.

[20] T. MAILLARD, T. FURON. *Towards Digital Rights and Exemptions Management Systems*, in "Computer Law and Security Report", vol. 20, n° 4, July 2004, p. 281-287.

[21] S. MARINKOVIC, C. GUILLEMOT. *Joint Source-channel Coding Based on Cosine Modulated Filter Banks for Erasure Resilient Signal Transmission*, in "accepted for publication in Eurasip journal on applied signal processing", 2004.

[22] N. MEMON, C. GUILLEMOT, R. ANSARI. *The JPEG lossless standard*, Chapter in the book The Handbook in Image and Video Processing, Eds Prentice-Hall, 2004.

[23] R. MOTWANI, C. GUILLEMOT. *Tree-structured oversampled filter banks as joint source-channel codes: Application to image transmission over erasure channels*, in "IEEE Transactions on Signal Processing", vol. 52, nº 9, Sept. 2004, p. 2584-2699.

[24] H. NICOLAS. *Shadow synthesis for video post-production applications*, in "To be published in IEEE Signal Processing Letters", 2004.

[25] G. RATH, C. GUILLEMOT. *Characterization of a class of error correcting frames for robust signal transmission over wireless communication channels*, in "accepted for publication in Eurasip journal on applied signal processing, special issue on design of cross-layer communication systems", 2004.

[26] G. RATH, C. GUILLEMOT. *Frame-theoretic analysis of DFT codes with erasures*, in "IEEE Transactions on Signal Processing", vol. 52, nº 2, Feb. 2004, p. 447-460.

[27] G. RATH, C. GUILLEMOT. *Recent advances on DFT codes based quantized frame expansions for erasure channels*, in "Elsevier journal on Digital Signal Processing", vol. 14, nº 4, July 2004, p. 332-354.

[28] G. RATH, C. GUILLEMOT. *Recent advances on DFT codes based quantized frame expansions for erasure channels*, in "Elsevier journal on Digital Signal Processing", vol. 14, nº 4, July 2004, p. 332-354.

[29] G. RATH, C. GUILLEMOT. *Subspace algorithms for error localization with quantized DFT codes*, in "To appear in IEEE trans. on communications", Dec. 2004.

[30] G. RATH, C. GUILLEMOT. *Subspace-Based Error and Erasure Correction with DFT Codes for Wireless Channels*, in "IEEE Trans on Signal Processing", vol. 52, nº 11, Nov 2004.

[31] J. VIÉRON, C. GUILLEMOT. *Real-time constrained TCP-compatible rate control for video over the Internet*, in "IEEE Transactions on Multimedia", vol. 6, nº 4, August 2004, p. 634-646.

[32] 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 "Eurasip Journal on Applied Signal Processing, special issue on Multimedia over wireless IP", nº 2, Feb. 2004, p. 158-176.

Publications in Conferences and Workshops

[33] A. MANOURY AND H. NICOLAS. *Détection et modélisation des ombres en mouvement et application à la compression vidéo*, in "Proceedings of WIAMIS, Lisboa, Portugal", April 2004.

[34] R. BALTER, P. GIOIA, L. MORIN. *Codage vidéo très bas débit basé sur des modèles 3D*, in "isivc'04, Second International Symposium on Image/Video Communications over fixed and mobile networks", june 2004.

[35] R. BALTER, P. GIOIA, L. MORIN, F. GALPIN. *Scalable and efficient coding of 3D model extracted from a video*, in "3DPTV'04, Second International Symposium on 3D Data Processing, Visualization, Thessaloniki, Greece", September 2004.

[36] G. BOISSON, E. FRANÇOIS, C. GUILLEMOT. *Accuracy-scalable motion coding for efficient scalable video compression*, in "IEEE International Conference on Image Processing, ICIP'04", Oct. 2004.

[37] G. BOISSON, E. FRANÇOIS, C. GUILLEMOT. *Efficient scalable motion coding for wide range scalable video compression*, in "European conference on signal processing, EUSIPCO'04", Sept. 2004.

[38] F. CAYRE, C. FONTAINE, T. FURON. *Watermarking Attack: Security of WSS Techniques*, in "Proc. of Int. Workshop on Digital Watermarking, Seoul, Corea", Springer-Verlag, IWDW'04, October 2004.

[39] V. CHAPPELIER, C. GUILLEMOT, S. MARINKOVIC. *Codage d'images par ondelettes unidimensionnelles orientées*, in "CORESA", May 2004, p. 117–120.

[40] V. CHAPPELIER, C. GUILLEMOT, S. MARINKOVIC. *Image coding with iterated contourlet and wavelet transforms*, in "IEEE International Conference on Image Processing, ICIP'04", Oct. 2004.

[41] F. GALPIN, R. BALTER, L. MORIN, K. DEGUCHI. *3D Models Coding and Morphing for Efficient Video Compression*, in "Proceedings of the Conference on Computer Vision and Pattern Recognition, CVPR'2004, Washington DC, USA", June 2004, p. 334–341.

[42] F. GALPIN, R. BALTER, L. MORIN, S. PATEUX. *Efficient and scalable video compression by automatic 3D model building using computer vision*, in "Picture Coding Symposium, PCS'2004, San Francisco, USA", Dec. 2004.

[43] B. L. GUEN, R. BALTER, L. MORIN. *Morphing de modèles 3D éstimés*, in "CORESA 2004, Compression et Représentation de Signaux Audio-visuels", May 2004.

[44] T. GUIONNET, C. GUILLEMOT. *Joint source-channel decoding of quasi-arithmetic codes*, in "Proc. IEEE Data Compression Conference, DCC", March 2004.

[45] H. NICOLAS. *Détection et modélisation des ombres en mouvement et application à la compression vidéo*, in "Proceedings of ISIVC, Brest, France", July 2004.

[46] H. NICOLAS. *Moving Shadow Representation Based on a Level Curves Decomposition*, in "Proceedings of ICIP, Singapore", Oct. 2004.

[47] H. NICOLAS, A. MANOURY, J. BENOIS-PINEAU, W. DUPUY AND D. BARBA. *Grouping Video Shots Into Scenes Based on 1D Mosaic Descriptors*, in "Proceedings of ICIP, Singapore", Oct. 2004.

[48] H. JÉGOU, C. GUILLEMOT. *Error-resilient and progressive transmission and decoding of compressed images*, in "Proc. WIAMIS'2004", Lisboa, Apr. 2004.

[49] H. JÉGOU, C. GUILLEMOT. *First-order multiplexed source codes for error-resilient entropy coding*, in "Proc. ISIT'2004", Chicago, June 2004.

[50] H. JÉGOU, C. GUILLEMOT. *Suffix-constrained codes for progressive and robust data compression: self-multiplexed codes*, in "Proc. EUSIPCO'2004", Vienna, Sept. 2004.

[51] H. JÉGOU, C. GUILLEMOT. *Variable length codes over noisy channel: New transmission and decoding strategies*, in "Picture Coding Symposium, PCS 2004", Dec. 2004.

[52] K. LAJNEF, C. GUILLEMOT, P. SIOHAN. *Codage de sources distribuées : analyse de performances et extension à trois sources*, in "CORESA", May 2004, p. 219–222.

[53] K. LAJNEFF, C. GUILLEMOT, P. SIOHAN. *Distributed Source Coding of Three sources using Punctured Turbo Codes*, in "IEEE Intl. workshop on Multimedia Signal Processing", Sept. 2004.

[54] M. CHAUMONT, S. PATEUX AND H. NICOLAS. *OBJECT-BASED VIDEO CODING USING DYNAMIC CODING APPROACHS*, in "Proceedings of ICIP, Singapore", Oct. 2004.

[55] S. MARINKOVIC, C. GUILLEMOT. *Erasure Resilience of Oversampled Filter Bank Codes Based on Cosine Modulated Filter Banks*, in "EEE Intl. conference on communication, ICC'04", May 2004.

[56] S. MARINKOVIC, C. GUILLEMOT. *Impulse noise correction in an image transmission system by means of an oversampled filter bank code*, in "European conference on signal processing, EUSIPCO 04", Sept. 2004.

[57] S. MARINKOVIC, C. GUILLEMOT. *Joint Source and Channel Coding/Decoding by Concatenating an Oversampled Filter Bank Code and a Redundant Entropy Code*, in "IEEE GLOBECOM conference", Dec. 2004.

[58] E. MORILLON, R. BALTER, L. MORIN. *2D/3D hybrid modeling for video sequence*, in "Wiamis 2004,International Workshop on Image Analysis for Multimedia Interavtive Services,april 2004", 2004.

[59] E. MORILLON, R. BALTER, L. MORIN. *Modélisation hybride 2D/3D de séquences vidéos*, in "CORESA 2004, Compression et Représentation de Signaux Audio-visuels", May 2004.

[60] R. MOTWANI, C. GUILLEMOT. *A para-pseudo inverse based method for reconstruction of filter bank frame-expanded signals from erasures*, in "IEEE International Conference on Image Processing, ICIP'04", Oct. 2004.