

*Projet TEMICS**Traitement, Modélisation d'Images et
CommunicationS**Rennes*

THÈME 3B

 *R* **apport
d'Activité**

2002

Table des matières

1. Composition de l'équipe	1
2. Présentation et objectifs généraux	2
3. Fondements scientifiques	3
3.1. Modélisation 3d par géométrie projective	3
3.2. Modélisation du mouvement par maillages actifs	4
3.3. Théorie débit-distorsion et formalisme MDL	5
3.4. Représentation non redondante des signaux de source	6
3.5. Représentation des signaux par bases de fonctions redondantes	7
3.6. Tatouage d'images	8
4. Domaines d'application	9
4.1. Panorama	9
4.2. Compression d'images fixes et animées avec fonctionnalités avancées	10
4.3. Communication multimédia	10
4.4. Protection, authentification et enrichissement de contenus images et vidéo	11
5. Logiciels	11
5.1. Plate-forme de communication vidéo	11
5.2. Codeur vidéo scalable	12
5.3. Tatouage robuste d'image	12
5.4. Codeur-décodeur vidéo basé sur une représentation par modèles 3d	12
6. Résultats nouveaux	12
6.1. Analyse et modélisation de séquences d'images	12
6.1.1. Modélisation 3d par analyse de séquences vidéo monoculaires	13
6.1.2. Segmentation d'objet vidéo pour le codage	13
6.1.3. Codage vidéo basé sur une représentation par maillages et mosaïques généralisées d'objets vidéo	15
6.1.4. Analyse des conditions d'illumination dans une séquence d'images	15
6.1.5. Structuration par analyse des trajectoires des objets vidéo	16
6.2. Codage et codage conjoint source-canal	16
6.2.1. Contrôle de congestion multipoint et régulation de débit	17
6.2.2. Codage scalable à grain fin	17
6.2.3. Transformations redondantes pour le codage conjoint source-canal	17
6.2.4. Codes sources robustes au bruit de transmission	18
6.2.5. Décodage conjoint source-canal de codes d'Huffman	19
6.2.6. Décodage conjoint source-canal de codes arithmétiques	20
6.3. Tatouage d'images	20
6.3.1. Schéma d'insertion et d'extraction de marques	22
6.3.2. Tatouage avec information adjacente et sécurité	22
7. Contrats industriels	22
7.1. Contrats industriels bi-latéraux	22
7.1.1. Convention Celar	22
7.2. Contrats nationaux	23
7.2.1. RNRT-V2net	23
7.2.2. RNRT-Cohrainte	23
7.2.3. RNRT-VIP	23
7.2.4. RNRT-DIPHONET	24
7.2.5. Projet RNTL DOMUS-VIDEUM	24
7.2.6. RNRT-EIRE	24

7.2.7.	RNRT-Cosocati	25
7.3.	Contrats européens	25
7.3.1.	IST-Busman	25
7.3.2.	IST-Song	26
7.3.3.	IST-Ozone	26
8.	Actions régionales, nationales et internationales	26
8.1.	Réseaux et groupes de travail internationaux bilatéraux	26
8.2.	Actions nationales	27
8.2.1.	ARC Télégeo	27
8.2.2.	Action CNRS tatouage	27
8.3.	Visites et invitations de chercheurs	28
9.	Diffusion des résultats	28
9.1.	Dépôt de brevets	28
9.2.	Animation de la communauté scientifique	28
9.3.	Enseignement	28
10.	Bibliographie	29

1. Composition de l'équipe

Responsable scientifique

Christine Guillemot [DR Inria]

Assistante de projet

Edith Blin-Guyot [TR Inria, jusqu'au 30 septembre 2002]

Stéphanie Lemaile [TR Inria, à partir du 1er octobre 2002]

Personnel Inria

Teddy Furon [CR, à partir du 1er Octobre 2002]

Claude Labit [DR, 1/5 temps depuis mars 1999]

Slavica Marinkovic [CR, à partir du 1er Octobre 2002]

Henri Nicolas [CR]

Stéphane Pateux [CR, détachement du corps des Télécoms]

Pierre Siohan [DR, détachement de France Télécom]

Personnel CNRS

Laurent Guillo [IR, (Atelier)]

Personnel Université

Luce Morin [maître de conférences, université Rennes 1]

Ingénieurs experts

Gabriele Dawini [IE Inria, jusqu'au 1er décembre 2002]

Bilal El Ali [IE Inria, jusqu'au 1er mai 2002]

Anne Manoury [IE Inria, à partir du 1er octobre 2002]

Chercheurs post-doctorants

Teddy Furon [Post-doctorat UCL jusqu'au 30 septembre 2002]

Slavica Marinkovic [Post-doctorat INRIA jusqu'au 30 septembre 2002]

Gagan Rath [Post-doctorat UNIVERSITÉ RENNES-1]

Claudio Weidmann [Post-doctorat INRIA]

Ingénieurs associés

Philippe Coval [jusqu'au 30 Septembre 2002]

Brendan Catherine

Chercheurs doctorants

Raphaële Balter [Bourse CIFRE France Télécom]

Fabien Catteau [Bourse MENRT, à partir du 1er octobre 2002]

Nathalie Camas [Bourse CIFRE France Télécom]

Vivien Chappelier [Bourse MENRT, à partir du 1er octobre 2002]

Dubhe Chavira-Martinez [Bourse CONACYT]

Marc Chaumont [Bourse MENRT]

Jonathan Delhumeau [Bourse INRIA, à partir du 1er octobre 2002]

Mireya Garcia-Vasquez [Bourse CONACYT]

Thomas Guionnet [Bourse INRIA]

Reda Hosny [Bourse MAE, jusqu'au 31 août 2002]

Hervé Jégou [Allocation ENS, à partir du 1er octobre 2002]

Guillaume Larignon [Bourse INRIA, jusqu'au 30 septembre 2002]

Gaetan Le Guelvout [Bourse INRIA-RÉGION]

Jean-Marie Pinel [Bourse MENRT]

Jérôme Viéron [Bourse INRIA]

2. Présentation et objectifs généraux

Les objectifs du projet sont de développer les concepts et les outils d'analyse, de modélisation, de codage, et de tatouage d'images, et plus généralement des informations vidéo manipulées en communication multimédia. Nos travaux portent plus particulièrement sur les problèmes suivants :

- l'interaction avec le contenu et la navigation dans des scènes vidéo 3D ;
- la représentation compacte et robuste aux bruits de transmission des images et des signaux vidéo ;
- le marquage (ou tatouage) des images et des signaux vidéo à des fins de protection contre les copies illicites, et à des fins d'authentification.

Pour répondre à ces objectifs dans un contexte en évolution forte et permanente, nos activités de recherche sont structurées en trois axes principaux :

- *Analyse et modélisation de séquences vidéo.* Dès que des fonctionnalités d'interaction avec le contenu vidéo ou de codage basé objet sont envisagées, une phase préalable d'analyse de la séquence est nécessaire. Cette phase consiste à segmenter les objets et à en extraire les principaux attributs (mouvement, texture, forme, orientation, illumination, ...). Les fonctions de navigation dans une scène vidéo nécessitent également la construction d'un modèle 3D de la scène. Les représentations des modèles et des attributs extraits doivent être faciles à manipuler. Lorsque des applications de communication sont envisagées, les représentations de ces modèles doivent en outre être compactes et robustes au bruit de transmission, rejoignant en cela le thème du codage conjoint source-canal. Pour une adaptation aisée à des contraintes variables liées au canal, ou pour une interaction fluide avec le contenu, il est en outre souhaitable que les modèles spécifiés permettent une description hiérarchique.
- *Codage conjoint source-canal.* Les environnements de communication hétérogènes (et à qualité de service non garantie) posent de nouveaux problèmes de représentation et de compression du signal image et vidéo. Les représentations codées d'une scène doivent être compactes et robustes au bruit de transmission. Le bruit de transmission peut se traduire par des pertes de paquets ou par des corruptions du train binaire. Minimiser l'impact du bruit sur la qualité du signal reconstruit, et optimiser la qualité globale de la chaîne de communication nécessitent d'abandonner les principes de séparabilité de codage de source et de codage de canal, pour se tourner vers des approches de codage conjoint source-canal. Nous nous intéressons aux techniques de codage et de décodage robuste par transformations et quantification redondantes, à la conception de codes sources robustes et performants en compression et à la conception de codes joints source-canal. Nos travaux portent également sur les techniques de décodage conjoint source-canal mettant en œuvre le principe turbo (inspiré des turbo-codes). Enfin, nous étudions le problème de l'allocation de débit optimale entre le codage de source et le codage de canal, et développons des techniques de codage scalable pour une adaptation aisée du débit de la source vidéo à la bande passante du réseau.
- *Tatouage.* L'apparition de réseaux ouverts comme l'Internet suscite de nouveaux dangers pour la propriété intellectuelle. Grâce - ou à cause - des technologies numériques, il est possible de faire des copies parfaites des contenus, la redistribution est devenue aisée et le traçage des redistributions difficile. L'objectif général du tatouage (watermarking en anglais) est de dissimuler des informations dans un contenu multimédia de manière invisible et robuste. De nombreuses applications sont envisageables parmi lesquelles celles visant à protéger les oeuvres multimédia : identifier leur origine, protéger les droits d'auteurs, ou bien encore vérifier leur intégrité. Cette marque doit être alors indécélable, infalsifiable et ineffaçable. Son extraction doit être robuste face à divers types d'attaques que le signal support (image, vidéo) pourrait subir (compression, retaille, filtrage, ...). Le tatouage peut aussi être utilisé pour enrichir un signal par insertion de métadonnées. Le problème du tatouage peut se formaliser comme un problème de communication

d'information dans un environnement bruité. Les travaux que nous menons s'appuient sur les bases scientifiques de l'équipe dans le domaine du codage conjoint source-canal et de l'analyse de séquences vidéo.

Compte-tenu de l'impact fort de la normalisation dans le domaine des applications multimédia communicantes, nous nous efforçons, dans nos orientations, de trouver un équilibre entre des approches compatibles avec les syntaxes normalisées et des approches se situant en rupture avec celles-ci. Le projet cherche également, en partenariat avec des industriels, à valoriser les résultats de ses travaux au travers de contributions à la normalisation (IETF, MPEG).

Les domaines d'application de nos travaux sont les applications multimédia en réseaux (Internet et réseaux sans fils) au travers de leurs multiples besoins en terme de compression, de transmission robuste ou en terme de fonctionnalités avancées telles que navigation, protection et authentification. Dans ces secteurs d'applications, nous avons des conventions de recherche ou de transfert avec France Télécom, le Celar, des projets multi-partenaires nationaux (projets RNRT COHRAINTE, COSOCATI, DIPHONET, EIRE, V2NET, VIP) et européens (projets IST-BUSMAN, IST-SONG et IST-OZONE).

3. Fondements scientifiques

3.1. Modélisation 3d par géométrie projective

Mots clés : *reconstruction 3D, vision par ordinateur, géométrie projective, projection perspective, modèle de caméra, équations de projection, auto-calibration, coordonnées homogènes, espace projectif, matrice fondamentale, matrice essentielle, contrainte épipolaire.*

La modélisation (ou reconstruction) 3D désigne l'estimation de la forme et de la position d'objets 3D à partir d'un ensemble d'images de ces objets. TEMICS s'intéresse plus particulièrement au problème de la modélisation de grandes scènes à partir de séquences vidéo monoculaires. De par sa nature, la modélisation 3D par géométrie projective est un problème inverse. Nous nous concentrons sur l'un des problèmes non résolus de la reconstruction 3D par géométrie projective, à savoir sur l'estimation des paramètres 3D de la caméra, rendue particulièrement délicate lorsque cette estimation doit se faire à partir de nombreux points de vue distants fournis par une caméra mobile. Nous nous intéressons également à la cohérence des observations entre les points de vue, à la gestion des parties cachées et des discontinuités de profondeur.

Les outils méthodologiques utilisés sont issus du domaine de la vision par ordinateur et de la géométrie projective. Les *équations de projection* pour une caméra \mathcal{C}_i modélisée par une *projection perspective* s'écrivent :

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

où \tilde{x} est un point 3D de coordonnées homogènes $\tilde{x} = (x \ y \ z \ 1)^t$ dans le repère de la scène \mathcal{R}_0 , et où $\tilde{p}_i = (X_i \ Y_i \ 1)^t$ représente les coordonnées de son projeté dans le plan image I_i . La matrice P_i est la *matrice de projection* associée à la caméra \mathcal{C}_i . Elle s'exprime en fonction de la matrice K représentant les *paramètres intrinsèques* de la caméra, et en fonction des transformations (rotation R_i et translation t_i) permettant de passer du repère de la scène \mathcal{R}_0 au repère de la caméra \mathcal{R}_i , par l'équation :

$$P_i = K(R_i|t_i).$$

La matrice K des *paramètres intrinsèques* de la caméra est définie par :

$$K = \begin{pmatrix} \frac{f}{t_x} & \frac{-f}{t_x \tan \theta} & X_c \\ 0 & \frac{f}{t_y \sin \theta} & Y_c \\ 0 & 0 & 1 \end{pmatrix},$$

où f est la *distance focale* de la caméra, et où l_x et l_y dénotent la taille en x et y des pixels. Le terme θ représente l'angle entre les axes x et y du plan image, et X_c et Y_c sont les coordonnées de la projection perpendiculaire du centre optique dans le plan image. Les matrices de rotation R_i et de translation t_i définissent les *paramètres extrinsèques* associés à la caméra \mathcal{C}_i .

Les paramètres intrinsèques et extrinsèques, aussi appelés paramètres de prise de vue, sont obtenus par des opérations de calibration ou d'auto-calibration. La *calibration* consiste à estimer les paramètres de prise de vue à l'aide de mires (objets fournissant des points 3D connus), et d'images de ces mires. L'*auto-calibration* consiste à estimer les paramètres de prise de vue uniquement à partir de données des images, préalablement mises en correspondance par identification et regroupement des projetés d'un même point 3D. Plus précisément, si on suppose les paramètres intrinsèques constants et connus, les paramètres extrinsèques R_i et t_i peuvent être estimés comme suit. Le repère de la scène est choisi arbitrairement confondu avec le repère \mathcal{R}_1 . Les matrices de rotation et de translation de la caméra \mathcal{C}_i sont donc celles qui font passer du repère \mathcal{R}_1 au repère \mathcal{R}_i de l'image I_i . Pour estimer ces matrices, on définit la *matrice essentielle* $E = T_i.R_i$ et la *matrice fondamentale* $F = K^{-t}.E.K^{-1}$. La matrice F est elle-même estimée à partir d'un ensemble de points mis en correspondance entre les images I_1 et I_i , par résolution du système formé de la *contrainte épipolaire* appliquée à chaque paire de points

$$p_1^t.F.p_i = 0.$$

La décomposition de E en valeurs singulières fait apparaître un produit de matrices qui permet de déduire la matrice de rotation R_i et le vecteur de translation t_i . Résoudre le problème de reconstruction 3D revient alors à trouver \tilde{x} connaissant les \tilde{p}_i , c'est-à-dire à résoudre le système d'équations (1) relativement aux coordonnées \tilde{x} . Comme toute résolution de problème inverse, la reconstruction 3D est très sensible aux erreurs de mesure. Sa résolution requiert une bonne précision des mesures faites dans l'image, ainsi qu'un choix judicieux des techniques d'optimisation numérique.

3.2. Modélisation du mouvement par maillages actifs

Mots clés : *maillage, éléments finis, mouvement.*

Les performances d'un algorithme de compression vidéo dépendent pour une large part de la capacité à modéliser de façon fiable le mouvement présent dans la scène. TEMICS privilégie la modélisation du mouvement et des déformations par maillages actifs. Les maillages actifs permettent une représentation continue et bien adaptée du mouvement observé. Ces modèles sont aussi utilisés dans les algorithmes de segmentation et de reconstruction 3D des objets de la scène.

Une séquence vidéo représente l'enregistrement d'une scène physique observée. Les variations observées dans la séquence sont liées aux variations des différents objets : variation de couleur, de forme, et principalement déplacements et déformations. Ces déplacements et déformations peuvent être décrites à l'aide du mouvement apparent \vec{v} mettant en relation les valeurs de deux images successives I_t et I_{t+1} comme suit :

$$I_{t+1}(p) = I_t(p - \vec{v}(p)) + n(p)$$

où le bruit n (en général faible) représente les variations liées aux autres phénomènes. Le champ de mouvement continu \vec{v} peut être défini par un ensemble de vecteurs de mouvement \vec{v}_i associés aux noeuds d'un maillage. Les maillages permettent de représenter correctement les mouvements locaux tels que translation, dilatation, rotations, ou étirement. Le mouvement apparent s'écrit alors sous la forme :

$$\vec{v}(p) = \sum_i \vec{v}_i \phi_i(p)$$

où $\phi_i(p)$ représente la fonction de forme associée au noeud i .

Les paramètres du modèle sont estimés en minimisant l'erreur quadratique moyenne des erreurs de prédiction ($\min_{\vec{v}_i} \sum_p [I_{t+1}(p) - I_t(p - \vec{v}(p))]^2$). Cette minimisation est souvent réalisée à l'aide d'algorithmes de type Gauss-Seidel ou de recherche exhaustive couplés à une représentation hiérarchique des modèles de mouvement et de la texture. La figure 1 illustre une telle estimation du mouvement apparent. Cette technique d'estimation du mouvement est bien adaptée au suivi temporel des objets, donc à la segmentation long terme ou encore à la reconstruction 3D des objets. Ce suivi des déformations fait apparaître un modèle auto-régressif le long des trajectoires de mouvement représentant les variations du signal : $I_{t+1}(p) = I_t(p - \vec{v}(p)) + n(p)$. L'utilisation de transformations le long des trajectoires de mouvement peut alors permettre de décorréler le signal dans la dimension temporelle.

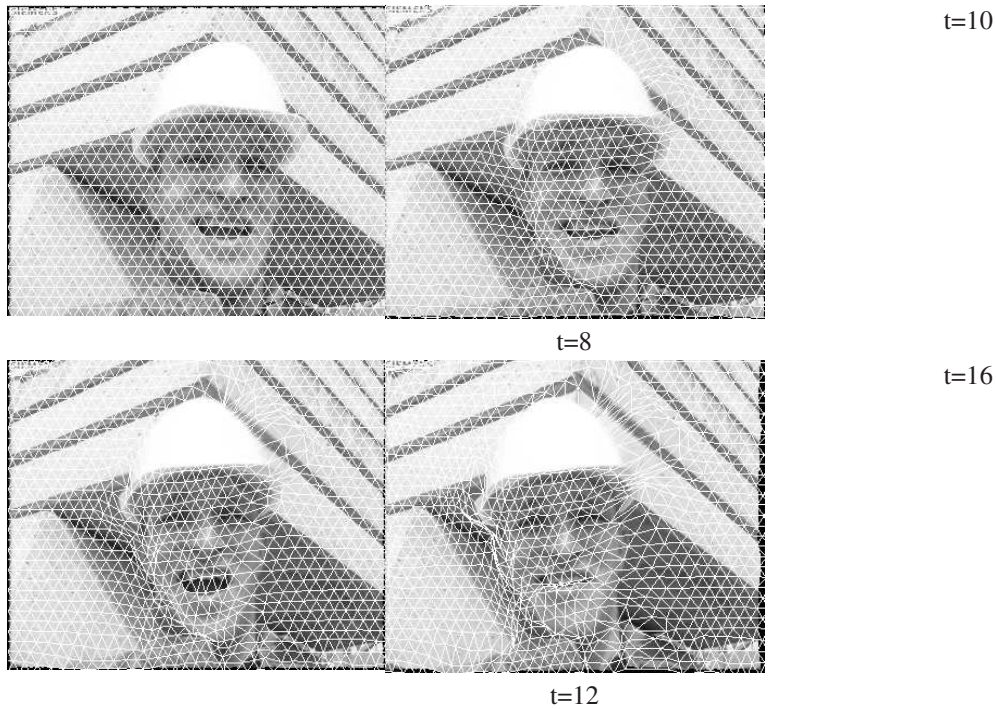


Figure 1. Exemple de modélisation des déformations au cours du temps au sein d'un vidéo via un maillage actif.

3.3. Théorie débit-distorsion et formalisme MDL

Mots clés : Limite OPTA (Optimum Performance Theoretically Attainable), Formalisme MDL (Minimum Description Length), allocation de débit, optimisation débit-distorsion, codage avec pertes, codage par descriptions multiples, modélisation de canal.

Les travaux de codage (avec ou sans pertes) et les travaux de codage conjoint source-canal de TEMICS reposent sur des concepts fondamentaux de la théorie de l'information, comme les notions de quantité d'information, d'entropie d'une source avec et sans mémoire, de capacité d'un canal et de limite de performance débit-distorsion. Les algorithmes de compression sont conçus de manière à approcher au mieux la limite de performance débit-distorsion $R(D)$ pour un signal considéré. Le coût de description d'un message peut aussi être représenté à l'aide du formalisme MDL (minimum description length) qui se rapproche des notions de complexité de Kolmogorov et de complexité stochastique.

Le théorème du codage de source définit les limites de performances du codage sans pertes et avec pertes. Dans le cas du codage sans pertes, la borne inférieure de débit est fournie par l'entropie de la source. Dans le cas du codage avec pertes, elle est donnée par une fonction débit-distorsion $R(D)$. Cette fonction $R(D)$ établit la quantité d'information minimale nécessaire à la représentation du message pour une distorsion donnée de ce message, ou à l'inverse, établit la distorsion minimale pour un débit (ou une quantité d'information) donné.

Cette fonction débit-distorsion limite, appelée OPTA (*Optimum Performance Theoretically Attainable*), est en général difficile à déterminer de manière analytique, sauf pour le cas de sources particulières (théoriques), comme par exemple pour des sources Gaussiennes. Dans le cas de sources réelles (images ou signaux vidéo), les algorithmes de compression nécessitent l'utilisation de procédures d'optimisation débit-distorsion. Pour ce faire, une minimisation Lagrangienne est généralement mise en œuvre :

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

Cette optimisation est réalisée en dérivant le Lagrangien, J , par rapport au paramètre à l'origine de la distorsion, c'est-à-dire souvent par rapport au pas de quantification Q .

Dans un cadre d'optimisation des performances de la qualité de service d'un système de communication, il est en outre nécessaire de prendre en compte les caractéristiques du canal et du codage de canal, en développant des stratégies de codage conjoint source-canal. On cherche alors à minimiser une distorsion globale, définie comme la somme de la distorsion due à la compression du signal et de la distorsion induite par le bruit de transmission sur le canal, sous la contrainte d'un débit global donné. Les techniques d'optimisation débit-distorsion conjointes source-canal permettent d'améliorer le compromis compression - robustesse au bruit de transmission, et d'avoir un partage optimal de la bande passante d'un canal de communication entre le débit utile alloué à la source et le débit de redondance alloué à la protection. La redondance nécessaire à la protection du signal peut être générée à l'aide de codes correcteurs traditionnels ou alors être introduite dans la représentation compressée du signal au travers de transformations par bases redondantes (voir section 3.5).

Le formalisme MDL (Minimum Description Length) constitue un outil méthodologique supplémentaire pour quantifier le coût de description global, ou la quantité d'information minimale, nécessaire pour représenter un message. Cet outil se rapproche de la notion de quantité d'information et de complexité stochastique définie par Kolmogorov. Il est basé sur une modélisation du signal par une suite de réalisations de variables aléatoires. En notant s une chaîne d'observations des réalisations de ces variables, et θ le vecteur des paramètres du modèle de représentation, le principe du formalisme MDL est de minimiser le coût de description de la chaîne s :

$$MDL(s) = \min_{\theta} \left\{ -\log_2 P_{\theta}(s) + DL(\theta) \right\},$$

où $DL(\cdot)$ représente le coût de description. Le terme $-\log_2 P_{\theta}(s)$ correspond au coût de codage idéal de la chaîne s connaissant le modèle de représentation θ ; $-\log_2 P_{\theta}(s) + DL(\theta)$ est le coût de description global de la chaîne s en utilisant le modèle θ . Le formalisme MDL est souvent utilisé dans des schémas de codage d'images fixes et animées, afin d'optimiser le modèle de représentation de l'image.

3.4. Représentation non redondante des signaux de source

Mots clés : Transformations par ondelettes, bancs de filtres multi-fréquences.

Les algorithmes de compression du signal présentent toujours plus ou moins la même architecture : une première étape de décorrélation suivie d'opérations de quantification et de codage. Les transformations (transformations en ondelettes ou par paquets d'ondelettes) qui remplacent le signal d'origine par sa représentation sur une base de fonctions, correspondent à une telle approche et constituent des éléments essentiels des schémas de codage développés par TEMICS. Ces transformations constituent en effet des outils privilégiés

pour une représentation scalable des signaux compressés, et donc pour permettre des fonctions de régulation de débit dans des environnements de communication hétérogènes.

Les techniques de représentation temps-fréquence (ou multi-résolution) du signal consistent à décomposer le signal par bancs de filtres multi-fréquences. Une structure itérée de décomposition par bancs de filtres va permettre de créer un ensemble de sous-signaux (aussi appelés sous-bandes) qui vont contenir l'information relative aux différentes bandes de fréquences contenues dans l'image ou le signal vidéo, comme illustré sur la figure 2. Ces structures de décomposition, conduisant naturellement à une représentation multi-échelle et scalable du signal, occupent une place de plus en plus importante dans les systèmes de compression et de communication. Elles constituent en effet des outils privilégiés pour la mise en place de mécanismes de contrôle de congestion et de régulation de débit dans des environnements de communication hétérogènes, et en particulier multi-points.

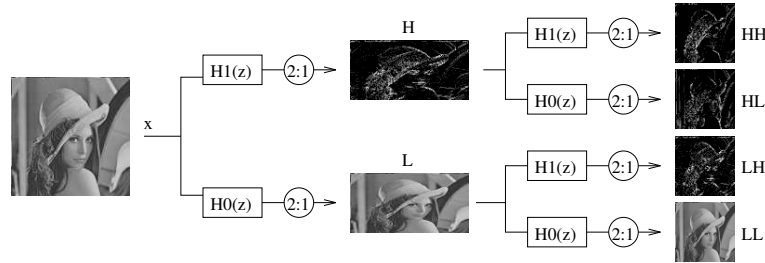


Figure 2. Analyse multi-résolution d'un signal 2D par un banc de filtres : le signal est décomposé sur 4 espaces orthogonaux entre eux, un espace d'approximation (correspondant au signal LL contenant les basses fréquences verticales et horizontales) et trois espaces de détail (correspondant aux signaux LH, HL et HH).

Les décompositions non redondantes par ondelettes ou bancs de filtres multi-fréquences sont nécessairement à échantillonnage critique, c'est-à-dire que la somme des échantillons contenus dans les sous-signaux est égale au nombre d'échantillons du signal à décomposer. L'application de ces transformations à une décomposition spatio-temporelle du signal vidéo conduit de façon naturelle à une représentation scalable du signal (scalabilité spatiale, temporelle et en qualité). Les structures dites « lifting » de mise en oeuvre de ces transformations facilitent en outre les traitements adaptatifs, ou encore l'introduction de traitements non-linéaires.

3.5. Représentation des signaux par bases de fonctions redondantes

Mots clés : Bancs de filtres suréchantillonnés, codes correcteurs, descriptions multiples, codage conjoint source-canal.

Les environnements de communication hétérogènes et à qualité de service non garantie posent de nouveaux problèmes de représentation des signaux. Pour minimiser l'impact du bruit de transmission sur la qualité du signal reconstruit, il s'avère nécessaire de maintenir une certaine corrélation ou redondance dans la représentation compressée du signal. L'une des approches étudiées par TEMICS repose sur l'utilisation de transformations redondantes.

Les décompositions sur des ensembles de fonctions redondantes (en anglais appelées *overcomplete frame expansions*) sont utilisées pour accroître la résistance d'un signal à un bruit additif. Nous les considérons ici pour accroître la robustesse du signal à des effacements (ou pertes) et à l'ajout de bruit impulsif sur des liens sans fils. Ces transformations peuvent être vues comme une décomposition du signal sur un ensemble de fonctions de bases $\Phi = \{\phi_k\}_{k=1}^m$ redondantes et être formalisées par une transformation linéaire $F : C^n \rightarrow C^m$ d'un espace de dimension n vers un espace de dimension $m \geq n$. Conceptuellement, l'approche s'apparente à un code de canal en bloc, défini sur le corps des réels, et mis en oeuvre au travers de la transformation. Les paramètres n/m définissant le taux de rendement du code et les fonctions de base de la transformation peuvent être adaptés aux caractéristiques du canal à un instant donné. L'interprétation de ces

transformations redondantes sous forme de codes de canal permet d'appliquer certains outils de décodage de canal (comme par exemple les algorithmes de décodage basés syndrome ou des estimateurs Bayesiens) pour exploiter la redondance et corriger les erreurs (ou pertes) dues à la transmission. Cependant, à la différence de codes correcteurs sur les corps finis, les mots de codes sont ici quantifiés à des fins de compression. Les techniques de décodage se heurtent donc à la difficulté de localiser et d'estimer le bruit introduit par le canal en présence d'un bruit de fond dû à la quantification des mots de code.

3.6. Tatouage d'images

Mots clés : *tatouage (watermarking), signature, copyright, authentification, détection, extraction robuste, étalement de spectre.*

L'objectif général du tatouage (*watermarking* en anglais) est de dissimuler des informations (une marque) dans un contenu multimédia. Cette marque doit être indécélable, infalsifiable et ineffaçable. Son extraction doit être robuste face à divers types d'attaques que l'image pourrait subir (compression, retaille, filtrage, transformations géométriques, ...). L'insertion du signal de tatouage est souvent réalisée dans un domaine transformé, permettant en particulier d'exploiter des fonctions de masquage du système visuel afin d'assurer l'invisibilité de la marque. Bien qu'il soit possible d'établir un parallèle entre le tatouage et une situation de communication classique, TEMICS privilégie un formalisme où le marquage tire profit de la connaissance a priori du contenu à tatouer. Ceci s'apparente à un problème de communication avec information adjacente.

Le premier problème posé est le choix du domaine transformé. Il s'agit de déterminer quelles composantes de l'image ou de la vidéo recevront le signal de tatouage. Soit $E(\cdot)$ cette fonction d'extraction allant de l'espace des contenus \mathcal{C} à l'espace des composantes, isomorphe à \mathbb{R}^N .

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

Savoir quantifier l'impact perceptif de l'incrustation de la marque dans ce domaine est indispensable. Cette mesure de distortion perceptive se traduit par une fonction de coût $d(\mathbf{V}_t - \mathbf{V})$ dans \mathbb{R}^N ne devant pas dépasser une certaine contrainte d_t . De plus, une modélisation de la distribution statistique des vecteurs \mathbf{V} permet une optimisation de la détection du signal de tatouage de faible puissance. Enfin, une modélisation de l'attaque $\mathbf{A} = \mathbf{V}_a - \mathbf{V}$ dans cet espace est nécessaire afin d'évaluer la robustesse théorique de la technique de tatouage et d'inventer des contre-attaques. Le coût de l'attaque est contraint par $d(\mathbf{V}_a - \mathbf{V}) < d_a$.

Une fois les contenus et l'attaque modélisés, il s'agit de trouver des schémas de communication adaptés à ces contraintes. Traditionnellement, un système de transmission à étalement de spectre à séquence directe (type multiporteuse CDMA) est choisi. Le facteur d'étalement fixe le compromis robustesse - capacité pour une distortion d'incrustation donnée. Il s'agit finalement d'un étiquetage $S(\cdot)$ associant à un message $m \in \mathcal{M}$ un signal de \mathbb{R}^N :

$$\begin{aligned} S(\cdot) : \quad \mathcal{M} &\longmapsto \mathbb{R}^N \\ m &\longrightarrow \mathbf{S} = S(m) \end{aligned}$$

À la détection, la fonction de décodage $S^{-1}(\cdot)$ travaille sur le signal reçu \mathbf{V}_a dans lequel le signal de tatouage interfère avec le bruit dû au contenu original (\mathbf{V}), aux attaques (\mathbf{A}) et enfin aux interférences entre symboles émis. Il faut donc trouver les fonctions $\{S(\cdot), S^{-1}(\cdot)\}$ telles que la chaîne de communication soit optimisée pour les contraintes $\{d_t, d_a\}$ données. Cela revient à maximiser la probabilité de décoder correctement le message caché :

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

Un nouveau paradigme est apparu récemment. Il stipule que le contenu original V ne doit pas être considéré comme un bruit mais comme un *état du canal* de transmission connu à l'incrustation seulement. Le signal de tatouage dépend alors de la connaissance de l'état du canal : $S = S(m, V)$. Ce nouveau paradigme est appelé communication avec information adjacente. Il induit l'invention de schémas de communication *non* classiques. S'il est vrai qu'une solution théorique optimale a été exhibée lorsque l'attaque se résume à l'addition de bruit Gaussien sur des vecteurs infiniment longs, la recherche de nouveaux codes réellement efficaces en pratique reste un défi pour la communauté tatouage à l'heure actuelle.

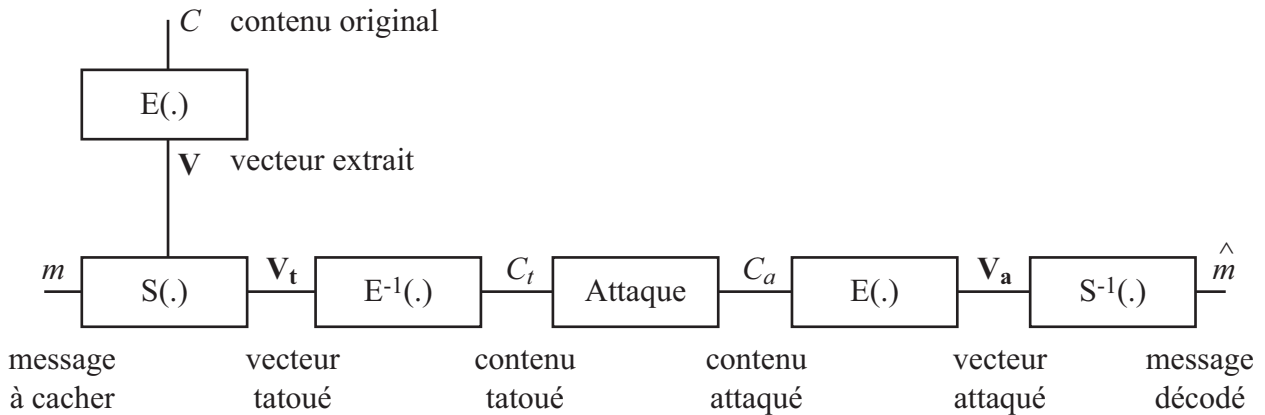


Figure 3. Schéma de tatouage classique.

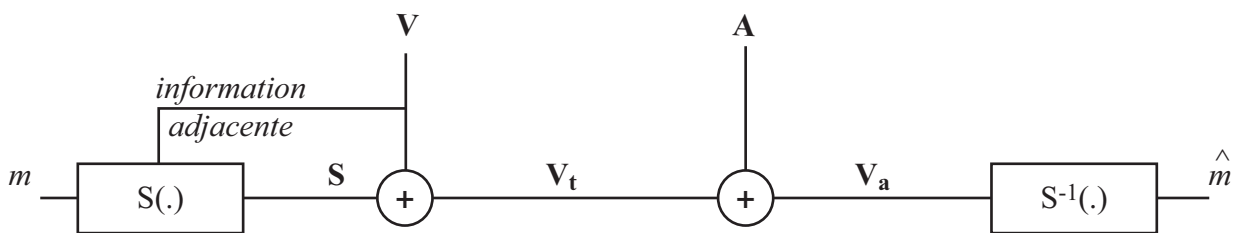


Figure 4. Nouveau paradigme : le tatouage comme une communication avec information adjacente.

4. Domaines d'application

4.1. Panorama

Nous nous intéressons à trois grands types d'applications. Le premier concerne la compression d'images fixes et animées. Nous explorons dans ce cadre les domaines de la compression scalable du signal vidéo et du couplage de la fonction de compression à des fonctionnalités telles que la navigation, l'indexation et le tatouage. Le second secteur concerne les applications de télécommunications incluant les applications de consultation multimédia (ex : visite virtuelle de sites), de diffusion (diffusion audiovisuelle sur Internet) et de communications de groupes (ex : visioconférence, réunions virtuelles). Nous avons dans ce domaine d'application de nombreuses collaborations nationales (cinq projets RNRT) et Européennes (deux projets IST). Le troisième secteur a trait à la protection et à l'enrichissement des contenus visuels (images fixes et vidéo) par insertion de données numériques dans le signal (tatouage). Enfin, certains travaux menés dans l'équipe peuvent avoir des retombées dans les domaines de la production, de la post-production vidéo et de la structuration de séquences vidéo pour de l'aide à la navigation, même si ces derniers domaines ne constituent pas des domaines d'application privilégiés par le projet.

4.2. Compression d'images fixes et animées avec fonctionnalités avancées

La compression d'image et du signal vidéo constitue un cadre applicatif naturel d'un grand nombre de techniques que nous développons. La compression occupe en effet une place prépondérante dans de nombreux domaines d'application nécessitant un stockage de l'information image ou vidéo, ou une transmission de cette information sur des réseaux à bande étroite, voire à qualité de service non garantie. Néanmoins, nos travaux ne cherchent pas à répondre seulement à des objectifs de compression mais cherchent à allier la compression à des fonctionnalités nouvelles telles que la navigation dans une scène par reconstruction de points de vues intermédiaires, la transmission fiable de ces données sur des réseaux à qualité de service non garantie et hétérogènes, ou encore telles que la protection et l'enrichissement des contenus par insertion de données numériques dans le signal (tatouage).

Ainsi, nos travaux algorithmiques d'analyse, de représentation par maillages (et en particulier par maillages hiérarchiques), de modélisation 3D de scènes réelles, et de création de vues intermédiaires permettent d'allier des fonctions de compression et de navigation dans des scènes virtuelles 3D issues de séquences vidéos naturelles. Ils s'appliquent donc directement à toute application multimedia mettant en œuvre des scènes virtuelles 3D, par exemple visite de musées virtuels, télé-présence, jeux,... Ce domaine d'application fait l'objet du projet RNRT-V2net.

De manière générale, le besoin d'interactions avec le contenu vidéo, d'accès et de navigation au sein du contenu demande de structurer l'information vidéo (objets, mosaïques, plans, ...) et de mettre au point des représentations compactes de cette structuration. Il semble donc naturel de chercher à concevoir des schémas de compression générant une information qui peut être aisément exploitée lors d'une structuration de la séquence vidéo et à l'inverse exploitant de manière efficace l'information de description ou de structuration (qui peut se présenter sous forme de métadonnées) à des fins de compression. L'accès à des bases de données vidéo structurées constitue ainsi un cadre applicatif naturel de nos travaux en compression avec fonctionnalités avancées.

4.3. Communication multimédia

Les applications diffusant du contenu multimédia sont confrontées aux problèmes de transmettre sur des réseaux aux performances variables des flux volumineux, et selon un modèle plus complexe que le traditionnel point-à-point homogène, comme par exemple en multi-point dans des environnements hétérogènes fixes ou mobiles. Pour réussir à transmettre les flux sur les réseaux, ces applications ont traditionnellement recours aux techniques relevant du domaine de la compression. Or les outils définis à des seules fins de compression présentent certaines limitations : ils sont très sensibles aux pertes ou aux erreurs de transmission. Nos travaux en codage conjoint source-canal ont pour objectifs de fournir des solutions au problème de l'optimisation de la qualité de service dans des environnements offrant un service de transport non garanti (comme par exemple, l'Internet d'aujourd'hui). Ils s'appliquent ainsi et de façon naturelle aux applications multimédia communicantes dans de tels environnements (Internet, liens sans fils). La transmission d'images ou de signaux vidéo dans des environnements hétérogènes (fixes et mobiles) en multi-point, constitue également un domaine d'application des travaux de recherche que nous menons. En particulier, la mise au point de schémas de codage scalables nous permet d'offrir des techniques élaborées de régulation de flux pour faire du contrôle de pertes et de gestion dans des contextes de diffusion multi-points.

Les domaines applicatifs plus particulièrement traités par le projet, au travers de ses partenariats industriels bilatéraux et européens, concernent tout d'abord les applications de consultation multimédia sur Internet (télé-achat, télé-visite, télé-enseignement). Dans ce type d'application, le codage des contenus multimédia (audiovisuels, 3D synthétique) est en général réalisé en différé, ce qui conduit à des difficultés supplémentaires pour assurer un contrôle de congestion efficace. Cette caractéristique va ainsi fixer un cadre spécifique pour une partie de nos travaux. Ce secteur applicatif constitue un point de rencontre avec les domaines de l'indexation, de la consultation et de la gestion de bases de données vidéo. Nos travaux concernent également les communications de groupes (ex : visioconférence) que nous traitons en collaboration avec France Télécom R&D et au sein du projet RNRT VIP. Enfin ils concernent les vidéocommunications mobiles, domaine

d'application que nous abordons en collaboration avec le Celar au travers d'une application de téléguidage d'un robot terrestre mobile, et au sein du projet RNRT Cosocati.

4.4. Protection, authentification et enrichissement de contenus images et vidéo

Les technologies numériques, qui permettent l'apport de fonctionnalités nouvelles dans les applications de communication vidéo et multimédia, ainsi que l'apparition de réseaux ouverts comme l'Internet ne sont pas sans susciter de nouveaux dangers pour la propriété intellectuelle et les droits de diffusion. De l'avis des professionnels de la photographie, le problème principal lié à la mise à disposition de photos au travers de réseaux ouverts est le pillage illicite de leur patrimoine. Il est en effet possible de faire des copies parfaites des contenus, la redistribution est devenue aisée et le traçage des redistributions est difficile. Ce problème, exprimé dans un contexte informatique, demande de mettre en place un environnement permettant d'une part de détecter qu'une image ou une vidéo fassent illégalement partie d'une collection d'images mises à disposition par un tiers (traçage de copies) et d'autre part d'exhiber une preuve de propriété irréfutable (authentification). Ces problèmes de vérification de propriété intellectuelle, de traçage d'usage illicite, et d'authentification des images numériques constituent ainsi un domaine d'application à forts enjeux de nos études de marquage (aussi appelé tatouage ou "watermarking" en anglais). Mais, dans ce contexte, la notion de sécurité s'ajoute aux caractéristiques classiques du tatouage (invisibilité, robustesse, capacité). Jusqu'à présent, l'attaque est modélisée par un bruit polluant le canal de transmission de la marque. Or, dans ces applications, l'adversaire est intelligent et il ne se contente pas d'ajouter un bruit aléatoire au contenu. Il y a un besoin urgent de définir une méthodologie pour analyser le niveau de sécurité que produit un schéma de tatouage.

Il existe aussi une grande interaction entre le tatouage et l'indexation par le contenu. Cette combinaison peut en effet permettre de couvrir les deux aspects que sont traçage et authentification, mais aussi de renforcer la puissance et la robustesse de chaque technique prise individuellement car l'une et l'autre s'épaulent dans le processus de recherche et de vérification. Ces domaines d'application vont être couverts dans le cadre des nouveaux projets RNRT Diphonet et IST Busman.

Néanmoins, la protection de copyright, l'authentification et le traçage de copies illicites ne sont pas les seuls domaines d'application du tatouage. Les outils d'insertion de données numériques dans un signal vidéo peuvent être utilisés pour transporter des métadonnées dans un environnement de diffusion hétérogène, assurant ainsi une certaine transparence du transport de ces métadonnées et une certaine persistance de ces données dans un cadre de transcodage inhérent à l'application de diffusion sur une infrastructure hétérogène de transport (DVB, Internet, accès mobiles, etc...) ou vers un parc hétérogène de terminaux. Cela conduit alors à des contenus enrichis par des métadonnées enfouies dans le signal. Elles 'voyagent' alors avec le contenu et elles sont exploitées de façon transparente par le terminal pour offrir différentes fonctionnalités à valeur ajoutée.

5. Logiciels

5.1. Plate-forme de communication vidéo

Participants : Gabriele Davini, Bilal El Ali, Christine Guillemot, Laurent Guillo [correspondant], Jérôme Viéron.

L'équipe TEMICS a développé, au travers de différents contrats (RNRT-VISI, IST-SONG) une plate-forme de communication vidéo. Elle permet, en émulant une large gamme de conditions de réseaux (congestion, délai perte, différenciation de services), de tester en conditions réalistes nos algorithmes de représentation, de codage et de codage conjoint source-canal. Les composants logiciels de cette plate-forme réalisés par l'équipe sont en cours de dépôt auprès de l'APP (Agence pour la Protection des Programmes). Une évolution de cette plate-forme est prévue en collaboration avec l'équipe ARMOR. Elle permettra de réaliser de nombreuses études intégrant l'état de l'art aussi bien pour les aspects de codage de séquences vidéos que pour le transport de l'information (différenciation de services, IPV6, streaming, ...).

5.2. Codeur vidéo scalable

Participants : Brendan Catherine, Christine Guillemot [correspondante], Stéphane Pateux, Jérôme Viéron.

Un logiciel de codage et de décodage vidéo scalable a été développé. Ce codeur permet d'atteindre des facteurs de compression très supérieurs à ceux fournis par le standard MPEG-4 tout en offrant des fonctionnalités de scalabilité à grain fin. Ces fonctionnalités sont essentielles pour permettre une adaptation dynamique du débit des signaux à transmettre à la bande passante disponible sur le réseau, celle-ci variant le temps.

5.3. Tatouage robuste d'image

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

Un logiciel de tatouage robuste d'images a été développé. Ce logiciel permet d'insérer dans une image un message de façon robuste avec différentes techniques de tatouage développées au sein du projet Temics. Ces techniques sont basées sur une insertion du message dans un domaine transformé ondelettes, l'utilisation de technique d'étalement de spectre avec modulation d'amplitude permettant de prendre en compte les propriétés de sensibilités perceptuelles, ainsi que de techniques de codage par codes correcteurs avec aou sans prise en compte de l'information de bord. Ce logiciel permet également d'extraire le message inséré et de donner des mesures de confiance dans l'extraction de ce message extrait.

5.4. Codeur-décodeur vidéo basé sur une représentation par modèles 3d

Participants : Raphaële Balter, Philippe Coval, Luce Morin [correspondante].

Un logiciel d'analyse, codage, décodage et visualisation de séquence vidéo a été développé. A partir d'une séquence vidéo de scènes statiques acquises par une caméra monoculaire en mouvement ce logiciel permet d'obtenir automatiquement une représentation par un flux de modèles 3D texturés. Les modèles sont extraits par des techniques de stéréovision couplées à une estimation de mouvement dense pour obtenir les correspondances. Un ajustement de faisceaux par fenêtre glissante assure une cohérence entre les modèles successifs. Une séquence virtuelle est alors reconstruite par reprojexion des modèles texturés. Les fonctionnalités induites par cette représentation permettent d'envisager de nombreuses applications de changement d'illumination, d'incrustation ou de navigation interactive. De plus un schéma de compression basé sur la représentation proposée permet d'atteindre des bas et très bas débit (16 à 256kb/s en CIF à 25Hz) avec une qualité visuelle satisfaisante.

6. Résultats nouveaux

6.1. Analyse et modélisation de séquences d'images

Mots clés : *mouvement, champs denses, disparité, segmentation, suivi temporel, triangulation, maillages 2D et 3D, maillages actifs, maillages hiérarchiques, contour actif, cartes de profondeur, géométrie projective, modèles par facettes, réalité augmentée, ombrage, modèles d'illumination, image mosaïque.*

Pour répondre à des besoins de navigation dans une scène, au travers d'un réseau de télécommunications, nous avons développé un schéma global de reconstruction et de représentation d'une séquence vidéo par un ensemble de modèles 3D. La représentation par modèles 3D permet la reconstruction de vues intermédiaires. Dans cette étude nous considérons des grandes scènes complexes et de longues séquences d'images, dont l'acquisition est faite par une caméra unique et mobile. Cette étude démarrée en 2000 s'est poursuivie cette année en cherchant à résoudre quelques points durs tels que la prise en compte des occultations dans l'estimation du mouvement ou encore la gestion des discontinuités entre les modèles. Les travaux de l'année ont aussi porté sur le développement d'un algorithme de segmentation et de codage conjoint par maillage actifs et mosaïques généralisées. Les maillages actifs sont exploités pour la création de mosaïques généralisées d'objets déformables.

6.1.1. Modélisation 3d par analyse de séquences vidéo monoculaires

Participants : Raphaële Balter, Luce Morin.

En 2001 (voir rapport d'activités), un schéma de reconstruction-représentation d'une séquence par création d'un ensemble de modèles 3D a été développé. L'originalité de l'approche est de ne pas reconstruire un modèle 3D unique de la scène, mais une suite de modèles 3D indépendants liés par des points de vue communs (images clés). Ce schéma génère un flux de modèles pouvant être transmis sur un réseau en vue d'une navigation à distance dans la scène.

La navigation libre correspond à un changement de trajectoire de la caméra par rapport au chemin original. Dans ce contexte, l'utilisation de plusieurs modèles 3D introduit des limites lorsque l'on s'éloigne de la trajectoire originale de la caméra. Pour garantir une transition douce entre deux modèles successifs, il est nécessaire de mettre en œuvre un post-traitement. Nos travaux ont porté cette année sur l'amélioration de la qualité visuelle des séquences reconstruites par introduction d'un morphing 3D. Le morphing consiste à créer une suite de maillages 3D intermédiaires permettant de passer continuellement d'un maillage source à un maillage cible. Il comporte deux étapes principales :

- la mise en correspondance de chaque point du maillage source avec un point du maillage cible (et vice et versa).
- l'interpolation qui définit la trajectoire entre chaque paire de points en correspondance.

Ces deux étapes doivent être entièrement automatiques. Elles s'appuient sur les cartes de profondeurs maillées et les positions de la caméra au cours de la séquence. Cette technique conduit à une amélioration très nette du rendu des images dans le cas de la navigation libre comme le montrent la figure 5. Elle présente cependant quelques limitations en présence de zones découvertes entre les deux modèles. Une étude de morphing bidirectionnel permettant de traiter symétriquement zones découvertes et recouvertes est en cours.



Figure 5. Image reconstruite correspondant à la position de caméra originale (a) et image extraite d'une séquence de navigation libre avec fondu 3D (b) et avec morphing 3D (c).

6.1.2. Segmentation d'objet vidéo pour le codage

Participants : Marc Chaumont, Dubhe Chavira-Martinez, Henri Nicolas, Stéphane Pateux.

Dans le cadre d'une application de codage, la segmentation idéale ne correspond pas forcément à une segmentation sémantique (un exemple typique concerne les objets immobiles qui n'ont pas besoin d'être segmentés). Nous avons ainsi développé une technique de segmentation basée sur une modélisation des objets sous la forme d'une texture variant peu au cours du temps, et ayant un mouvement continu. Pour chaque objet, un modèle de déformation par maillages actifs est utilisé. Un suivi long-terme de ces maillages permet de réaliser une segmentation des différents objets présents basée sur des mesures de cohérence long-terme, et non plus uniquement court-terme. Le problème de la segmentation est ainsi posé comme un problème d'étiquetage minimisant une fonctionnelle énergétique (modèle Markovien). Les termes énergétiques utilisés, ainsi que les modèles proposés sont directement liés à des modèles de codage. L'objectif final est d'obtenir une segmentation optimisée pour le codage avec la sélection adéquate des modèles de représentation pour chaque objet. Les travaux de cette année ont porté sur le développement de différentes techniques d'initialisation,



Figure 6. Segmentation obtenue à l'aide d'une modélisation par maillages actifs des différents objets.

d'introduction de modèles hiérarchiques, et d'extension des modèles utilisés notamment via l'introduction d'une classe de rejet pour rejeter toutes les zones ne collant pas aux modèles proposés.

6.1.3. Codage vidéo basé sur une représentation par maillages et mosaïques généralisées d'objets vidéo

Participants : Nathalie Cammas, Marc Chaumont, Stéphane Pateux.

Les maillages actifs permettent de reconstituer la texture des différents objets présents dans la scène, créant ainsi un ensemble de mosaïques généralisées dynamiques pour les différents objets. Ceci nous a conduit à développer un schéma de codage vidéo basé objet de type analyse-synthèse. Les différentes informations nécessaires à la reconstruction de la vidéo sont codées : texture, mouvement et forme. Les informations de texture et de mouvement sont codées par l'intermédiaire d'une décomposition en ondelettes $2D + t$. L'information de forme est codée via l'utilisation d'une représentation paramétrique spatio-temporelle hiérarchique. Un schéma de reconstruction de type synthèse d'image est utilisé pour reconstruire chaque image de la scène vidéo. Les premiers résultats obtenus montrent un apport pour le codage à bas débit de flux vidéos par rapport à des approches classiques. La segmentation en objets permet d'avoir une représentation efficace des différents objets dans la scène, nécessitant ainsi moins de débit pour les représenter. Par ailleurs, l'allocation d'un débit plus élevé sur les informations de texture permet d'avoir une meilleure qualité visuelle grâce à une diminution des artefacts (absence de phénomènes de blocs) ainsi qu'à une meilleure finesse et stabilité temporelle. Ces résultats peuvent encore être affinés en optimisant l'allocation du débit parmi les différentes entités à coder (forme, texture et mouvement) et objets présents dans une scène.

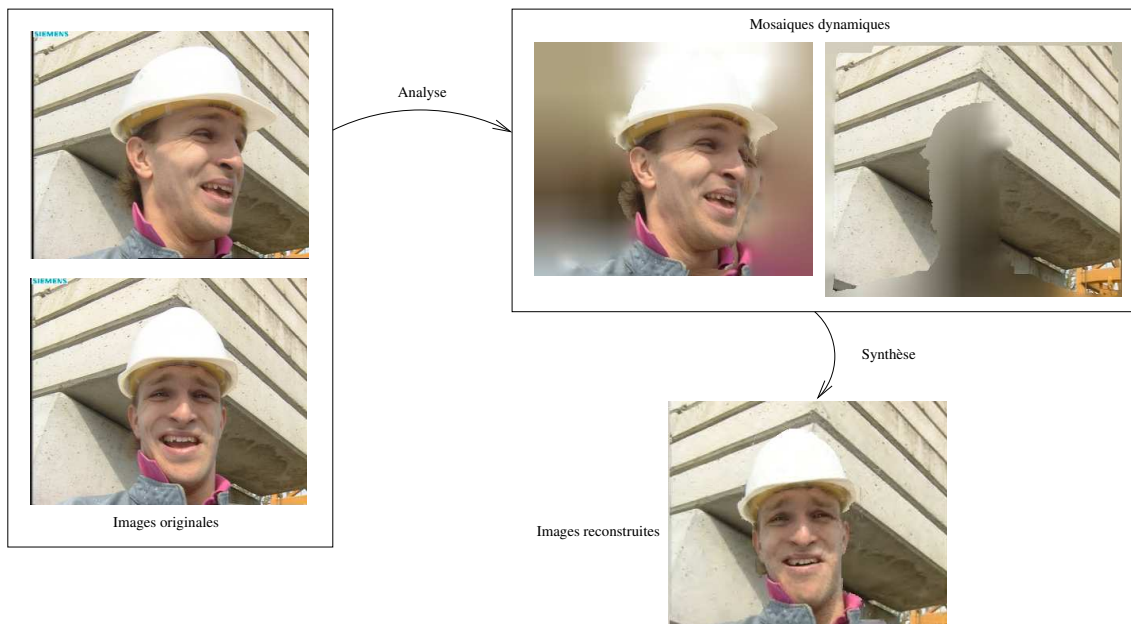


Figure 7. Exemple de représentation par objets et mosaïques dynamiques généralisées.

6.1.4. Analyse des conditions d'illumination dans une séquence d'images

Participants : Henri Nicolas, Jean-Marie Pinel.

L'objectif est ici de modéliser et de modifier les conditions d'éclairage dans les séquences d'images naturelles pour des applications de compression ou d'édition vidéo. Il faut alors estimer les caractéristiques (intensité et position) des sources de lumière (supposées ponctuelles), obtenir une segmentation des zones d'ombres générées par les objets en mouvement, et modéliser les variations temporelles d'éclairage dans

ces zones. Les méthodes d'estimation de la position de la source de lumière et de segmentation des contours des ombres développées précédemment présentent deux limitations : d'une part il est nécessaire de disposer des contours des objets dont on cherche à segmenter l'ombre. En outre, le modèle d'éclairage ne prend pas en compte les zones de pénombres dans lesquelles le modèle d'ombre n'est plus valide. Pour remédier à ces problèmes, une nouvelle approche de détection/segmentation des ombres a été développée. Cette approche se base sur l'utilisation d'une image de mosaïque. Les zones d'ombre en mouvement sont alors détectées par comparaison à l'image de référence et au modèle d'ombre. De plus, le modèle d'ombre a été amélioré de manière à détecter les zones de pénombre existantes aux frontières des ombres, ce qui permet une meilleure précision dans la définition de ces frontières. Un algorithme de compression vidéo a été développé dans lequel les ombres en mouvement sont tout d'abord détectées puis éliminées des images originales. L'information supplémentaire à coder et liée à l'extraction des ombres concerne, d'une part, les contours des ombres, et d'autre part, les paramètres correspondant au modèle d'illumination. L'algorithme a aussi été appliqué à l'édition (ajout, suppression d'objets) de séquences pour accroître le réalisme de la séquence de synthèse.

6.1.5. Structuration par analyse des trajectoires des objets vidéo

Participants : Mireya Garcia-Vasquez, Henri Nicolas.

Il existe de nombreux algorithmes d'estimation des mouvements de translation, des zooms ainsi que des rotations autour d'un axe perpendiculaire à l'axe optique. Cependant, l'estimation des paramètres des mouvements de rotations autour d'axes parallèles au plan image est plus difficile. Dans ce contexte, une méthode d'estimation en trois phases a été développée. Tout d'abord, la validité d'un modèle de mouvement de type affine est évaluée entre deux instants clés consécutifs. Si la qualité de la compensation de mouvement est jugée suffisante, le modèle plan est considéré comme valide. Sinon, une recherche du mouvement non-plan est effectuée. La deuxième phase d'analyse consiste alors à déterminer la direction de l'axe de rotation de l'objet. Enfin, dans un troisième temps, l'angle de rotation non-plan de l'objet est estimé. Cette méthode a été validée expérimentalement sur plusieurs séquences synthétiques et naturelles comprenant des objets en rotation non-plans rigides (des voitures), ou semi-rigides (des visages). Les résultats obtenus ont permis d'obtenir une estimation relativement précise et stable dans le temps du mouvement non-plan. Néanmoins cette méthode nécessite la connaissance des contours (avec une précision raisonnable) des objets vidéo aux instants clés. Le modèle de représentation des mouvements de rotation non-plans a été testé dans un schéma de compression vidéo basé objet. Les performances de ce modèle, en terme de qualité de compensation de mouvement, ont été comparées à celles des modèles affines. Les résultats mettent en évidence un gain en terme de qualité de la compensation du mouvement, ceci au prix d'un coût supplémentaire de codage des paramètres de mouvement relativement faible.

6.2. Codage et codage conjoint source-canal

Mots clés : *compression, codage et décodage robustes, contrôle de congestion, descriptions multiples, scalabilité, transformations en ondelettes, multi-ondelettes, transformations par bases de fonctions redondantes, quantification, théorie débit-distorsion, régulation, progressivité, codes correcteurs, modélisation stochastique, estimation Bayésienne, inférence probabiliste, propagation de croyance, principe turbo, modèles de canal, Internet, mobilité, qualité de service.*

L'objectif de cet axe de recherche est l'optimisation de la qualité de service sur réseaux hétérogènes tels que l'Internet filaire ou sans fils. Nos travaux ont ainsi porté sur le contrôle de congestion et de débit des flux vidéo en nous concentrant sur les transmissions multi-points. Pour répondre à ces besoins de contrôle de congestion sur réseaux hétérogènes et non stationnaires, nous avons développé un algorithme de codage à bas débit et scalable à grain fin de séquences vidéo. L'approche est basée sur une représentation multi-résolution spatio-temporelle compensée en mouvement. Plusieurs approches de codage robuste à la fois à des pertes de paquets et à des perturbations binaires ont par ailleurs été étudiées. La résistance au bruit de transmission nécessite souvent de réintroduire une corrélation ou redondance dosée dans le train binaire de la source comprimée. La question posée est sous quelle forme ? Nous avons développé plusieurs solutions complémentaires : la

première repose sur une décomposition du signal sur des bases de fonctions redondantes (transformations par blocs et bancs de filtres sur-échantillonnés) qui s'apparentent à des codes correcteurs sur les réels. La seconde approche concerne le décodage robuste de codes statistiques largement utilisés dans les systèmes de compression. Une modélisation sous forme d'automate stochastique de codeurs arithmétiques et quasi-arithmétiques nous a ainsi conduit à des solutions de codage et de décodage robuste de ces codes. Enfin, une nouvelle famille de codes sources robustes a été introduite et protégée par dépôt de brevet. Celle-ci ouvre de nombreuses perspectives dans le domaine des communications multimédia sur des réseaux sans fils.

6.2.1. *Contrôle de congestion multipoint et régulation de débit*

Participants : Christine Guillemot, Reda Hosny, Jérôme Viéron.

Dans un contexte de transmission multi-point, un contrôle de congestion (ou de débit) réactif et précis peut exiger un nombre élevé de niveaux de scalabilité, aboutissant à des performances de compression moindres. Ce problème peut être contourné en adaptant dynamiquement le débit de chaque niveau de scalabilité sous la contrainte d'un débit global constant. Dans notre approche, cette adaptation dynamique du débit est réalisée par une optimisation débit-distorsion conjointe des paramètres de quantification et des modes de codage, associée à un modèle de régulation hybride émetteur-récepteur. Le modèle de régulation hybride est associé à un mécanisme de classification (ou de *clustering*) développé par le projet PLANETE. La classification repose sur des métriques relatives aux taux de pertes ou aux délais perçus par les différents récepteurs de la liaison multi-point et permet de mettre en place une prédiction de bande passante par groupe de récepteurs. Le modèle de prédiction de bande passante repose sur des principes dits de « TCP-compatibilité », en d'autres termes sur des équations modélisant le trafic des connexions TCP, afin d'obtenir un partage équitable de la bande passante du réseau entre les données multimédia et les applications téléinformatiques. Cette approche permet d'adapter, de manière dynamique et avec une granularité suffisamment fine, le débit reçu par chacun des récepteurs aux conditions de transmission qu'il perçoit. L'approche permet ainsi une utilisation optimale de la bande passante pour l'ensemble des récepteurs de la session multi-point.

6.2.2. *Codage scalable à grain fin*

Participants : Christine Guillemot, Stéphane Pateux, Jérôme Viéron.

L'optimisation de la qualité de service d'une application de communication vidéo passe par l'adaptation dynamique du débit des signaux compressés à la bande passante du réseau de transmission. Les réseaux de transmission à qualité de service non garantie sont en effet non stationnaires et caractérisés par des variations dans le temps de la bande passante, du taux de pertes, et des délais. Dans un contexte d'application de consultation où les contenus sont pré-encodés et mis à disposition sur un serveur, cette adaptation nécessite une *représentation scalable à grain fin* des données. Il doit en effet être possible d'adapter le nombre de niveaux de scalabilité et le débit par niveau en venant sélectionner des portions du train binaire pré-encodé et stocké sur le serveur, et ceci de manière dynamique en fonction des variations des conditions de transmission. Pour répondre à cet objectif, nous avons développé un algorithme de codage basé sur une décomposition par ondelettes spatio-temporelles de la séquence compensée en mouvement. La décomposition multi-résolution est suivie d'un codage arithmétique par plans de bits tenant compte d'une information contextuelle pour mieux exploiter les dépendences spatio-temporelles. L'estimation et la compensation de mouvement reposent sur une approche hiérarchique basée sur une décomposition en *quadtree* contrainte en débit. Le schéma ainsi obtenu permet alors d'obtenir un niveau de performance comparable au schéma de codage H26L (TML version 6.0), tout en offrant une granularité fine de représentation.

6.2.3. *Transformations redondantes pour le codage conjoint source-canal*

Participants : Christine Guillemot, Slavica Marinkovic, Gagan Rath.

Les décompositions sur bases de fonctions redondantes (en anglais *overcomplete frame expansions*) ont été récemment introduites comme représentations redondantes d'un signal pour accroître sa robustesse aux effacements (ou pertes de paquets) ou à l'ajout de bruit impulsif. Nous avons étudié une classe de fonctions construites à partir de codes DFT (qui peuvent être vus comme des codes de canal sur les réels). À la différence

des codes correcteurs sur les corps finis, les mots de code sont quantifiés, cette opération introduisant un bruit additionnel. La difficulté consiste donc à localiser et à estimer le bruit introduit par le canal de transmission en présence de celui de quantification vu comme un bruit de fond. Considérant des approches de décodage par syndrome, nous nous sommes dans un premier temps intéressés à modéliser l'erreur de reconstruction en présence de bruit de quantification (lorsque les mots de codes sont quantifiés) pour différents modèles de pertes et d'erreur sur le canal de transmission. À la différence avec les codes correcteurs sur les corps finis, la performance des codes sur les réels quantifiés dépend du nombre et de la localisation des pertes (ou des erreurs) dans le mot de code. Nous avons montré l'équivalence en terme d'erreur de reconstruction entre tous les codes BCH issus de codes DFT et donc, par la suite, considéré uniquement des codes passe-bas. L'analyse de performance ayant montré que ces codes sont plus sensibles aux rafales de pertes que à des pertes aléatoires, nous avons développé une approche de décodage récursif basé syndrome permettant de réduire l'erreur de reconstruction lors de rafales de pertes. Cette étude a été étendue au cas des bancs de filtres sur-échantillonnés, lesquelles permettent de traiter un plus grand nombre de *patterns* de pertes.

Ces décompositions sur bases de fonctions redondantes peuvent être vues comme des codes joints source-canal. Ces codes joints peuvent être utilisés sur des canaux caractérisés par des erreurs dont on ne connaît pas la localisation. Cependant la fiabilité de l'étape de localisation des erreurs dépend fortement du bruit de quantification. Lorsque les mots de code sont quantifiés, le problème de la localisation des erreurs est analogue à un problème d'estimation de directions d'arrivées d'ondes planes incidentes sur une antenne linéaire equirépartie. Cette observation nous a conduit à adapter les techniques d'estimation DOA pour la localisation d'erreurs. Nous avons aussi étendu l'approche au cas où nous avons à faire face à la fois à des pertes et à des erreurs, ce qui est le cas des réseaux sans fils.

6.2.4. Codes sources robustes au bruit de transmission

Participants : Christine Guillemot, Hervé Jégou.

Tout système de compression de signaux multimédia (image, vidéo, audio, parole) fait appel à des codes statistiques aussi appelés codes à longueurs variables qui permettent d'obtenir des débits approchant l'entropie du signal. Les codes les plus utilisés dans les systèmes existants (en particulier dans les standards) sont les codes de Huffman. Plus récemment on a vu un regain d'intérêt pour les codes arithmétiques en raison de leurs performances accrues en terme de compression. Ils permettent en effet de découpler le processus d'encodage du modèle supposé de la source. Ceci permet aisément d'utiliser des modèles statistiques d'ordre supérieur. Néanmoins ces codes sont très sensibles au bruit de transmission : toute erreur dans le train binaire peut engendrer une désynchronisation du décodeur et donc une propagation des erreurs sur la suite des informations décodées.

Pour pallier ce problème de propagation les standards de premières générations (H.261, H.263, MPEG-1, MPEG-2) ont incorporé dans le train binaire des marqueurs de synchronisation. Cela conduit à structurer le train binaire en paquets délimités par ces marqueurs de synchronisation et permet de confiner la propagation des erreurs au sein du paquet. Cependant si une erreur intervient en début de paquet la suite du paquet peut être perdue. En outre, la périodicité de ces marqueurs de synchronisation doit être restreinte pour éviter une perte trop grande en efficacité de compression. Les nouveaux standards (H.263+ et MPEG-4) ont alors fait appel à des codes à longueurs variables réversibles (RVLC). Si une erreur s'est produite en milieu de paquet, la symétrie des codes permet de confiner la propagation des erreurs sur un segment en milieu de paquet au lieu de la propager jusque la fin du paquet délimité par un marqueur de synchronisation. Cependant, la symétrie du code engendre une perte en efficacité de compression par rapport à un code d'Huffman de l'ordre de 10 %. En outre ils n'évitent pas complètement le problème de propagation des erreurs.

Nous avons conçu une nouvelle famille de codes sources performants en compression, c'est à dire permettant d'approcher l'entropie du signal, tout en étant très robustes au bruit de transmission. Ces codes, appelés *codes multiplexés*, permettent de confiner le phénomène de désynchronisation aux données moins prioritaires, tout en permettant d'atteindre asymptotiquement l'entropie pour l'ensemble des données. L'idée

consiste à créer des codes à longueur fixe pour l'information de plus grande importance¹ et d'utiliser la redondance inhérente aux codes pour décrire l'information de moindre importance, d'où le nom de codes « multiplexés ». La figure 8 illustre les performances de ces codes par rapport à celles obtenues avec des codes à longueurs fixes et avec des codes d'Huffman. Ces codes ont fait l'objet d'un dépôt de brevet.



Figure 8. PSNR et qualité visuelle obtenus respectivement avec codes à longueurs fixes, codes d'Huffman et codes multiplexés. Les taux d'erreurs bits sont 0.005 (en haut) et 0.05 (en bas).

6.2.5. Décodage conjoint source-canal de codes d'Huffman

Participants : Brendan Catherine, Christine Guillemot, Pierre Siohan, Claudio Weidmann.

Les codes d'Huffman largement utilisés dans les standards de compression sont très sensibles au bruit, une erreur pouvant entraîner une dé-synchronisation du décodeur. Une modélisation sous forme d'automates stochastiques des codeurs à longueurs variables a permis, durant l'année 2001, de construire un algorithme de décodage robuste de codes à longueurs variables (codes d'Huffman). Ce travail avait été mené en collaboration avec le projet SIGMA2. Le décodage basé sur une estimation Bayésienne permet de réduire fortement les taux

¹comme les basses fréquences si une transformation en ondelettes est utilisée

d'erreurs symboles et d'accroître la capacité de resynchronisation du décodeur. Les algorithmes ont d'abord été validés avec des sources *théoriques* de Gauss-Markov (voir rapport d'activités 2001).

Ces travaux ont été poursuivis en 2002 en s'intéressant à des sources réelles issues d'un codeur vidéo. Lors d'une première phase, nous avons analysé les propriétés de distance des codes VLC de type Huffman, ce qui a confirmé le fait que, pour des dimensions de code données, tout gain de codage dépend essentiellement d'une meilleure exploitation de la statistique de source. Le décodage nécessite une information *a priori* sur les statistiques de la source. De nombreux travaux en décodage conjoint source-canal font l'hypothèse d'une connaissance parfaite de la distribution de la source. Cette hypothèse étant peu réaliste en pratique, nous avons élaboré une méthode d'estimation paramétrique de la statistique de la source à partir du train binaire bruité (reçu par le décodeur). Par ailleurs, afin de limiter la complexité de décodage, nous avons conçu une méthode d'élagage du treillis de décodage de source. Ces méthodes d'estimation et d'élagage, évaluées dans un premier temps dans le cas de sources théoriques, vont à présent être testées et adaptées pour prendre en compte le cas de grandes tables de codes VLC utilisées pour le décodage des vecteurs de mouvement issus d'un codeur vidéo.

6.2.6. Décodage conjoint source-canal de codes arithmétiques

Participants : Christine Guillemot, Thomas Guionnet.

Les codes arithmétiques, comme indiqué plus haut, tendent à occuper une place de plus en plus importante dans les schémas de compression émergents, incluant les standards JPEG-2000, H.264 et MPEG-4. Couplés avec des modèles statistiques de source d'ordres supérieurs, ils permettent d'atteindre des facteurs de compression nettement supérieurs à ceux obtenus avec les codes d'Huffman. La contrepartie de ces facteurs de compression élevés est une sensibilité accrue au bruit de transmission. Il est donc essentiel de concevoir des solutions de décodage robuste pour de tels codes.

Tout comme les codes d'Huffman, les codes arithmétiques peuvent être représentés par un arbre. Cependant, les codes arithmétiques posent un certain nombre de problèmes spécifiques. La première difficulté est liée au fait que certaines transitions dans l'arbre n'engendrent pas l'émission de bits. En outre le codage d'un symbole dépend de tous les symboles passés. Par conséquent, la taille de l'arbre croît de manière exponentielle avec la séquence de symboles codés.

Décodage séquentiel : Nous avons développé une première solution basée sur un décodage séquentiel avec une sortie *souple*, couplée à un mécanisme d'élagage de l'arbre afin de contrôler la complexité de l'algorithme de décodage. Les modèles développés sont aussi bien adaptés à l'introduction de mécanismes de synchronisation souples, tels que ceux que nous avons étudié (voir rapport d'activités 2001) dans le cas des codes d'Huffman. Cet algorithme de décodage a été intégré dans un décodeur d'images fixes reposant sur le standard JPEG2000. La figure 9 illustre le gain en performance apporté par cette solution par rapport au décodage instantané utilisé habituellement.

Codes quasi-arithmétiques : La technique de décodage séquentiel reste néanmoins sous-optimale, du fait de l'élagage nécessaire pour limiter la complexité du décodage. Nous nous sommes alors intéressés à l'utilisation de codes arithmétiques à précision réduite, dits *codes quasi-arithmétiques*, modélisables sous forme d'automates à nombre d'états fini. Une modélisation appropriée de l'ensemble source et codeur nous a permis de définir un algorithme de décodage souple optimal, pour une perte minimale en compression. Cette modélisation permet l'utilisation de mécanismes de synchronisation souples, ainsi que l'association avec un code convolutif dans un schéma itératif similaire aux turbo codes en série. Cette approche surpasse le décodage souple de codes d'Huffman et le décodage séquentiel de codes arithmétiques pour des sources faiblement corrélées et des taux d'erreurs importants.

6.3. Tatouage d'images

Mots clés : *tatouage, signature, copyright, étalement de spectre, codage canal.*

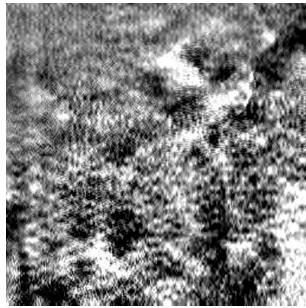
Nos travaux ont porté sur la conception d'un schéma de tatouage basé sur l'utilisation de techniques d'étalement de spectre et de codage avec information adjacente. Le gain de l'information adjacente a par ailleurs été étudié sous deux aspects : le décodage de symbole enfoui et la détection d'une présence de tatouage.



(a) Original.



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



(c) JPEG-2000 coded, AWGN channel ($E_b/N_0 = 5$ dB), PSNR = 11.25 dB.



(d) JPEG-2000 avec décodeur arithmétique robuste, canal AWGN ($E_b/N_0 = 5$ dB), PSNR = 33.26 dB.

Figure 9. Impact des erreurs de transmission sur la qualité visuelle (0.5 bpp).

6.3.1. Schéma d'insertion et d'extraction de marques

Participants : Gaetan Le Guelvouit, Stéphane Pateux.

En considérant une gamme d'attaques de type ajouts de bruit et filtrages pour des signaux modélisés par des mixtures de Gaussienne, nous avons développé un schéma de tatouage basé sur l'utilisation de techniques d'étalement de spectre et de codage avec information adjacente. L'information adjacente permet de limiter l'interférence avec le signal hôte. Les différents paramètres intervenant dans ce schéma de tatouage (force d'insertion, stratégie de l'attaquant) ont alors été optimisés au travers de la modélisation du problème sous la forme d'un jeu entre un attaquant et un défenseur. Une expression analytique des différents paramètres du schéma de tatouage a été obtenue permettant également d'exprimer les performances du système en fonction des distortions d'insertion et de dégradations subies et du nombre de bits insérés. Cette résolution fait apparaître la dualité entre le problème du tatouage et la transmission d'informations sur un canal Gaussien bruité avec information adjacente. Le schéma de tatouage développé a été appliqué au cas du tatouage d'images fixes, et des expérimentations ont été menées afin de valider l'approche proposée face à des manipulations usuelles d'images. Les travaux actuels portent sur l'application de ces résultats au tatouage de signaux vidéo ainsi que sur l'extension du modèle d'attaques, en modélisant les phénomènes de désynchronisation.

Table 1. Résultats expérimentaux obtenus suite à diverses attaques réalisées par Stirmark. L'image tatouée est l'image *Lena* (512×512). Les mesures dans cette table correspondent aux rapports signaux à bruit du canal de transmission.

6.3.2. Tatouage avec information adjacente et sécurité

Participant : Teddy Furon.

Les schémas connus de tatouage avec information adjacente reposent sur l'hypothèse que l'incrustation et le décodage connaissent un dictionnaire de signaux de référence (ou constellation). Il est connu que si l'attaque est un ajout de bruit blanc Gaussien, alors cette solution est optimale au sens où la capacité du canal ne dépend que des données $\{d_t, d_a\}$. La puissance du vecteur \mathbf{V} n'intervient pas. C'est 'comme si' l'interférence entre le signal de tatouage et le contenu originale était annulée. Cependant, ces schémas ne sont guère exploitables en pratique. La taille du dictionnaire commun est prohibitive. Aucune robustesse contre un bruit multiplicatif n'est garantie. Il s'avère que pour un fort niveau d'attaque, ces schémas ne concurrencent pas la méthode traditionnelle de l'étalement de spectre type CDMA. Les travaux en cours visent donc l'invention de schémas mixtes où les régions critiques associées à un symbole sont nettement plus compactes que dans les schémas usuels.

Une classification des attaques est à l'étude. Elle induit une méthodologie visant à définir et quantifier le niveau de sécurité d'une technique de tatouage. Nous nous inspirons de la cryptographie pour cette étude. En effet, même si cryptographie et tatouage travaillent sur des objets radicalement différents (*e.g.*, des nombres et des signaux), au moins la notion de sécurité face à un adversaire intelligent est commune.

7. Contrats industriels

7.1. Contrats industriels bi-latéraux

7.1.1. Convention Celar

Participants : Christine Guillemot, Gagan Rath.

- Numéro de la convention : 1 99 C 459 00 31324 061 et 1 00 C 0302 00 31324 06 1.
- Intitulé : *Liaison vidéo numérique en robotique mobile*.
- Activité de recherche concernée : § 6.2.5.
- Partenaire : Inria, Celar.

- Financement : Celar.
- Période : 18 mois à partir de septembre 2000.

Le problème posé consiste à transmettre un signal vidéo sur un réseau sans fil à partir d'un équipement d'acquisition mobile, pour une application de téléguidage d'un robot terrestre mobile. Le réseau support est ainsi caractérisé par une faible bande passante et une qualité de service non garantie qui peut en outre varier dans le temps. Afin de permettre une utilisation optimale de la bande passante disponible à un instant donné et variant dans le temps, la source vidéo émettrice doit prendre en compte les caractéristiques du canal. Le codeur doit donc émettre un débit relativement élevé mais robuste aux erreurs de transmission. Nous avons mis en place dans un codeur vidéo H263+ des outils de scalabilité et de dissimulation d'erreurs.

7.2. Contrats nationaux

7.2.1. RNRT-V2net

Participant : Raphaèle Balter, Luce Morin.

- Numéro de la convention : 2 00 C 0339.
- Intitulé : Projet RNRT-V2NET, *Visite virtuelle en réseau*.
- Activité de recherche concernée : § 6.1.1.
- Partenaires : Thomson Multimédia, Inria, France Télécom R&D.
- Financement : RNRT.
- Période : 36 mois à partir de janvier 2000.

L'objectif du projet V2NET est de développer et intégrer de nouveaux outils pour une navigation optimale dans des mondes 3D réels et virtuels dans un contexte d'applications client-serveur. L'optimalité est définie ici comme la recherche d'un meilleur compromis qualité d'image/fluidité de navigation. Les études portent plus précisément sur la modélisation et la représentation de scènes 3D complexes réelles et virtuelles, la structuration de la scène et les protocoles applicatifs, la visualisation de scènes 3D complexes réelles et virtuelles, et la mise-en-place d'une plate-forme d'expérimentation permettant de valider les approches proposées. Nos contributions ont porté sur des outils de représentation de scènes par séquences de modèles 3D.

7.2.2. RNRT-Cohrainte

Participant : Stéphane Pateux.

- Numéro de la convention : 299C3220031324011.
- Intitulé : Projet RNRT-Cohrainte, *Codage hiérarchique et robuste de sources audiovisuelles et application à l'Internet*.
- Activité de recherche concernée : § 6.1.3.
- Partenaires : France Télécom R&D, Philips-Lep, Irisa/Inria-Rennes, ENST-Paris.
- Financement : RNRT.
- Période : 36 mois à partir de mars 1999.

L'objectif du projet Cohrainte est de développer des solutions de codage de flux vidéo et audio avec une granularité fine afin de permettre une adaptation des flux lors de la transmission sur des réseaux à qualité de service variable. Cette année les travaux ont porté sur le développement d'un schéma de codage vidéo basé sur une représentation par maillages actifs et ondelettes $2D + t$.

7.2.3. RNRT-VIP

Participant : Christine Guillemot.

- Numéro de la convention : 2 01 A 0650 00 000MC 01 1
- Intitulé : Visiophonie sur IP sans fils

- Activité de recherche concernée : § 6.2.3 et § 6.2.6.
- Partenaires : Comsys, ENSEA, France Télécom R&D, Irista/Inria-Rennes, INRIA-Sophia, ENST-Paris, Université Paris-6, Thalès Communication.
- Financement : RNRT.
- Période : 30 mois à partir de novembre 2001.

L'objectif du projet est de concevoir de nouvelles solutions de codage robuste et de codage conjoint source-canal pour la transmission vidéo sur des réseaux IP sans fils. Nos contributions ont porté sur l'étude de transformations redondantes (ou codes sur les réels) et l'étude de solutions de décodage robuste de codes arithmétiques.

7.2.4. **RNRT-DIPHONET**

Participants : Gaëtan Le Guelvouit, Stéphane Pateux.

- Numéro de la convention : 2 01 A 0704 00 000MC 01 1
- Intitulé : Diffusion de Photographies à travers l'Internet
- Activité de recherche concernée : § 6.3.1.
- Partenaires : Canon, CNRS (L2S), INRIA, Andia Press
- Financement : RNRT.
- Période : 30 mois à partir de janvier 2002.

L'objectif du projet Diphonet est de proposer des outils de protection et de traçage de photos pour la diffusion sur Internet de photos professionnelles. Pour se faire en coopération avec des techniques d'indexation, l'utilisation de technique de tatouage est proposée afin d'insérer dans une image des informations de copyright, mais aussi des informations permettant de retrouver l'origine d'une photo. Cette première année, une étude bibliographique sur les différentes techniques de tatouage a été menée ainsi qu'une évaluation des techniques de tatouage existantes.

7.2.5. **Projet RNTL DOMUS-VIDEUM**

Participants : Anne Manoury, Henri Nicolas.

- Numéro de la convention : 2 02 C 0100 00 00 MPR 01 1
- Intitulé : *Sélection, indexation, visualisation avancée et recherche par le contenu des documents vidéo.*
- Activité de recherche concernée : § 6.1.5.
- Partenaires : Inria, Thomson Multimedia, Ecole polytechnique de Nantes, INA, SFRS.
- Financement : projet RNTL.
- Période : 30 mois à partir de novembre 2001.

L'objectif du projet DOMUS VIDEUM consiste à développer des outils pour la sélection, l'indexation, la visualisation avancée et la recherche par le contenu des documents vidéo pour les plate-formes multimedia domestiques avec apprentissage du profil utilisateur. La participation du projet TEMICS consiste à développer des outils permettant, à partir d'une décomposition en plan-séquence d'une vidéo, de regrouper les plans-séquences non-consécutifs mais représentant le même contenu, ce qui permet d'obtenir une décomposition en hyperplans de la vidéo originale. Les approches en cours de développement se base sur la construction d'images mosaïques pour chaque plan. De manière à permettre une construction rapide d'images mosaïques, les seules informations de mouvement utilisées sont les vecteurs de mouvement disponibles dans le flux compressé MPEG.

7.2.6. **RNRT-EIRE**

Participants : Christine Guillemot, Henri Nicolas, Jérôme Vieron.

- Numéro de la convention : 2 02 C00990031324011

- Intitulé : EIRE (Etudes d'optimisations algorithmiques de JPEG 2000).
- Activité de recherche concernée : § 6.2.2.
- Partenaires : CRIL Technology, ENSTA, I3S, Inria, IRCOM, Thalès.
- Financement : projet RNRT.
- Période : 27 mois à partir de novembre 2001.

L'objectif du projet EIRE est de développer et d'optimiser des solutions de codage d'images fixes compatibles JPEG2000 et vidéo. Notre contribution porte sur la définition d'un algorithme scalable basé sur une décomposition en sous-bandes spatio-temporelles compensée en mouvement suivie d'un codage arithmétique par plan de bits inspiré de l'algorithme EBCOT.

7.2.7. RNRT-Cosocati

Participants : Pierre Siohan, Claudio Weidmann.

- Numéro de la convention : pas encore affecté.
- Intitulé : *Codage conjoint Source-canal pour la Transmission d'Images*.
- Activité de recherche concernée : § 6.2.5
- Partenaires : France Télécom R&D, Thalès (en remplacement d'Alcatel Space Industries), Irisa/Inria-Rennes CNES, GET, I3S.
- Financement : RNRT.
- Période : 36 mois à partir de mars 2000.

Le projet exploratoire Cosocati concerne l'application de techniques de codage et décodage conjoint source-canal pour de la transmission d'images dans le cas de systèmes de télécommunications utilisant des canaux de type radio-mobile, et dans celui de systèmes d'observation spatiale par satellite. Par rapport à des systèmes de communications classiques, on espère tirer parti d'avantages potentiels du codage conjoint : un gain de complexité de calcul ou de matériel pour un niveau de performances donné, ou encore une zone de fonctionnement nominale étendue. Le rôle plus précis de Temics concerne l'étude de méthodes de décodage conjoint source-canal. Nous examinons en particulier différentes techniques d'estimation de la statistique de source après transmission dans un canal bruité.

7.3. Contrats européens

7.3.1. IST-Busman

Participants : Teddy Furon, Christine Guillemot, Stéphane Pateux.

- Numéro de la convention : 1 02 C 0186 00 00MPR 00 5.
- Intitulé : Bringing user satisfaction to media access networks.
- Activité de recherche concernée : § 6.3.1.
- Partenaires : British Telecom, Frametool, HHI, INRIA, Motorola, QMUL, Telefonica, University of Munich.
- Période : 30 mois à partir de avril 2002.
- Financement : CEE.

Le projet porte sur le développement de solutions d'indexation et de marquage des contenus vidéo pour la création de services multimédia à valeur ajoutée. Nos contributions portent sur les solutions de tatouage qui vont être utilisées pour insérer des méta-données (création de *smart-vidéos*) dans un flux avec différents niveaux de robustesse, de sécurité et de capacité. Un premier prototype d'insertion de données dans un signal vidéo a été développé. Une réflexion a aussi été menée pour réaliser un système de tatouage joint symétrique/asymétrique.

7.3.2. IST-Song

Participants : Bilal ElAli, Christine Guillemot.

- Numéro de la convention : 1 99 A 1821 00 00MP R00 5.
- Intitulé : IST-SONG, *Portals of Next Generation*.
- Activité de recherche concernée : § 5.1.
- Partenaires : France Télécom, DT-Berkom, Blaxxun, Csel, Enst-br, Epfl, GMD, Inria, RUS, Telenor, Univ. Genève.
- Période : 34 mois à partir de janvier 2000.
- Financement : CEE.

L'objectif du projet est de développer de nouvelles technologies d'interfaces hommes-machines (avatars, agents intelligents, etc.), de nouveaux algorithmes de représentation et de communication de contenus audiovisuels et de mondes virtuels sur Internet et sur des liens sans fils, pour des « *portails* » de seconde génération. En collaboration étroite avec le CSELT, France Télécom et RUS, TEMICS a contribué à la définition de stratégies de transport de données multimédia sur Internet, à l'étude d'algorithmes de contrôle de congestion et de codage conjoint source canal sur l'Internet « *best-effort* » et dans un cadre de différenciation de services.

7.3.3. IST-Ozone

Participants : Gabriele Davini, Christine Guillemot, Henri Nicolas.

- Numéro de la convention : 1 01 A0672 00 000MC 00 5.
- Intitulé : New technologies and services for emerging nomadic societies.
- Activité de recherche concernée : § 5.1.
- Partenaires : Epictoid, IMEC, INRIA, Philips, Thomson Multimedia, Université d'Eindhoven.
- Période : 30 mois à partir de septembre 2001.
- Financement : CEE.

Le projet IST-Ozone a comme objectif le développement d'une infrastructure de de communication, capable d'offrir aux individus des informations et des services divers ; l'accès aux ressources sera le plus vaste possible (*anywhere and anytime*). Notre contribution au projet IST-Ozone s'inscrit, en particulier, dans le contexte de la transmission de flux audiovisuels, dont les performances sont essentielles pour garantir aux usagers une qualité de service. En particulier, on a étudié et développé les modules de contrôle de congestion et de contrôle des pertes, avec les objectifs suivants :

- éviter les congestions sur les réseaux de communication ;
- garantir la fluidité des transmissions, en respectant les contraintes " temps-réel " des flux audiovisuels ;
- minimiser l'effet des pertes (et des délais) sur les réseaux de communication ;
- garantir un accès équitable aux ressources disponibles pour plusieurs flux concurrents.

Dans ce contexte, nous avons contribué à la conception de l'architecture de la plate-forme Ozone, à celle d'un des trois démonstrateurs, ainsi qu'à la rédaction des spécifications correspondantes.

8. Actions régionales, nationales et internationales

8.1. Réseaux et groupes de travail internationaux bilatéraux

Le projet Temics participe également à plusieurs collaborations bi-latérales :

- programme France-Maroc (participants : C. Labit, L. Morin, S. Pateux) : cette action intégrée franco-marocaine (responsable français : L. Morin) est menée en collaboration avec l'université de Rabat, Maroc (Professeur D. Aboutajdine, professeur assistant de l'INPT-Rabat A. Tamtaoui) ;

- programme France-Mexique (participants : L. Morin, H. Nicolas, S. Pateux, C. Guillemot) : cette collaboration est menée avec l'Uonam-Mexico (professeur Garcia-Ugalde) dans le cadre d'une convention CNRS-Conacyt ;
- participation au colloque France-Taiwan (participants : C. Guillemot), Sophia-Antipolis, mars 2002.
- collaboration avec R. Ansari de l'Université d'Illinois à Chicago (UIC) et avec P. Duhamel du CNRS sur le thème du codage conjoint source-canal ; il s'agit d'une ARC IRISA dans le cadre de laquelle nous accueillons S. Marinkovic sur un statut postdoctoral ;
- collaboration avec l'Université Catholique de Louvain (UCL) ; il s'agit d'une collaboration avec B. Macq sur le thème du tatouage, dans le cadre de laquelle nous avons accueilli T. Furon sur un statut postdoctoral jusque septembre 2002.
- collaboration avec le Computer Science Department of the University College of Dublin. Ce programme d'échange appelé *Ulysse* est financé par le CNRS et Enterprise Ireland. Deux *lecturers*, Neil Hurley et Guenole Silvestre, ont été accueillis pendant une semaine en septembre. La collaboration porte sur le développement de nouveaux schémas de communication avec information adjacente pour le tatouage.

8.2. Actions nationales

8.2.1. ARC Télégéo

Participants : Raphaèle Balter, Luce Morin.

- Intitulé : Télégéo : Géométrie et Télécommunications.
- Activité de recherche concernée : § 6.1.1.
- Partenaires : Creatis-Insa de Lyon, ENST Paris, INRIA (ISA, TEMICS, PRISME).
- Financement : INRIA.
- Période : un an à partir de juin 2002.

L'objectif de cette ARC est de créer une synergie dans le domaine de la transmission en réseaux d'objets géométriques et d'étudier des techniques de représentation d'objets géométriques pour leur transmission sur des réseaux hétérogènes. La contribution de l'équipe TEMICS porte d'une part sur des méthodes de compression de maillages surfaciques non structurés, et permet d'autre part de partager des connaissances sur les approches de compression progressive/scalable et la prise en compte de critères de rendu visuel.

8.2.2. Action CNRS tatouage

Participant : Stéphane Pateux.

- Intitulé : Tatouage et dissimulation de données pour les communications audiovisuelles.
- Activité de recherche concernée : § 6.3.1.
- Partenaires : UTC, INPG, ENST, Supelec, LIFL.
- Financement : CNRS.
- Période : janvier-décembre 2002.

Cette action spécifique a pour but de permettre à différents acteurs académiques nationaux de mettre en commun des compétences complémentaires, afin d'assurer une prospective de recherche sur le thème de la protection des données audiovisuelles numériques par des techniques de tatouage (marquage de la propriété de leurs ayants droit). L'étude de différents protocoles de sécurisation possibles à base de technique de tatouage sont étudiées dans ce cadre. Les partenaires de l'action s'intéressent également aux méthodes plus générales de dissimulations de données, permettant l'authentification des documents audiovisuels, leur traçage, et de vérifier leur intégrité.

8.3. Visites et invitations de chercheurs

- Visite du professeur Ravi Motwani, IIT, Indian Institute of Technology, Kanpur, Inde, pour trois mois à compter de Mai 2002. La collaboration avec le professeur Motwani a porté sur le thème du codage conjoint source-canal et en particulier sur la conception de transformations redondantes transmission vidéo robuste sur des réseaux à qualité de service non garantie.

9. Diffusion des résultats

9.1. Dépôt de brevets

- Hervé Jégou et Christine Guillemot, "compression de données numériques robuste au bruit de transmission", brevet français No. 02 09287 déposé le 22 Juillet 2002.

9.2. Animation de la communauté scientifique

- T. Furon a été *co-chair* d'une session sur le thème du tatouage et de la sécurité à EUSIPCO'02.
- C. Guillemot est éditeur associé du journal IEEE transactions on Image Processing ;
- C. Guillemot est membre du comité de programme des conférences Spie-VCIP 2002, IEEE-ICIP 2002, ACM Multimedia 2002, CORESA 2002 et *chair* de session à IEEE-ICIP 2002 ;
- C. Guillemot est membre du comité international IEEE IMDSP (Image and MultiDimensional Signal Processing Technical Committee) ;
- C. Guillemot est membre de la commission 2 du RNRT ;
- C. Guillemot est nommée, à partir de septembre 2002, suppléante du PCP (président du comité des projets) à la commission d'évaluation de l'INRIA ;
- C. Guillemot est membre du comité de pilotage du RTP "Réseaux ambiants" ;
- L. Morin et S. Pateux ont participé activement (par le biais de présentations) aux réunions du GDR Isis ;
- H. Nicolas est membre nommé (suppléant) de la commission de spécialistes en 27ième section ;
- H. Nicolas est membre du comité de programme technique de la conférence ICIP 2002 ;
- S. Pateux est coordinateur scientifique au niveau de l'IRISA du CNRT (Centre National de Recherche Technologique) TIM-Bretagne ;
- S.Pateux est co-responsable avec F. Davoine de l'UTC pour la rédaction d'un ouvrage collectif sur le tatouage de signaux multimédia dans la collection traité IC2 de Hermès.

9.3. Enseignement

- Diic-LSI, Ifsic, université de Rennes 1 (H. Nicolas : codage) ;
- Diic-INC, Ifsic, université de Rennes 1 (L. Morin, C. Labit, H. Nicolas, F. Garcia Ugalde, C. Guillemot : traitement d'images, vision 3D, mouvement, codage, compression d'images, cryptographie, communication) ;
- DEA Stir, option Image, université de Rennes 1 (C. Labit, H. Nicolas : codage) ;
- DEA Informatique, université de Rennes 1 (C. Guillemot, S. Pateux : Compression vidéo et communication) ;
- DEA TIS, université de Cergy-Ensea (C. Guillemot : codage) ;
- Enic, Villeneuve-d'Ascq (C. Guillemot : Vidéo sur IP) ;
- Ensar Rennes, option informatique (L. Morin : fondements du traitement d'image, morphologie mathématique) ;
- ENSTBR (C. Guillemot : Les normes de codage vidéo de MPEG-1 à MPEG-4) ;
- DEA ÉLECTRONIQUE, INSA, université de Rennes 1 (P. Siohan : Télévision numérique) ;

- Diic-INC, Ifsic, Dess-Mitic, université de Rennes 1 (S. Pateux : Tatouage de documents multimédia) ;
- ESIGETEL Fontainebleau, (S. Pateux : Tatouage de signaux multimédias ; C. Guillemot : Compression et communication vidéo) ;

10. Bibliographie

Bibliographie de référence

- [1] C. GUILLEMOT, T. TURLETTI. *Systèmes multimédia communicants*. Éditions Hermès, 2000, chapitre Algorithmes de codage et de contrôle de transmission sur réseaux hétérogènes, pages 25-82.
- [2] A. GUYADER, E. FABRE, C. GUILLEMOT, M. ROBERT. *Joint Source-Channel Turbo Decoding of Entropy-Coded Sources*. in « IEEE Journal on Selected Areas in Communications, special issue on the turbo principle : from theory to practice », numéro 9, volume 19, sept., 2001, pages 1680-1696.
- [3] J. P. LEDUC, J. M. ODOBEZ, C. LABIT. *Adaptive motion-compensated wavelet filtering for image sequence coding*. in « IEEE Transactions on Image Processing », numéro 6, volume 6, June, 1997, pages 862-879.
- [4] R. MOHR, L. MORIN, E. GROSSO. *Relative positioning with uncalibrated cameras*. in « in : Geometric Invariance in Computer Vision, Editors : A. Zisserman J.L. Mundy, MIT Press », 1992, pages 440-460.
- [5] H. NICOLAS, C. LABIT. *Motion and illumination variation estimation using a hierarchy of models : application to image sequence coding*. in « Journal of Visual Communication and Image Representation », numéro 4, volume 6, December, 1995, pages 303-316.
- [6] P. RAULT, C. GUILLEMOT. *Symmetric delay factorization : genralized framework for paraunitary filter banks*. in « IEEE Trans. on signal processing », numéro 12, volume 47, Dec., 1999, pages 3315-3325.
- [7] F. MOREAU DE SAINT-MARTIN, P. SIOHAN, A. COHEN. « *Biorthogonal filterbanks and energy preservation in image compression* ». in « IEEE Transactions on Image Processing », numéro 2, volume 8, février, 1999, pages 168-178.

Thèses et habilitations à diriger des recherche

- [8] F. GALPIN. *Représentation 3D de séquences vidéo ; Schéma d'extraction automatique d'un flux de modèles 3D, applications à la compression et à la réalité virtuelle*. thèse de doctorat, Thèse de doctorat en Informatique, Université de Rennes 1, France, 1, 2002, <http://www.irisa.fr/bibli/publi/theses/2002/galpin/galpin.html>.
- [9] X. HÉNOCQ. *Contrôle d'erreur pour transmission de flux vidéo temps réel sur réseaux de paquets hétérogènes et variant dans le temps*. thèse de doctorat, Université de Rennes 1, Mai, 2002.
- [10] JEAN-MARIE PINEL. *Etude des conditions d'éclaircissement dans une séquence d'images et application à la composition et au codage vidéo*. thèse de doctorat, Thèse de l'université de Rennes I, 28 novembre, 2002.

Articles et chapitres de livre

- [11] J. BENOIS-PINEAU, H. NICOLAS. *A new method for region-based depth ordering in a video sequence : application to frame interpolation.* in « Journal of Visual Communications and Image Representation », numéro 3, volume 13, September, 2002, pages 363-385.
- [12] T. FURON, P. DUHAMEL. *An asymmetric watermarking method.* in « Accepté pour publication dans IEEE Transactions on Signal Processing, special issue on "Signal Processing for Data Hiding in Digital Media and Secure Content Delivery" », 2002.
- [13] F. GALPIN, L. MORIN. *Sliding adjustment for 3D video representation.* in « EURASIP Journal on applied Signal Processing - Special issue on 3D Imaging and Virtual Reality », à paraître.
- [14] L. LUCAT, P. SIOHAN, D. BARBA. *Adaptive and global optimization methods for weighted median filters.* in « Signal Processing : Image communication », volume 17, 2002, pages 509-524.
- [15] L. OISEL, É. MÉMIN, L. MORIN, F. GALPIN. *1D dense disparity estimation for 3D reconstruction.* in « IEEE transactions on Image Processing », à paraître..
- [16] G. RATH, C. GUILLEMOT. *Frame-theoretic analysis of DFT codes with erasures.* in « Accepté pour publication dans IEEE Transactions on Signal Processing », 2002.
- [17] G. RATH, C. GUILLEMOT. *Performance analysis and recursive syndrome decoding of DFT codes for bursty erasure recovery.* in « Accepté pour publication dans IEEE Transactions on Signal Processing », 2002.
- [18] P. SIOHAN, C. SICLET, N. LACAÏLLE. *Analysis and design of OFDM/OQAM systems based on filterbank theory.* in « IEEE Transactions on Signal Processing », numéro 5, volume 50, mai, 2002, pages 1170-1183.
- [19] J. VIÉRON, C. GUILLEMOT. *Real-time constrained TCP-compatible rate control for video over the Internet.* in « Accepté pour publication dans IEEE Transactions on Multimedia », 2002.

Communications à des congrès, colloques, etc.

- [20] M. CHAUMONT, S. PATEUX, H. NICOLAS. *Segmentation of non-rigid video objects using long-term temporal consistency.* in « Proceedings of ICIP, Rochester, USA », September, 2002.
- [21] D. DEUFF, I. CHANCLOU, H. NICOLAS, M. MAZIÈRE, M. MAURIZOT. *Building Interactive Videos : a simple authoring tool for teachers.* in « Proceedings of SCI, Orlando, USA », July, 2002.
- [22] F. GALPIN, L. MORIN, K. DEGUCHI. *Compression performance of computer vision based coding.* in « Proceedings of the IAPR Workshop on Machine Vision Applications, MVA'2002, Nara, Japan », décembre, 2002.
- [23] F. GALPIN, L. MORIN, K. DEGUCHI. *La vision par ordinateur pour la compression vidéo.* in « JST'2002, Journées pour la Sciences et la Technologie, Tokyo, Japan », novembre, 2002.

- [24] M. GARCIA, H. NICOLAS. *Video object trajectory analysis*. in « Proceedings of ICIP, Rochester, USA », September, 2002.
- [25] G. L. GUELVUIT, S. PATEUX, C. GUILLEMOT. *Information-theoretic resolution of perceptual watermarking of non i.i.d. Gaussian signals*. in « Proceedings of the European Signal Processing Conference, EUSIPCO'02 », Septembre, 2002.
- [26] G. L. GUELVUIT, S. PATEUX, C. GUILLEMOT. *Perceptual watermarking of non I.I.D. signals based on wide spread spectrum using side information*. in « Proceedings of the IEEE Int. conference on Image Processing, ICIP'02 », Septembre, 2002.
- [27] T. GUIONNET, C. GUILLEMOT, E. FABRE. *Soft decoding of multiple descriptions*. in « Proceedings of the Int. Conf. on Multimedia and Expo, ICME'02 », Août, 2002.
- [28] M. JEANNE, P. SIOHAN, J. CARLACH. *Source and joint source-channel decoding of variable length codes*. in « Proc. International Conference on Communications (ICC) », New-York, EU, avril, 2002.
- [29] H. NICOLAS, F. DENOVAL. *Interactive Modifications of Video Object Trajectories in Natural Video Sequences for post-production applications*. in « Proceedings of ICIP, Rochester, USA », September, 2002.
- [30] H. NICOLAS, F. DENOVAL. *Interactive Object-based Video Editing for Post-production*. in « Proceedings of SCI, Orlando, USA », July, 2002.
- [31] D. PINCHON, P. SIOHAN, C. SICLET. *A fast design method for biorthogonal modulated filter banks*. in « Proc. of 11th European Signal Processing Conference », Toulouse, France, septembre, 2002.
- [32] D. PINCHON, P. SIOHAN, C. SICLET. *A fast design method for orthogonal modulated filter banks*. in « Proc. International Conference on Acoustics, Speech and Signal Processing », Orlando, Florida, mai, 2002.
- [33] J. PINEL, H. NICOLAS. *Shadows Analysis and Synthesis in Natural Video Sequences*. in « Proceedings of ICIP, Rochester, USA », September, 2002.
- [34] G. RATH, C. GUILLEMOT. *Error localization performance analysis of DFT codes in presence of quantization error*. in « Proceedings of the IEEE Int. Information Theory Workshop, ITW'02 », Octobre, 2002.
- [35] G. RATH, C. GUILLEMOT. *On the optimal performance of quantized DFT codes with erasures*. in « Proceedings of the IEEE Int. Symposium on Information Theory, ISIT'02 », Juillet, 2002.
- [36] G. RATH, C. GUILLEMOT. *Recursive syndrome decoding of DFT codes for bursty erasures*. in « Proceedings of the IEEE Int. Conference on Speech and Signal Processing, ICASSP'02 », Mai, 2002.
- [37] G. RATH, C. GUILLEMOT. *Syndrome decoding and performance analysis of DFT codes with bursty erasures*. in « Proceedings of the Int. Conference on Data Compression, DCC'02 », Mars, 2002.
- [38] C. SICLET, P. SIOHAN, D. PINCHON. *Oversampled orthogonal and biorthogonal multicarrier modulations with perfect reconstruction*. in « Proc. International Conference on Digital Signal Processing », Santorini,

Greece, juillet, 2002.

- [39] J. VIÉRON, C. GUILLEMOT, S. PATEUX. *Motion-compensated 2D+t wavelet representation for FGS video coding*. in « Proceedings of the Tyrrenian Int. workshop on digital communications, IWDC'02 », Septembre, 2002.
- [40] J. VIÉRON, T. TURLETTI, X. HÉNOCCQ, C. GUILLEMOT, K. SALAMATIAN. *TCP-compatible rate control for FGS layered multicast video transmission based on a clustering algorithm*. in « Proceedings of the IEEE Int. Symposium on Circuits and Systems, ISCAS'02 », Mai, 2002.
- [41] C. WEIDMANN. *New Upper Bounds on the Expected Length of One-to-One Codes*. in « Proc. IEEE Int. Symp. Information Theory », pages 179, Lausanne, Switzerland, June, 2002.
- [42] C. WEIDMANN. *Variable-Length Codes getting Framed*. in « Proc. 2nd Asian-European workshop on Information theory », pages 104-107, Breisach, Germany, June, 2002.