

# *Projet TEMICS*

*Traitement, Modélisation d'Images et CommunicationS*

*Rennes*

THÈME 3B



*R*apport  
*d'Activité*

2000



## Table des matières

<b>1</b>	<b>Composition de l'équipe</b>	<b>3</b>
<b>2</b>	<b>Présentation et objectifs généraux</b>	<b>4</b>
<b>3</b>	<b>Fondements scientifiques</b>	<b>6</b>
3.1	Modélisation 3D par géométrie projective . . . . .	6
3.2	Estimation de mouvement et segmentation . . . . .	7
3.3	Approximation par maillages hiérarchiques et éléments finis . . . . .	8
3.4	Théorie débit-distorsion et formalisme MDL . . . . .	9
3.5	Représentation temps-fréquence et par bases de fonctions redondantes . . . . .	12
3.6	Tatouage d'images . . . . .	13
<b>4</b>	<b>Domaines d'applications</b>	<b>15</b>
4.1	Panorama . . . . .	15
4.2	Compression d'images fixes et animées . . . . .	15
4.3	Communication multimédia . . . . .	16
4.4	Production et post-production vidéo . . . . .	16
<b>5</b>	<b>Logiciels</b>	<b>17</b>
5.1	Logiciel de segmentation et de suivi interactif d'objets vidéo . . . . .	17
5.2	Logiciel de post-production vidéo . . . . .	17
<b>6</b>	<b>Résultats nouveaux</b>	<b>19</b>
6.1	Analyse, modélisation et édition de séquences d'images . . . . .	19
6.1.1	Segmentation et suivi temporel d'objets vidéo . . . . .	19
6.1.2	Modélisation 3D par analyse de séquences vidéo monoculaires . . . . .	20
6.1.3	Représentation par maillages hiérarchiques adaptatifs . . . . .	21
6.1.4	Création d'objets mosaïques 2D . . . . .	22
6.1.5	Édition vidéo basée objet . . . . .	25
6.2	Codage et codage conjoint source-canal . . . . .	26
6.2.1	Codage orienté régions, modèles 3D et textures . . . . .	26
6.2.2	Codage robuste et scalable, et régulation de débit associée . . . . .	27
6.2.3	Allocation de débit conjointe source/canal pour transmission vidéo sur Internet . . . . .	29
6.2.4	Décodage conjoint source-canal . . . . .	29
6.2.5	Codage conjoint source-canal . . . . .	30
6.2.6	Régulation de débit pour la transmission d'images satellitaires sur canal à débit fixe . . . . .	31
6.3	Tatouage d'images . . . . .	31

<b>7</b>	<b>Contrats industriels (nationaux, européens et internationaux)</b>	<b>32</b>
7.1	Contrats industriels bi-latéraux . . . . .	32
7.1.1	CTI-FT RD . . . . .	32
7.1.2	Convention Celar . . . . .	33
7.1.3	Convention Microprocess . . . . .	34
7.1.4	Convention Thomson-TCO . . . . .	34
7.1.5	Convention Cnes . . . . .	35
7.1.6	Convention Thomson Multimédia . . . . .	35
7.2	Contrats nationaux . . . . .	36
7.2.1	RNRT-V2net . . . . .	36
7.2.2	RNRT-Osiam . . . . .	36
7.2.3	RNRT-Cohrainte . . . . .	37
7.2.4	RNRT-Visi . . . . .	37
7.3	Contrats européens . . . . .	38
7.3.1	Esprit-Nemesis . . . . .	38
7.3.2	IST-Openise . . . . .	38
7.3.3	IST-Song . . . . .	39
<b>8</b>	<b>Actions régionales, nationales et internationales</b>	<b>39</b>
8.1	Réseaux et groupes de travail internationaux . . . . .	39
8.2	Réseaux et groupes de travail bilatéraux . . . . .	40
8.3	Visites et invitations de chercheurs . . . . .	40
<b>9</b>	<b>Diffusion de résultats</b>	<b>40</b>
9.1	Animation de la communauté scientifique . . . . .	40
9.2	Enseignement . . . . .	41
<b>10</b>	<b>Bibliographie</b>	<b>41</b>

---

# 1 Composition de l'équipe

## Responsable scientifique

Christine Guillemot [DR Inria]

## Assistante de projet

Edith Blin-Guyot [TR Inria]

## Personnel Inria

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

Henri Nicolas [CR]

Stéphane Pateux [CR, détachement du corps des Télécoms]

## Personnel CNRS

Laurent Amsaleg [CR]

Thierry Fauconnier [IR, (Atelier)]

## Personnel Université

Luce Morin [maître de conférences, université Rennes 1]

## Collaborateurs Extérieurs

Francisco Garcia-Ugalde [PAST]

Philippe Robert [PAST, jusqu'au 30 septembre 2000]

## Ingénieurs experts

Bilal El Ali [IE Inria, depuis le 1er février 2000]

Lionel Oisel [IE Inria, jusqu'au 31 mars 2000]

## Chercheurs post-doctorants

Zhikui Chen

Gagan Rath [post-doctorat à partir du 15 avril 2000]

**Ingénieurs associés**

Philippe Coval [à partir du 1er octobre 2000]

Jeanne Guillou [à partir du 1er octobre 2000]

**Chercheurs doctorants**

Dubhe Chavira-Martinez [Bourse CONACYT]

Marc Chaumont [Bourse MENRT, à partir du 1er octobre 2000]

Franck Denoual [Bourse CIFRE]

Franck Galpin [Bourse INRIA-RÉGION]

Mireya Garcia-Vasquez [Bourse CONACYT]

Thomas Guionnet [Bourse INRIA]

Xavier Henocq [Bourse INRIA]

Reda Hosny [Bourse MAE]

Gaetan Le Guelvouit [Bourse INRIA-RÉGION, à partir du 1er octobre 2000]

Delphine Le Guen [Bourse CNES, jusqu'au 31 décembre 2000]

Fabrice Le Léanec [Bourse MENRT, jusqu'au 31 décembre 2000]

Gwenaëlle Marquant [Bourse MENRT, jusqu'au 31 décembre 2000]

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;
- le mélange réaliste de données visuelles 2D ou 3D, naturelles ou synthétiques ;
- 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, modélisation et édition 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 vidéo est nécessaire. Cette phase consiste à segmenter les objets vidéo 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 (déformations géométriques ou changements photométriques). 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 soient bien adaptés à une description multi-grille ou 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 ou de ses attributs 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 robuste, de décodage robuste et conjoint source-canal mettant en œuvre le principe turbo (inspiré des turbo-codes). Nos travaux concernent également le problème de l'allocation de débit optimale entre le codage de source et le codage de canal, les techniques de codage scalable pour une adaptation aisée du débit de la source vidéo à la bande passante du réseau, et les techniques de codage par bases redondantes et par descriptions multiples.
- *Tatouage.* L'apparition de réseaux ouverts comme l'Internet suscite de nouveaux dangers pour la propriété intellectuelle conduisant à la réticence des possesseurs de contenus envers l'Internet, ou envers d'autres environnements ouverts de distribution. Grâce - ou à cause - des technologies numériques, il est possible de faire des copies parfaites des contenus, la redistribution est ainsi devenue aisée et le traçage des redistributions difficile. Nous nous intéressons principalement au tatouage de séquences vidéo à des fins de protection de copie et d'authentification. Nous privilégions le tatouage multi-niveau, la détection et l'extraction robuste de marque exploitant les résultats du projet en codage conjoint source-canal.

Les domaines d'application de nos travaux sont l'imagerie satellitaire au travers de ses besoins en terme de compression, les applications multimédia en réseaux (Internet et réseaux sans fils), la production audiovisuelle. Dans ces secteurs d'applications, nous avons des conventions de recherche ou de transfert avec France Télécom, le Celar, le Cnes, Microprocess, des projets multi-partenaires nationaux (projets RNRT COHRAINTE, OSIAM, V2NET, VISI) et européens (projets ESPRIT-NEMESIS, IST-SONG et IST-OPENISE). Le projet cherche également à valoriser les résultats de ses travaux au travers de contributions à la normalisation (IETF, MPEG).

### 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.

**Résumé :** *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 :

$$P_i = K(R_i | t_i). \quad (2)$$

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

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

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  $\theta$  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

$$p_1^t.F.p_i = 0. \quad (3)$$

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 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 Estimation de mouvement et segmentation

**Mots clés :** modèles affines, modèle de caméra, prédiction.

**Résumé :** *L'estimation du mouvement de la caméra, de régions ou d'objets au sein de séquences vidéo naturelles, ainsi que le calcul des segmentations correspondantes constituent des aspects essentiels sur lesquels s'appuient plusieurs axes de recherche du projet. La connaissance du mouvement permet en effet d'analyser le contenu dynamique d'une scène, ce qui facilite les manipulations du contenu objet des séquences ou sa reconstruction 3D. L'estimation de mouvement est un*

*maillon très important dans les schémas de compression vidéo à base de prédiction temporelle, et dans les algorithmes de segmentation et de modélisation 3D.*

Le mouvement apparent directement visible dans les images représente le mouvement relatif entre l'objet et la caméra. L'estimation du mouvement propre des objets nécessite une compensation du mouvement de la caméra qui peut être réalisée en estimant le mouvement apparent d'une zone fixe dans l'image. L'estimation du mouvement peut être réalisée par pixel ou par région. Dans ce dernier cas, il est nécessaire de procéder à une segmentation des images considérées. TEMICS privilégie les approches par régions. Elles nécessitent l'utilisation d'une modélisation du mouvement 2D qui s'appuie souvent sur le modèle affine suivant, bien que des modèles plus simples (le modèle translationnel), ou des modèles plus complexes (le modèle homographique), puissent également être utilisés :

$$\begin{aligned} dx &= t_x + k.(x - x_g) - \theta(y - y_g) + h_1(x - x_g) + h_2(y - y_g), \\ dy &= t_y + k.(y - y_g) + \theta(x - x_g) + h_1(y - y_g) + h_2(x - x_g), \end{aligned} \quad (4)$$

où  $t_x$  et  $t_y$  représentent les termes translationnels, et  $k$  le facteur de divergence. Le terme  $\theta$  est l'angle de rotation plan, et  $h_1$  et  $h_2$  sont les termes hyperboliques. Il est important de noter que la vidéo ne restituant qu'une projection bidimensionnelle du mouvement tridimensionnel des objets, ce modèle ne permet d'obtenir qu'une représentation approximative du mouvement apparent 2D. L'estimation de ces paramètres de mouvement est utile non seulement pour l'analyse du mouvement proprement dit, mais aussi pour le recalage des images dans un référentiel de représentation commun dans le cadre de la construction d'images mosaïques 2D, ou de la reconstruction 3D d'une scène. Le type de régions utilisé varie selon l'application visée. Classiquement, les méthodes de compression vidéo utilisent des segmentations en régions carrées, alors que des techniques d'édition vidéo basée objet ont besoin des contours précis des objets vidéo. Des formats de segmentation intermédiaires tels que les régions de forme polygonale ou les maillages triangulaires sont également utilisés en compression vidéo ou en analyse de scènes.

### 3.3 Approximation par maillages hiérarchiques et éléments finis

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

**Résumé :** *L'approximation par maillage et éléments finis constitue un outil méthodologique permettant une représentation souple et adaptative d'une image à des fins d'analyse ou de compression. Une image fixe peut être approximée par morceaux. Une séquence vidéo pourra être représentée par une succession de déformations appliquées à une représentation par maillage initiale.*

La méthode des éléments finis (MEF) peut être considérée comme une méthode générale de discrétisation et d'approximation de problèmes continus. Un milieu continu est subdivisé en un nombre fini d'éléments dont le comportement est régi par un nombre fini de paramètres. De

façon pratique, la méthode des éléments finis est basée sur l'utilisation d'un maillage approprié du système à modéliser (les mailles usuelles étant généralement triangulaires ou quadrilatères).

Bien que destinée initialement à la modélisation et à l'analyse de problèmes physiques continus, une utilisation de la méthode des éléments finis dans le cadre du codage vidéo se justifie pour plusieurs raisons :

- la MEF offre une représentation souple et adaptative, elle se prête ainsi parfaitement au caractère non stationnaire du signal vidéo en espace et en temps, en alternant zones de luminance, de mouvement, uniformes et/ou perturbées,
- la MEF permet une analyse numérique robuste et, par analogie avec la physique, permet une large compréhension des problèmes et des solutions envisageables.

Le type de modèle utilisé sur les mailles peut être de nature différente suivant la complexité ou bien encore le type de continuité désirée (éléments finis de Lagrange réalisant une interpolation affine par maille et de continuité  $C^0$ , éléments finis de Hermitte réalisant une interpolation par des polynômes de degré 3 par maille et de continuité  $C^1$ ). L'interpolation de Lagrange pour un maillage triangulaire peut être définie par la donnée d'un ensemble de valeurs définies sur les nœuds du maillage, et l'interpolation de ces valeurs de façon affine sur les différentes mailles. Ainsi, si l'on considère un triangle  $(P_1 P_2 P_3)$ , avec des valeurs aux nœuds  $(V_1, V_2, V_3)$ , l'interpolation affine pour tout point de ce triangle s'exprime comme :

$$V(P) = w_1(P).P_1 + w_2(P).P_2 + w_3(P).P_3$$

avec

$$w_1(P) = \frac{\overrightarrow{P_2 P} \wedge \overrightarrow{P_2 P_3} \cdot \vec{z}}{\overrightarrow{P_2 P_1} \wedge \overrightarrow{P_2 P_3} \cdot \vec{z}},$$

où  $\wedge$  et  $\cdot$  représentent respectivement le produit vectoriel et le produit scalaire de deux vecteurs, et  $\vec{z}$ , le vecteur orthogonal au plan de l'image. Les poids  $w_2(P)$  et  $w_3(P)$  sont obtenus de façon similaire par permutation circulaire sur les indices.

Dans le cadre d'une image fixe, l'image est approximée par morceaux de façon continue sur les différentes mailles (cf. figure 1.a,b). Dans le cas d'une image issue d'une séquence, le maillage est utilisé afin de déformer une image de référence dans le but d'obtenir une image aussi proche que possible de l'image courante (cf. figure 1.c,d). Les différents paramètres contrôlant le maillage sont alors obtenus selon des critères d'optimisation de la reconstruction de l'image désirée.

### 3.4 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.

**Résumé :** *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*

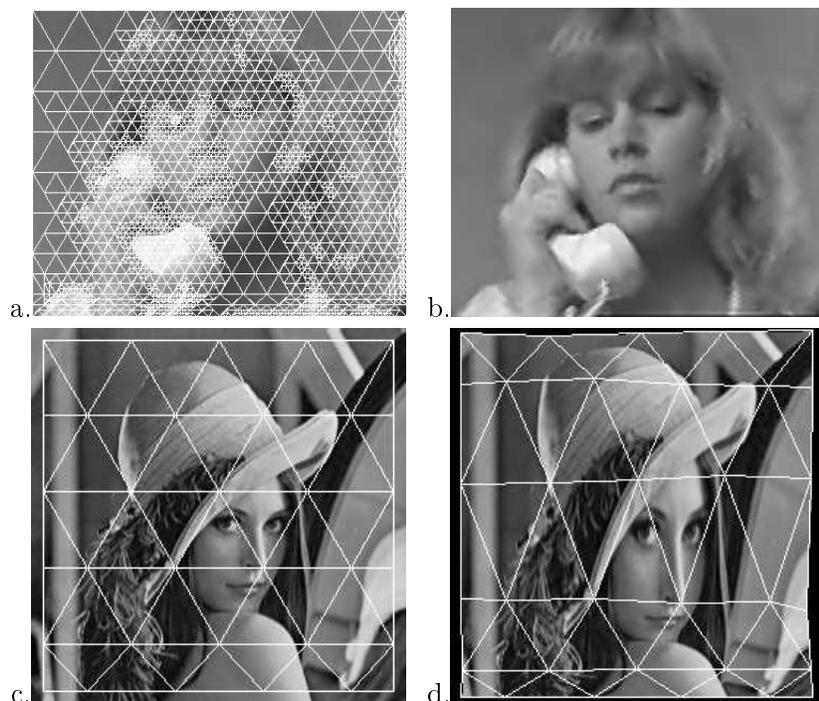


FIG. 1 – Exemples d'utilisation des maillages. (a,b) : exemple d'utilisation d'un maillage pour représenter une image fixe. (c,d) : exemple d'utilisation d'un maillage pour déformer une image.

*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. Elle répond à l'équation suivante :

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

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 contrainte d'un débit global donné. La fonction de coût à minimiser devient alors

$$J = (R_s + R_c) + \lambda(D_s + D_c).$$

Le débit global est alors défini comme la somme du débit de source  $R_s$  et du débit de redondance  $R_c$  introduit par le codeur de canal (ou codes correcteurs). Les techniques d'optimisation débit-distorsion conjointes source-canal permettent d'améliorer le compromis compression - robustesse au bruit de transmission, 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), ou encore par des approches de représentation du signal par descriptions multiples. Le *codage par descriptions multiples* peut être vu comme une généralisation à un environnement de transmission multi-canal du codage avec critère de fidélité.

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  $\theta$  le vecteur des paramètres du modèle de représentation, le principe du formalisme MDL est de minimiser le coût de description de la chaîne  $s$ :

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

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  $\theta$ ;  $-\log_2 P_{\theta}(s) + DL(\theta)$  est le coût de description global de la chaîne  $s$  en utilisant le modèle  $\theta$ . 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 basé sur un découpage en régions.

### 3.5 Représentation temps-fréquence et par bases de fonctions redondantes

**Mots clés :** Transformations par ondelettes, bancs de filtres multi-fréquences, échantillonnage, codes correcteurs, descriptions multiples.

**Résumé :** *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. Dans un contexte de transmission sur des canaux bruités, 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 sur des bases de fonctions redondantes.*

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.

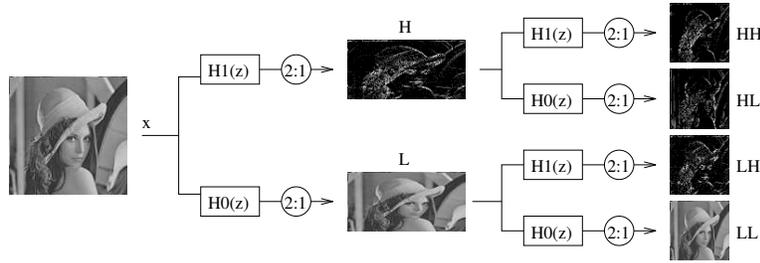


FIG. 2 – Analyse multi-résolution d'un signal 2D par un banc de filtres : le signal est décomposé sur 4 espaces orthogonaux entre eux, un espace d'approximation (correspondant au signal LL contenant les basses fréquences verticales et horizontales) et trois espaces de détail (correspondant aux signaux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

Les décompositions par bancs de filtres multi-fréquences utilisées en compression de signal sont à échantillonnage critique, c'est-à-dire que la somme des échantillons contenus dans les sous-signaux est égale au nombre de pixels contenus dans l'image. Dans un contexte de transmission sur des canaux bruités, il s'avère nécessaire de maintenir une certaine redondance ou corrélation dans la représentation multi-résolution de l'image. L'une des approches considérées dans le projet TEMICS consiste à utiliser des transformations redondantes, comme par exemple des bancs de filtres sur-échantillonnés. Conceptuellement, l'approche s'apparente à un code de canal en bloc, défini sur le corps des réels, et mis en œuvre 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. 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  $m \geq n$ . 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 Bayésiens) pour exploiter la redondance et corriger les erreurs (ou pertes) dues à la transmission.

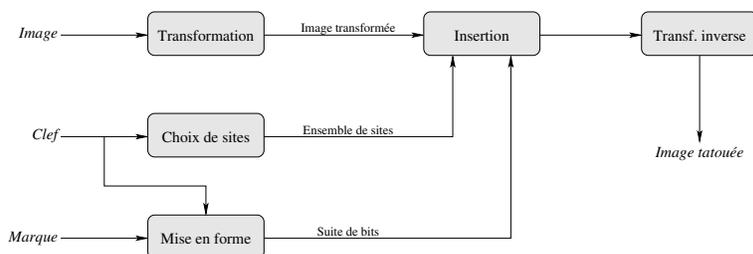
### 3.6 Tatouage d'images

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

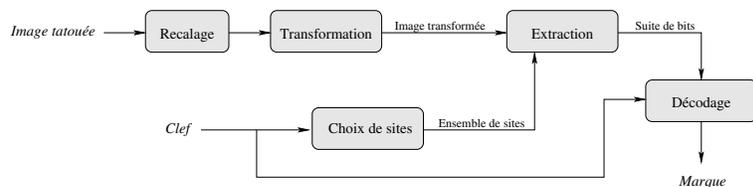
**Résumé :** *L'objectif général du tatouage (ou watermarking en anglais) est de dissimuler des informations dans une image afin de pouvoir identifier son origine, de protéger son copyright, ou bien encore de vérifier son intégrité. Cette marque doit être indécélable, infalsifiable et ineffaçable. Son extraction doit être robuste face à divers types d'attaques que l'image pourrait subir (compression, retaille, filtrage, transformations géométriques, ...). Il est possible d'établir un parallèle entre le tatouage et une situation de communication classique. TEMICS privilégie le marquage*

dans le domaine transformé, qui s'apparente en communication aux techniques de transmission multi-porteuses. Cette approche permet en particulier d'exploiter des fonctions de masquage du système visuel afin de s'assurer que la marque sera non visible.

Les techniques d'insertion de marques peuvent être classées en plusieurs catégories : marquage de sites choisis dans une image (dans le domaine spatial), marquage dans le domaine transformé (par exemple après transformation en cosinus discrète), ou par utilisation de codes pseudo-aléatoires pour répartir la marque dans tout le signal. Le schéma général d'insertion d'une marque dans le domaine transformé et d'extraction de cette marque est illustré par la figure 3.



(a) Schéma du principe d'insertion de la marque.



(b) Schéma du principe d'extraction de la marque.

FIG. 3 – Schéma de principe pour l'insertion et l'extraction de marque dans une image.

Une fois choisi le site d'insertion, il faut concevoir une technique d'insertion de la marque. Il existe plusieurs méthodes d'insertion. Nous nous intéressons principalement aux techniques d'insertion par étalement de spectre, qui offrent une meilleure robustesse aux attaques que peut subir la marque.

Soit un message  $m$  à insérer défini par la suite de bits  $\{b_j, j = 1, \dots, M\}$ , et soit  $\{V_i, i = 1, \dots, N\}$  l'ensemble des valeurs des coefficients correspondant aux différents sites d'insertions. La technique d'étalement de spectre consiste alors, à partir d'un jeu de vecteurs  $\{g^j, j = 1, \dots, M\}$  de dimension  $N$  aléatoires et orthogonaux (par exemple défini à partir de M-séquences),

à insérer le message en modifiant les valeurs originales des coefficients :

$$V'_i = V_i + \alpha * \sum_{j=1..M} b_j \cdot g_i^j.$$

Par la suite, on peut extraire le bit  $b_k$  par corrélation :  $\frac{\sum_{i=1..N} V'_i \cdot g_i^k}{\sum_{i=1..N} (g_i^k)^2}$ , puisque les vecteurs  $g^j$  sont orthogonaux entre eux, et qu'ils sont décorrélés de l'image. Le coefficient  $\alpha$  permet de régler l'amplitude du bruit généré dans l'image par l'insertion de la marque. Ce coefficient doit être choisi suffisamment faible afin que la marque ne soit pas visible, mais aussi suffisamment fort pour que la détection par corrélation fonctionne correctement même en présence d'attaques.

## 4 Domaines d'applications

### 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 d'images satellitaires (en collaboration avec le Cnes), et de la compression scalable du signal vidéo (en collaboration avec FT RD). 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 (quatre projets RNRT) et Européennes (deux projets IST). Le dernier secteur a trait à la production et à la post-production vidéo. Plus précisément, les domaines d'application de nos travaux sont les outils de création de contenus multimédia. Dans ce domaine, nous avons des collaborations avec Thomson Optronique, Microprocess et une collaboration Européenne (projet Esprit Nemesis).

### 4.2 Compression d'images fixes et animées

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.

Les travaux du projet visent à améliorer certains schémas de compression spécifiques qui sont encore peu satisfaisant à l'heure actuelle. En particulier, le projet se penche sur des schémas de compression quasi sans pertes avec prise en compte de modèles basés psychovision, sur des schémas de compression sélective, sur des schémas robustes aux pertes ou aux erreurs de transmission, sur des schémas où le besoin d'interactions avec le contenu vidéo demande de mettre au point des représentations compactes de la structuration d'une séquence en objets ou en régions. D'une façon générale, la qualité des schémas de compression que le projet met au point est accrue par nos travaux théoriques concernant la quantification et du codage entropique aux applications de codage à bas débits (inférieures à 0.5bpp).

Une collaboration avec le Cnes nous permet d'appliquer nos travaux à la compression d'images satellitaires. Dans ce cadre, en plus d'une compression efficace, nous devons maîtriser la nature des pertes et caractériser les dégradations afin de garantir une certaine qualité de reconstruction.

### 4.3 Communication multimédia

Les applications diffusant du contenu multimédia sont notamment 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. Aussi, les travaux du projet évoqués dans le paragraphe précédent s'appliquent naturellement à cet aspect des applications multimédia. Or les outils définis à des seules fins de compression présentent certaines limitations : ils permettent des degrés limités d'interaction avec le contenu de l'information multimédia, 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 codages scalables nous permet d'offrir des techniques élaborées de régulation de flux pour faire du contrôle de pertes et de congestion 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, de même que le cadre de normalisation MPEG-7, constituent des points 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 FT RD et au sein du projet RNRT Visi. 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.

### 4.4 Production et post-production vidéo

Nous pouvons distinguer deux types fondamentaux de manipulations d'une séquence vidéo acquise par une caméra. D'une part, l'édition vidéo pour laquelle on souhaite supprimer des

objets dans la scène ou ajouter de nouveaux objets réels ou synthétiques. D'autre part, la modification artificielle de la position, de la focale ou du mouvement de la caméra permettant de changer *a posteriori* le cadrage de la scène.

Actuellement, le processus de manipulation d'une séquence vidéo est essentiellement manuel. Au stade de la post-production, un opérateur humain installe ses modifications manuellement, image après image, jusqu'à l'obtention de l'effet désiré. Aussi, les travaux du projet visant à mettre au point des outils automatiques ou semi-automatiques pour aider, faciliter et simplifier l'édition vidéo s'appliquent ici directement. En particulier, nos outils cadrent avec les spécifications MPEG-4 relatives à la création d'effets spéciaux et la composition de scène. Nous appliquons le résultat de nos recherches au travers du projet Esprit Nemesis.

Les applications (télé-présence, jeux,...) mettant en œuvre des scènes virtuelles font appel à des outils de création de contenus qui vont permettre de construire des modèles 3D. Ces modèles 3D peuvent être générés par des outils de synthèse d'images ou par analyse de séquences vidéo filmant des scènes naturelles. Pour qu'il soit possible de naviguer de façon fluide dans la scène modélisée, ces modèles doivent avoir des représentations compactes, voire robustes au bruit de transmission, si la navigation doit se faire au travers d'un réseau de communication. 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 s'appliquent donc directement à toute application mettant en œuvre des scènes 3D virtuelles issues de scènes naturelles et mettant en œuvre les fonctions de navigation et de communication associées. Ce domaine constitue le cadre applicatif de nos travaux au sein du projet RNRT-V2net.

## 5 Logiciels

### 5.1 Logiciel de segmentation et de suivi interactif d'objets vidéo

**Participant** : Stéphane Pateux.

**Mots clés** : segmentation, suivi, masque d'objet.

Dans le cadre du projet RNRT Osiam, un logiciel de segmentation interactive d'objets vidéo a été développé. Ce logiciel s'appuie sur un noyau en C++ autonome permettant de réaliser un suivi automatique d'objets vidéo. Une interface graphique en JAVA a été développée sur ce noyau de base afin d'offrir une interactivité dans la phase de segmentation (voir figure 4). Via cette interface, il est possible de définir à l'aide de la souris une segmentation en objet, puis de lancer le programme de suivi et de corriger ensuite les éventuelles dérives pouvant apparaître. Ce logiciel est en cours de dépôt à l'APP.

### 5.2 Logiciel de post-production vidéo

**Participants** : Henri Nicolas, Lionel Oisel.

**Mots clés** : post-production, image panoramique, segmentation statique, segmentation

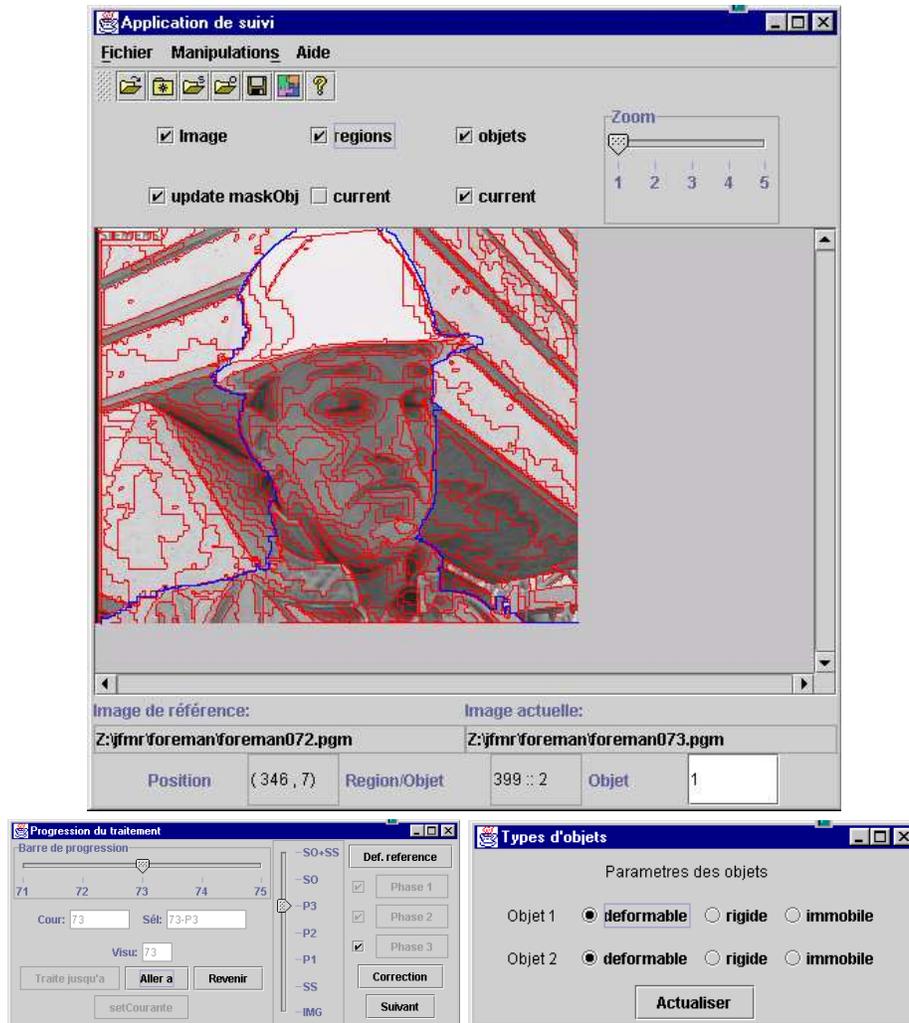


FIG. 4 – Logiciel de segmentation interactive d'objets vidéo. Les contours bleus délimitent les objets définis, les contours rouges les régions spatiales utilisées pour définir les objets. Les différentes fenêtres permettent de contrôler et de lancer le programme de suivi.

dynamique, recalage 3D/2D.

Dans le cadre du projet européen Nemesis (voir section 7) a été développé un logiciel de post-production vidéo s'adressant tant aux professionnels du domaine qu'aux amateurs avertis. Il s'agit en fait d'outils intégrés à un logiciel de post-production déjà largement diffusé, « *Adobe After Effect* ». Outre sa large diffusion, ce dernier présente l'avantage d'offrir une interface dotée de nombreux outils de gestion de séquences vidéo. Les algorithmes développés dans Nemesis se présentent donc sous la forme de fonctions optionnelles (« plug-ins ») qui utilisent l'interface d'« *After Effect* ». Il s'agit d'outils de segmentation statique (laboratoire ICS, Grèce), de segmentation dynamique (laboratoire HHI, Allemagne), de recalage 2D/3D et d'étalonnage interactif (Inria Rocquencourt, France), et de création d'images panoramiques (Irisa-projet Temics).

## 6 Résultats nouveaux

### 6.1 Analyse, modélisation et édition 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.

**Résumé :** *Les travaux de l'année ont porté sur le développement d'une technique de suivi temporel d'objets segmentés. L'algorithme repose sur une projection arrière et basé mouvement d'une segmentation spatiale initiale. Il permet une bonne localisation des frontières des objets et le suivi d'objets déformables. Une représentation par maillages triangulaires hiérarchiques et adaptatifs associée à une représentation progressive du mouvement a conduit au développement d'un codeur vidéo scalable permettant d'adapter le signal vidéo compressé aux capacités du réseau ou des terminaux. Pour répondre à des besoins de navigation dans une scène, au travers d'un réseau de télécommunications, nous avons en outre 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. Une autre spécificité de l'étude repose sur le fait que nous ne disposons pas d'information a priori sur le contenu de la scène, ni sur les paramètres de calibration de la caméra.*

#### 6.1.1 Segmentation et suivi temporel d'objets vidéo

**Participant :** Stéphane Pateux.

Dans le cadre du projet RNRT OSIAM, des techniques de suivi d'objets vidéo ont été développées, permettant de répondre aux besoins d'applications diverses (codage MPEG-4, sur-

veillance, ...). Les objets suivis peuvent être quelconques et déformables. La technique proposée consiste dans un premier temps à effectuer une segmentation spatiale de chaque image. Une des approches possibles de segmentation spatiale consiste à exploiter l'information de couleur [42]. Une projection arrière de la segmentation spatiale obtenue sur un masque de référence permet de trouver à quel objet (défini par le masque de référence) appartient cette région [43]. La projection tient compte du mouvement de la région. Cette technique permet une bonne localisation des frontières des objets suivis, et un traitement aisé des objets déformables qui sont généralement difficiles à suivre. Par ailleurs la complexité de l'algorithme permet d'enviesager des techniques rapides de suivi. Diverses améliorations de cet algorithme de suivi ont été apportées afin de palier les problèmes de segmentation spatiale ou bien encore d'occultations entre objets. Une technique de redécoupage spatial des régions a été développée afin d'identifier et de traiter les zones à problèmes (par exemple où se produisent des occultations). L'algorithme obtenu permet de suivre automatiquement et de façon efficace des objets présents dans une scène et définis préalablement, par segmentation spatiale.

### 6.1.2 Modélisation 3D par analyse de séquences vidéo monoculaires

**Participants :** Franck Galpin, Luce Morin, Philippe Robert.

Les approches de modélisation, basées sur la mise en correspondance de points de vue distants dans les séquences, reposent sur des techniques d'estimation de champs denses de mouvement respectant la contrainte épipolaire. Cette dernière est identifiée par l'extraction de points singuliers, suivie d'une mise en correspondance et d'une estimation robuste de la matrice fondamentale codant la géométrie projective. La scène observée peut alors être approximée par un ensemble de régions planes à partir d'un maillage initial arbitraire, de telle sorte que par exemple les facettes correspondent à un même modèle homographique de mouvement. La scène peut aussi être décrite par une ou plusieurs mosaïques (planaires, cylindriques, sphériques,...) auxquelles peuvent être associées des informations de profondeur.

La reconstruction 3D est très sensible aux erreurs de mesure dans l'image. Nous avons donc dans un premier temps cherché à améliorer l'estimation du champ de disparité. Nous avons comparé deux méthodes, l'une basée sur l'estimation d'un flot optique contraint par la géométrie épipolaire, l'autre basée sur un modèle de mouvement affine par maillage triangulaire (voir section 6.1.3). La première méthode s'avère plus adaptée aux scènes naturelles, la seconde plus adaptée aux scènes de type polyédrique. Nous avons ensuite développé un schéma global de reconstruction-représentation d'une séquence par création d'un ensemble de modèles 3D. 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) [25], [23]. Ce schéma génère un flux séquentiel de modèles VRML pouvant facilement être transmis sur un réseau en vue d'une navigation à distance dans la scène. Le choix des images clés est crucial pour la qualité des modèles reconstruits. Nous avons défini un critère de sélection des images clés basé sur deux types de mesures : d'une part l'amplitude du mouvement apparent entre deux images clés, d'autre part la cohérence entre le mouvement estimé de la caméra et les images initiales. Une interface de navigation a été développée pour tester ce schéma (voir figure 5). Elle permet de rejouer la séquence initiale par enchaînement des modèles ou de spécifier

un trajet virtuel et de construire la séquence correspondante.

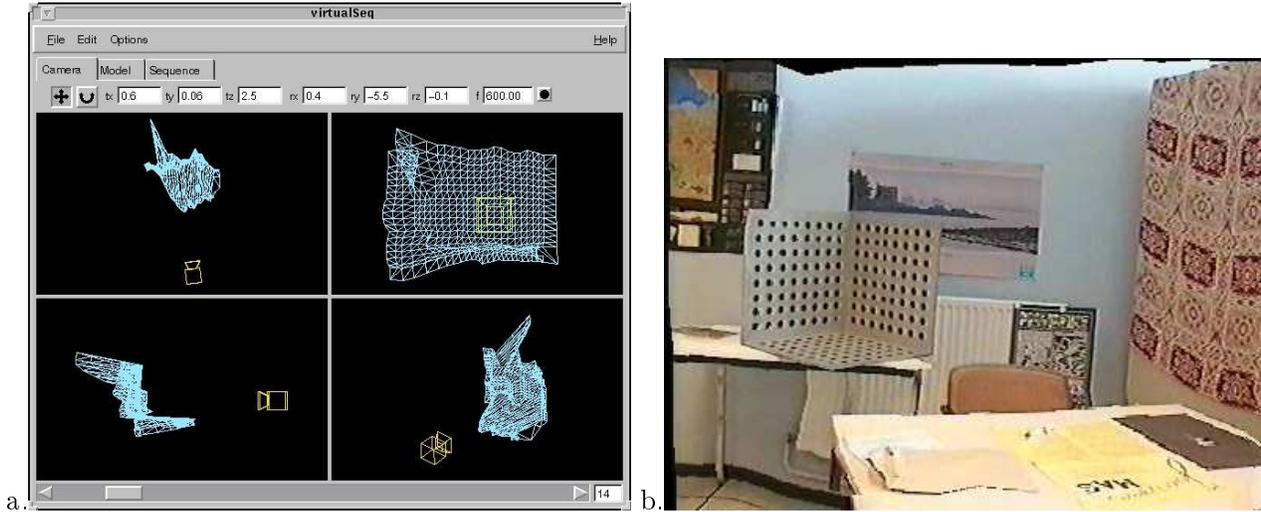


FIG. 5 – Modélisation 3D par analyse de séquences monoculaires: Interface de navigation. (a) vues du modèle 3D et de la caméra virtuelle; (b) image virtuelle générée par la caméra.

### 6.1.3 Représentation par maillages hiérarchiques adaptatifs

**Participants :** Claude Labit, Gwenaëlle Marquant, Stéphane Pateux.

L'objectif de cette étude est de représenter les objets d'une séquence vidéo par un ensemble de maillages hiérarchiques adaptatifs. Le maillage des objets est optimisé en terme de performance débit-distorsion. Cette année nous avons plus particulièrement travaillé sur la représentation par maillages triangulaires scalables pour des applications de codage et de communication à des débits et qualités variables [40]. Dans le prolongement de travaux effectués en mode de codage intra image, nous avons étendu certaines notions au codage inter image. Afin d'estimer le mouvement entre deux images successives d'une séquence, nous avons mis en œuvre une approche combinant les maillages hiérarchiques et les représentations multi-résolutions. Ceci permet, avec un maillage grossier et une image sous échantillonnée, d'estimer d'abord les mouvements dominants présents dans la scène, avant de les raffiner à l'aide de niveaux de hiérarchie plus fins. L'approche que nous avons adoptée optimise le mouvement pour chaque niveau de la hiérarchie, et permet une représentation progressive du mouvement. Ceci permet un rendu « personnalisé » en fonction des disponibilités du réseau ou des capacités des terminaux. De même que pour les images fixes, le critère pour établir l'ordre d'importance des triangles dans la représentation prend en compte à la fois la qualité de représentation en terme de PSNR, mais aussi le coût de codage engendré par une telle division. Ceci garantit d'avoir une représentation optimale pour un coût de codage minimum.

Dans le cadre de l'estimation du mouvement, nous avons développé une méthode « hybride » couplant les avantages des méthodes d'estimation « *forward* » et « *backward* » largement employées. Les résultats obtenus montrent un gain de la qualité par rapport à l'approche classique « *forward* ». Enfin, afin de gérer les problèmes d'occultations dus aux mouvements et aux profondeurs relatives des objets dans les scènes, nous avons effectué un suivi indépendant de chaque objet de la scène (le contexte MPEG-4 nous fournissant les masques des objets). Cette représentation par maillages a été développée en introduisant la notion de ligne de rupture externe [37]. Ceci permet de ne pas chercher à représenter le contour polygonal de l'objet avec le maillage, et par conséquent, pour un coût de codage égal, à positionner plus de nœuds dans les zones de mouvement complexe. Ceci est illustré sur la figure 6.

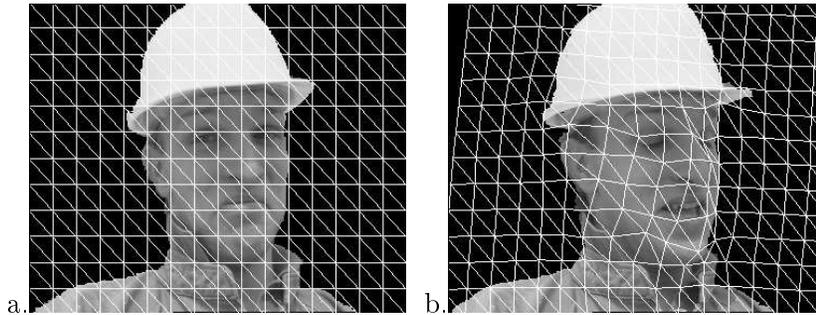


FIG. 6 – Exemple de suivi de maillage pour un objet avec prise en compte des lignes de rupture.

#### 6.1.4 Création d'objets mosaïques 2D

**Participants :** Henri Nicolas, Jean-Marie Pinel.

L'objectif de l'étude est de définir une méthode de création d'images mosaïques à partir d'une séquence vidéo sans connaissance *a priori* du mouvement ou des paramètres de la caméra. Nous avons développé une technique d'estimation de la position d'une image dans le référentiel du panoramique qui ne nécessite pas une estimation du mouvement de la caméra. Cette technique est basée sur l'utilisation de modèles de mouvement de complexité variable (modèle à trois paramètres, modèle affine à six paramètres et modèle homographique). Elle repose également sur des techniques de suivi temporel et sur la mise en correspondance multi-résolution d'images. Elle a permis d'obtenir de très bons résultats sur des séquences complexes ayant un mouvement de caméra rapide (par exemple la séquence « stefan ») et sur des séquences de test représentatives du domaine de la post-production vidéo. Afin de gérer efficacement le problème de recouvrement des images projetées dans le référentiel du panoramique, de tenir compte de leurs résolutions respectives (provenant des variations de la focale de la caméra), des conditions d'éclairage, et des effets éventuels de faux contours au niveau des raccordements entre les images, nous avons défini un critère original permettant d'optimiser efficacement les coefficients de pondération entre la nouvelle image et l'image panoramique [17].

La création d'images mosaïques pose le problème de l'harmonisation des conditions d'éclairage. Une telle harmonisation peut être réalisée si l'on dispose d'information sur les conditions d'éclairage des objets vidéo manipulés dans les séquences d'origine. La méthode d'analyse des conditions d'illumination en cours de développement se décompose en deux phases principales:

1. *Détection de la position 2D d'une source lumineuse.* En présence d'une (ou éventuellement de plusieurs) source de lumière localisée, les conditions d'éclairage d'un objet sont fortement liées non seulement à l'intensité de la source lumineuse, mais aussi à sa position par rapport à l'objet. La méthode d'estimation de la position de la source de lumière s'appuie sur une étude des contours respectifs de l'objet et de l'ombre correspondante à un ou plusieurs instants clés. Il est montré [44] qu'il est possible de localiser la position 2D (c'est-à-dire la projection de la position réelle dans le plan image) de la source de lumière avec une précision dépendant de la forme de l'objet, de son ombre ainsi que de son mouvement entre les instants considérés.
2. *Détection des ombres.* La détection des contours de l'ombre du ou des objets considérés à un instant donné est réalisée en deux temps. Tout d'abord, la zone dans laquelle l'ombre est susceptible de se situer est estimée grâce à une approche combinant une méthode de suivi temporel de la position de l'ombre et l'utilisation d'une contrainte physique liée à la position estimée de la source lumineuse. Dans un deuxième temps, les contours réels de l'ombre sont obtenus en analysant les variations d'illumination entre les deux instants considérés dans la zone de vraisemblance précédemment estimée.

Ce problème de compensation et d'harmonisation d'illumination lors de création d'images mosaïques est illustré sur la figure 7



FIG. 7 – *Compensation d'illumination d'une mosaïque obtenue à partir d'images fixes.*



FIG. 8 – Exemples d'insertion et de suppression d'objets dans une séquence avec gestion de l'ombre.

Une version semi-automatique de la méthode de création d'images mosaïques a également été développée dans le but de garantir (sous réserve de validité des modèles de mouvement utilisés) le haut niveau de qualité requis par les applications de post-production. Cette technique a de plus été optimisée en temps de calcul et implantée au sein du logiciel de post-production du projet européen NEMESIS.

### 6.1.5 Édition vidéo basée objet

**Participants :** Laurent Amsaleg, Franck Denoual, Mireya Garcia-Vasquez, Henri Nicolas.

L'objectif de cette étude est de développer des méthodes semi-automatiques d'édition de scènes à partir de plusieurs objets vidéo ou mosaïques. Dans le domaine de la post-production vidéo, cette édition est souvent réalisée manuellement.

**Édition.** Une technique de classification des différentes instances d'un objet vidéo appelées VOP (*video object plane*) en fonction de leur angle de prise de vue a aussi été développée [20]. Cette méthode repose sur l'utilisation d'un critère de vraisemblance permettant d'évaluer le caractère significatif des paramètres d'un modèle de mouvement affine décrivant le mouvement relatif objet/caméra. Chaque transition entre deux VOPs successifs est étiquetée comme "constante" (pas de modification de l'angle de vue) ou "variable" (modification du point de vue). À chaque transition est associé le type de mouvement correspondant. L'objet vidéo est alors caractérisé par une succession de segments temporels avec ou sans modification du point de vue. Tous les VOPs d'un segment étiqueté "constant" peuvent être considérés similaires, à une transformation plane près (si l'on ne prend pas en compte les effets d'illumination ou d'occlusion). Pour chacun de ces segments, un VOP représentatif (par exemple avec le facteur d'échelle le plus grand) peut alors être choisi.

La trajectoire que doit suivre l'objet dans la scène est découpée en un ensemble de segments, séparés par des instants clés. La trajectoire entre les instants clés est supposée linéaire. Ces instants clés sont obtenus en analysant les trajectoires originales et les nouvelles trajectoires de l'objet. Si la qualité du résultat final n'est pas suffisante, l'opérateur a la possibilité de rajouter manuellement de nouveaux instants clés. De plus, une méthode basée sur une calibration initiale du fond de la scène permet de déterminer automatiquement le facteur d'échelle associé à n'importe quel point de la zone calibrée. Le positionnement d'un objet sur une ligne de niveau permet alors de lui associer automatiquement le facteur d'échelle correspondant. Cette calibration utilise la notion de lignes de niveau, où une ligne de niveau est définie comme l'ensemble des points de l'image se situant à une distance identique de la caméra. La phase de calibration est réalisée de manière interactive en début de traitement. Cette méthode a été validée sur des séquences réelles. Des modifications significatives de la trajectoire d'objets non-rigides (des personnages) ont ainsi pu être réalisées tout en gardant un bon niveau de réalisme. L'insertion et la suppression d'objets est illustrée sur la figure 8.

**Prise en compte des liens sémantiques entre objets vidéo.** Comme présenté ci-dessus, les manipulations que l'on fait sur une séquence vidéo considèrent, pour en maintenir la cohé-

rence, les conditions d'illumination de la séquence mixée et les caractéristiques dynamiques des objets que l'on manipule. Néanmoins, manipuler une séquence demande de considérer d'autres conditions qui participent également à la cohérence globale d'une scène. En particulier, certains objets appartenant à une séquence peuvent être associés par une relation logique, par une relation de cause à effet, qui n'est ni liée à l'illumination ni aux trajectoires.

De façon générale, ces relations logiques peuvent être vues comme des liens sémantiques reliant divers objets d'une scène. Pour aborder ce problème, il est important de classifier les différents types d'interaction sémantiques entre les objets, par exemple, les relations décalées temporellement, les interactions impliquant un contact physique dans l'image, les relations bijectives ou non, les relations entre un évènement sonore et un évènement visuel, etc.

## 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.

**Résumé :** *Les travaux ont porté sur le codage à bas débit basé d'une part sur une représentation par régions de la vidéo et basé d'autre part sur un modèle 3D de la scène. Nous avons par ailleurs développé un algorithme de codage robuste et de codage conjoint adapté à une transmission du signal vidéo sur l'Internet avec pertes de paquets. Ces mécanismes ont été développés d'abord dans un cadre de transmission point-à-point (unicast) puis étendus à une transmission multi-point (multicast) caractérisée par une hétérogénéité des canaux de transmission. Afin d'adapter le signal vidéo aux caractéristiques du canal, nous avons développé des modèles de prédiction de bande passante et des modèles de régulation du débit des données à transmettre, qui tiennent compte des contraintes en termes de délais caractérisant des signaux multimédia temps-réel. Dans le domaine des communications sur des liens sans fils, en collaboration avec le projet SIGMA2, une modélisation sous forme d'automate stochastique d'un codeur statistique a permis de développer un algorithme de décodage robuste de codes à longueurs variables. Le décodage repose sur des techniques d'estimation Bayésienne. Un décodeur conjoint source-canal, mettant en œuvre le principe turbo (ou de propagation de croyance) et les techniques d'estimation Bayésienne, a aussi été testé sur des sources théoriques de Gauss-Markov.*

### 6.2.1 Codage orienté régions, modèles 3D et textures

**Participants :** Dubhe Chavira-Martinez, Franck Galpin, Luce Morin, Stéphane Pateux.

Nous avons développé un schéma de compression de signaux vidéo basé sur une modélisation 3D de la scène. Ce schéma est destiné à coder à bas débit des scènes fixes telles que

le fond fixe (ou l'image panoramique). Le schéma développé suit une structure I-P-B (Intra-Prédit-Bidirectionnel) classique. La séquence vidéo est représentée par un ensemble d'images clés codées en mode I, par un ensemble de modèles 3D valides entre 2 images-clés, et par une suite des positions de la caméra [23]. Une image intermédiaire entre 2 images-clés est reconstruite par projection du modèle 3D texturé par une image-clé, sur le point de vue de la caméra. L'information correspondante peut être soit transmise (images de type P) ou interpolée (images de type B). Les informations à transmettre (modèles 3D, positions caméra) sont issues de l'analyse automatique de la séquence présentée dans la section 6.1.2. Une comparaison quantitative (mesure de PSNR) et qualitative (qualité visuelle subjective) avec des schémas de codage normalisés (MPEG et H.263+), sur des séquences réelles d'intérieur et d'extérieur, a montré la validité de cette approche pour des débits de l'ordre de 100 kbit/s (format CIF) (voir figure 9). Noter que cette approche permet, par rapport aux schémas normalisés, de reconstruire d'autres points de vue de la scène.

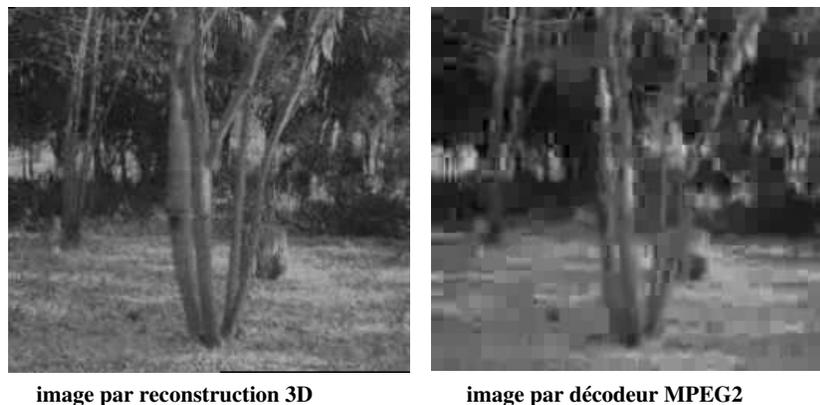


FIG. 9 – *Codage par modèles 3D: comparaison avec MPEG-2; Au même débit notre codeur donne un PSNR similaire mais une qualité visuelle supérieure et des fonctionnalités de navigation dans la scène.*

### 6.2.2 Codage robuste et scalable, et régulation de débit associée

**Participants :** Christine Guillemot, Reda Hosny, Fabrice Le Léannec, Jérôme Viéron.

L'algorithme adaptatif de choix de modes de codage (intra ou inter), développé en 1999, a été enrichi de modèles débit-distorsion fiables des performances de la quantification. Ces modèles permettent une régulation fine du débit et un compromis optimum entre les performances en compression et la résistance aux pertes de paquets, dans un cadre de codage vidéo à base de prédiction temporelle (MPEG ou H.26x) [35]. L'algorithme a été étendu à un mode de représentation scalable (scalabilité de type SNR, c'est-à-dire en rapport signal à bruit) ou multi-niveaux des signaux vidéo [36]. En plus de la distorsion due à la quantification et de la distorsion induite par le canal sur chaque niveau de scalabilité, il tient compte des dépendances entre les niveaux de scalabilité et de l'effet de propagation des pertes inter-niveaux. Une tech-

nique de dissimulation des pertes a en outre été développée afin de minimiser l'impact de la propagation des pertes entre les niveaux de scalabilité.

L'optimisation de la qualité de service passe aussi 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 de bande passante, de taux de pertes, et de délais. Il est donc nécessaire de mettre en place dans les schémas de codage des modèles de prédiction de bande passante et de régulation du débit des signaux transmis. Les approches de contrôle de débit (ou de congestion) publiées à ce jour ne tiennent pas compte des caractéristiques débit-distorsion des signaux et de leur contrainte de délai. Un modèle de régulation de débit qui s'adapte à la fois aux modèles débit-distorsion des signaux compressés, à leur contrainte de délai maximum, et à divers modèles de prédiction de la bande passante disponible sur le réseau, a été développé. Cette étude a permis également de révéler certaines faiblesses dans les approches de prédiction de bande passante dites « TCP-compatible » publiées à ce jour (estimation du processus de pertes, hypothèse sur la taille fixe des paquets non valide pour des signaux vidéo). Ce modèle de régulation de débit a d'abord été validé dans un cadre de transmission point-à-point.

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 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 [36]. Le modèle de régulation hybride est associé à un mécanisme de classification (ou de *clustering*). 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 mécanisme de classification a été développé par le projet PLANETE. 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 s'inspirant des concepts de réseaux actifs 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.

Les algorithmes ci-dessus sont aussi étudiés dans un environnement de différenciation de services. En particulier, un mécanisme de marquage sémantique des données contenues dans un flux vidéo compressé est étudié afin de garantir un traitement différencié de ces données au niveau des mécanismes de gestion des files d'attente à l'étude pour une évolution de l'Internet vers une offre de services différenciés.

### 6.2.3 Allocation de débit conjointe source/canal pour transmission vidéo sur Internet

**Participants :** Bilal Elali, Christine Guillemot, Xavier Hénocq.

Les techniques de contrôle de congestion et de codage robuste permettent d'adapter les flux vidéo transmis aux caractéristiques du réseau, et ceci afin de contribuer à l'amélioration de la qualité de la transmission et de minimiser l'impact des erreurs ou des pertes sur la qualité du signal reconstruit. Néanmoins, il s'avère souvent nécessaire de faire appel à des techniques complémentaires de correction de pertes, afin d'accroître la qualité apparente de la transmission. Le problème posé est alors le partage de la bande passante prédite entre le débit utile (des données vidéo) et le débit de redondance. Nous avons développé un mécanisme d'allocation de débit conjoint source-canal : connaissant la probabilité effective des pertes après décodage, la répartition de débit entre la source et la redondance (ou codage de canal) est réalisée de manière à minimiser la distorsion globale (source+canal) [14]. Cet algorithme vient conditionner les taux de rendements des codes correcteurs d'erreurs utilisés et contrôler la corrélation à introduire ou à maintenir au niveau de la représentation compressée des données vidéo. Cette optimisation globale de la chaîne de communication, incluant codage de source et codage de canal, avec une protection inégale des différentes portions de flux vidéo, permet un gain en terme de qualité (PSNR) du signal reconstruit. Cet algorithme a été étendu à un mode de représentation scalable des signaux vidéo et adapté au contrôle de congestion hybride émetteur/récepteurs associé au mécanisme de classification évoqué ci-dessus [29].

### 6.2.4 Décodage conjoint source-canal

**Participants :** Christine Guillemot, Francisco Garcia-Ugalde.

Cette étude est menée en collaboration avec le projet SIGMA2 (Éric Fabre). Les objectifs de forte compression conduisent à l'utilisation de codes statistiques à longueurs variables. Ces codes 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 de construire un algorithme de décodage robuste de codes à longueurs variables (codes d'Huffman) et de codes à longueurs variables réversibles (RVLC). Le décodage basé sur une estimation Bayésienne permet de réduire fortement les taux d'erreurs symboles et d'accroître la capacité de resynchronisation du décodeur mesurée à l'aide de la distance de Levenshtein<sup>1</sup>. La modélisation sous forme de chaînes de Markov des dépendances entre les symboles de source d'une part, et entre les bits de la séquence codée par le code à longueur variable d'autre part, fournit un cadre naturel à l'introduction de petits mots de synchronisation permettant une synchronisation « douce » du décodeur. Par rapport à des codes auto-synchronisants, ces mots de synchronisation ne doivent pas vérifier les conditions de préfixe des codes à longueurs variables et présentent donc une redondance moindre. L'application du principe turbo (ou de propagation de croyance) entre le décodeur de codes à longueurs variables et le décodeur de canal utilisés en alternance selon une structure série a permis d'atteindre des taux d'erreurs

---

1. A. Guyader, E. Fabre, C. Guillemot, M. Robert, 'Joint source-channel turbo decoding of entropy-coded sources', soumis pour publication, IEEE journal of selected areas in communications

symboles très faibles pour des taux d'erreurs bits au niveau du canal de l'ordre de  $5 \times 10^{-2}$ . Les algorithmes ont été testés sur des sources de Gauss-Markov. L'intégration de ces décodeurs au sein d'un décodeur vidéo est en cours afin de valider les approches sur des signaux réels.

### 6.2.5 Codage conjoint source-canal

**Participants :** Zhikui Chen, Christine Guillemot, Thomas Guionnet, Xavier Hénocq, Stéphane Pateux, Gagan Rath.

Les communications sans fil sont caractérisées par des transmissions fortement bruitées avec des erreurs ponctuelles ou en rafales. L'approche classique qui consiste à considérer les deux problèmes du codage de source et de la protection (codage de canal) de manière séparée et à mettre en cascade le codeur de source, conçu pour un canal sans perte, avec des codes correcteurs d'erreurs (ex: codes convolutionnels, turbo codes, correction par anticipation, ...) s'avère sous-optimale. Malgré les bonnes performances des codes correcteurs d'erreurs, les contraintes d'un système réel (complexité, délai de décodage,...) conduisent à des taux d'erreurs résiduels non négligeables. Il est donc apparu nécessaire d'optimiser les deux opérations de codage (source et canal) conjointement.

Les opérateurs de compression ont été conçus à ce jour de manière d'une part à décorréler le signal (transformations) et d'autre part à optimiser les performances débit-distorsion de quantification. Le codage de canal a pour objectif l'ajout de redondance dans le train binaire transmis sur le réseau afin de pouvoir corriger les erreurs de transmission. Dans le modèle utilisant des codes correcteurs associés aux flux vidéo transmis, les performances du système de communication peuvent aussi être améliorées par la mise en place de techniques de décodage conjoint source-canal où une connaissance *a priori* des statistiques de la source est exploitée au niveau du décodage de canal (voir section 6.2.4). Il est aussi possible d'introduire de la redondance directement dans le flux compressé, en concevant des codeurs de source qui puissent s'adapter à des caractéristiques et modèles du canal variant dans le temps. On peut par exemple concevoir des transformations et des quantificateurs maintenant une quantité « contrôlée » de corrélation ou d'informations redondantes. Cette idée est à l'origine des méthodes de codage par descriptions multiples et des transformations redondantes (comme par exemple les bancs de filtres sur-échantillonnés, ou les transformées multi-ondelettes). Noter que le codage par descriptions multiples est une extension du codage avec critère de fidélité, suivant la théorie débit distorsion, au cas multi-canal.

Nous avons étudié les principes de codage par descriptions multiples, en considérant d'une part les transformées multi-ondelettes et d'autre part la quantification redondante. Pour répondre à des objectifs supplémentaires de transmission progressive des signaux à des fins d'adaptation du débit aux caractéristiques variables du canal de transmission, nous avons mis en place un schéma de codage d'images fixes à descriptions multiples qui s'inspire des concepts de progressivité de JPEG2000. Nous avons également démarré une étude consistant à transposer certains concepts de la théorie des codes correcteurs au niveau des transformations multi-résolutions utilisées dans les schémas de compression.

### 6.2.6 Régulation de débit pour la transmission d'images satellitaires sur canal à débit fixe

**Participants :** Delphine Le Guen, Stéphane Pateux, Claude Labit.

Nous avons mis en place un schéma de régulation locale basé sur une transformée en ondelettes de l'image, une classification automatique de ces coefficients d'ondelettes et une transmission progressive de l'information. Notre technique permet d'améliorer la qualité locale sur l'image en faisant en sorte que les zones d'intérêt, définies via la classification, soient codées en priorité. Pour ce faire, nous avons développé une méthode de pondération optimale des coefficients d'ondelettes, pondération qui intervient préalablement à un codage de type EZW (*embedded Zerotree Wavelet*). Les poids, appliqués aux coefficients d'ondelettes, dépendent des filtres de synthèse utilisés lors de la transformation en ondelettes et de l'appartenance ou non du coefficient à une zone d'intérêt. Ces zones d'intérêt sont mises en évidence en intégrant la gêne effective liée à l'apparition de deux artéfacts caractéristiques d'un codage basé ondelettes. Nous définissons de ce fait deux zones d'intérêt, l'une relative aux zones de fortes transition et l'autre relative aux zones uniformes. Les résultats obtenus sur l'image Gênes sont illustrés sur les figures 10 et 11. La carte de classification présente les zones de fortes transitions en gris et les zones uniformes en noir. Pour un taux de compression de 6, nous comparons nos résultats à ceux obtenus par l'algorithme EZW sans classification. Les images reconstruites sont visualisées après zoom et égalisation d'histogramme. Nous pouvons remarquer une atténuation des rebonds sur les digues ainsi que sur la piste d'atterrissage (bas de l'image).

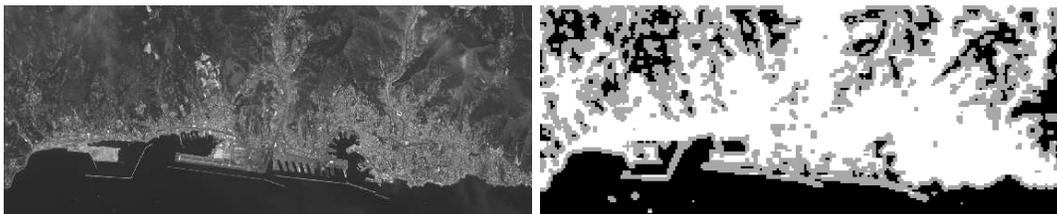


FIG. 10 – Gênes : image originale SPOT5 (©Cnes) (2000x800) et carte d'étiquettes

### 6.3 Tatouage d'images

**Participants :** Gaëtan Le Guelvouit, Stéphane Pateux, Christine Guillemot.

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

**Résumé :** *Notre approche en tatouage cherche à exploiter d'une part, et à mieux formaliser, voire quantifier d'autre part, le parallèle existant entre les techniques de tatouage d'images et de transmission d'informations à travers un canal bruité. Le message à insérer peut être vu comme un message à transmettre sur un canal bruité, à savoir l'image, qui peut subir des attaques de formes diverses (compression, filtrage, transformations géométriques, etc). Ce parallèle conduit naturel-*

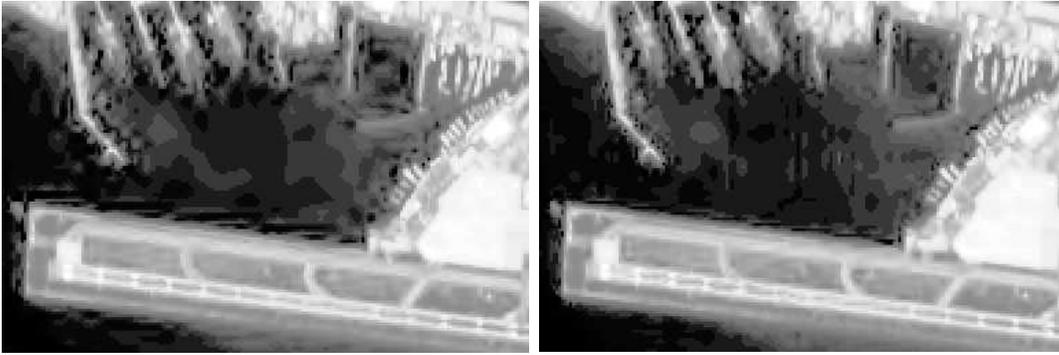


FIG. 11 – *Technique de codage EZW standard et codage de type EZW classifié pour un taux de 6, sur une zone particulière de l'image*

*lement à considérer les codes correcteurs comme outil de protection de la marque, de manière à permettre in fine une détection et une extraction robuste. L'apport des codes correcteurs par rapport à la robustesse du tatouage face à diverses attaques (compression JPEG, retaille, filtrage,...) a été étudié et quantifié.*

Une étude comparative des différentes techniques existantes de tatouage nous a conduits à définir un schéma de tatouage dans un domaine transformé (transformation en ondelettes) à l'aide d'une technique d'insertion par étalement de spectre. Dans un premier temps, une caractérisation des perturbations observées lors de l'extraction d'une marque insérée dans une image suite à différentes attaques a été effectuée. Cette étude a permis de définir un modèle du bruit présent lors de la phase d'extraction et de mettre à profit cette connaissance afin d'améliorer le décodage de la marque insérée. Ceci est réalisé grâce à l'utilisation de codes correcteurs convolutionnels et à l'utilisation d'un décodeur *soft* de Viterbi à l'entrée duquel on injecte une mesure de confiance sur les bits extraits. Cette nouvelle technique de tatouage nous a permis d'évaluer l'intérêt des codes correcteurs d'une part et d'estimer la capacité de stockage (ou volume de la marque) d'une image en fonction du niveau d'attaque toléré d'autre part.

## 7 Contrats industriels (nationaux, européens et internationaux)

### 7.1 Contrats industriels bi-latéraux

#### 7.1.1 CTI-FT RD

**Participants :** Christine Guillemot, Xavier Hénocq, Fabrice Le Léannec.

- Numéro de la convention : 1 98 C 391 00 31324 01 2.
- Intitulé : *Transmission multi-point de données multimédia sur réseaux hétérogènes.*

- Activité de recherche concernée : § 6.2.2 et § 6.2.3.
- Partenaire : Inria, FT RD.
- Financement : CTI-FT RD.
- Période : 36 mois à partir de septembre 1998.

Notre contribution porte sur les thèmes de recherche suivants :

- l'étude de techniques de codage vidéo robustes aux pertes de paquets ;
- l'étude de modes de représentation scalable permettant un bon compromis entre décorrélation et robustesse aux pertes de paquets ;
- l'étude de mécanismes de codage conjoint source-canal, permettant un partage de bande passante optimal entre le débit utile et le débit de redondance ;
- l'étude de mécanismes de masquage de pertes adaptés à l'Internet ;
- l'étude et le développement de mécanismes modifiés et adaptés à un Internet multi-service.

### 7.1.2 Convention Celar

**Participants :** Thierry Fauconnier, 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 pour la phase 1 : 6 mois à partir de septembre 1999.
- Période pour la phase 2 : 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 nous intéressons dans ce contexte au problème du décodage conjoint source-canal, et au codage par descriptions multiples.

### 7.1.3 Convention Microprocess

**Participants** : Henri Nicolas, Franck Denoual.

- Numéro de la convention : 1 97 C 531 00 31315 012.
- Intitulé : *Développement de méthodes semi-automatiques de composition de scènes et implantation sur noyau temps-réel.*
- Activité de recherche concernée : Thèse de Franck Denoual, § 6.1.5.
- Partenaires : Inria, Microprocess.
- Financement : Convention Cifre Microprocess.
- Période : 36 mois à partir de mai 1997.

Une collaboration est en cours avec la société Microprocess, spécialisée dans les problèmes temps-réel et exploitant les concepts objets, sur le thème : « développement de méthodes semi-automatiques de composition de scènes et implantation sur un noyau temps-réel ». L'objectif pour la société Microprocess est d'étendre ses activités dans le domaine du temps-réel au monde de la compression vidéo MPEG et de la composition de scènes vidéo. Les décodeurs MPEG 1,2 et 4 ont été implantés sur le noyau temps-réel SOFT KERNEL développé par Microprocess selon la méthodologie OMT. De plus, une nouvelle méthode semi-automatique de mixage d'objets vidéo a été développée. Elle permet de modifier le contenu basé objet d'une scène en enlevant ou en ajoutant un objet dans une séquence vidéo, ou encore en modifiant la trajectoire d'objets existants.

### 7.1.4 Convention Thomson-TCO

**Participant** : Henri Nicolas.

- Numéro de la convention : 1 98 C 307 00 31324 012.
- Intitulé : *Création interactive d'objets mosaïques 2D1/2 et 3D.*
- Activité de recherche concernée : Thèse d'Agnès De Simon, § 6.1.4.
- Partenaires : Inria, Thomson-TCO.
- Financement : Convention Cifre Thomson-TCO.
- Période : 36 mois à partir de mai 1998.

Cette convention Cifre avec Thomson Optronique concerne l'encadrement de la thèse d'Agnès De Simon sur le thème : « création interactive d'objets mosaïques 2D1/2 et 3D ». L'objectif de cette étude est de développer de nouveaux outils permettant la reconstitution interactive de scènes 3D à partir d'outils de calibration interactive et d'algorithmes de création d'images mosaïques.

### 7.1.5 Convention Cnes

**Participants :** Claude Labit, Delphine Le Guen, Stéphane Pateux.

- Numéro de la convention : 0 98 C 314 00 31324 011.
- Intitulé : *Régulation de débit pour systèmes de compression embarqués sur satellite.*
- Activité de recherche concernée : Thèse de Delphine Le Guen, § 6.2.6.
- Partenaires : Inria, Cnes, Alcatel Aérospatiale Cannes.
- Financement : CNES.
- Période : 36 mois à partir de octobre 1997.

L'étude a pour objectif la définition de mécanismes de régulation de débit pour la transmission d'images satellitaires, en collaboration avec le Cnes (bourse Cnes pour le doctorat de D. Le Guen) et Alcatel Aérospatiale Cannes. Cette recherche s'appuie sur un schéma de régulation réalisé antérieurement par le Cnes pour l'application Spot5. Celui-ci est basé sur une transformation DCT suivie d'une quantification scalaire avec un module de régulation intégré, activé par « fauchées » de 8 lignes d'images. Les résultats sont satisfaisants pour des taux de compression allant jusqu'à 2, mais la recherche de taux de compression plus élevés nous pousse à revoir les modules de compression et de régulation.

### 7.1.6 Convention Thomson Multimédia

**Participants :** Luce Morin, Philippe Robert.

- Activité de recherche concernée : Thèse de Yannick Nicolas, § 6.1.2.
- Partenaires : Inria, Thomson Multimédia.
- Financement : Convention Cifre Thomson Multimédia.
- Période : 36 mois à partir de octobre 1998.

L'objectif de cette convention Cifre est d'étudier une chaîne de navigation à distance dans des modèles de scènes 3D réelles. Plus précisément, il s'agit d'étudier et de développer les éléments d'une chaîne de modélisation, de transmission et de rendu pour la navigation virtuelle dans des scènes réelles complexes, dans un contexte de diffusion et d'applications client-serveur. Les systèmes actuels permettent la navigation dans des conditions limitées quant à la qualité des images et au réalisme de la scène, et à la flexibilité de navigation, elle-même restreinte par le type de mouvement de la caméra virtuelle (panoramique, zooms, translation, etc.).

## 7.2 Contrats nationaux

### 7.2.1 RNRT-V2net

**Participants** : Franck Galpin, Luce Morin.

- Numéro de la convention : 2 00 C 0339.
- Intitulé : Projet RNRT-V2NET, *Visite virtuelle en réseau*.
- Activité de recherche concernée : § 6.1.2.
- Partenaires : Thomson Multimédia, Inria, FT RD.
- 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.

### 7.2.2 RNRT-Osiam

**Participant** : Stéphane Pateux.

- Numéro de la convention : 298C5620031324011.
- Intitulé : Projet RNRT-Osiam, *Outils de segmentation spatio-temporelle pour MPEG-4/MPEG-7*.
- Activité de recherche concernée : § 6.1.1.
- Partenaires : Philips-Lep, Ireste, Irisa/Inria-Rennes, i3s, Timc, Creatis, FT RD.
- Financement : RNRT.
- Période : 36 mois à partir de décembre 1998.

Le projet exploratoire RNRT-Osiam a pour objectif la spécification d'outils de segmentation spatio-temporelle pour MPEG-4/MPEG-7. Diverses approches complémentaires de segmentation hiérarchique et de suivi d'objets vidéo sont mises en oeuvre, comparées, intégrées au sein d'une même plate-forme logicielle. Les domaines applicatifs visés sont la vidéosurveillance, la télémédecine et la composition d'objets vidéo pour des applications de post-production, et d'indexation de contenus vidéo.

### 7.2.3 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 : FT RD, Philips-Lep, Irisa/Inria-Rennes, ENST-Paris.
- Financement : RNRT.
- Période : 36 mois à partir de mars 1999.

Le projet exploratoire Cohrainte s'intéresse à la spécification de sources scalables à granularité fine pour la représentation de flux de données audio-visuelles. L'application visée étant le transport de ces flux sur réseaux IP, une attention particulière est portée à l'allocation conjointe, et négociée dynamiquement, des ressources aux flux audio et vidéo.

### 7.2.4 RNRT-Visi

**Participants** : Bilal ElAli, Christine Guillemot, Jérôme Viéron.

- Numéro de la convention : 2 99 A 461 00 000MC 011.
- Intitulé : *vidéo interactive sur Internet*.
- Activité de recherche concernée : § 6.2.2.
- Partenaires : Thomson Multimédia, Inria, Edixia, Eurecom, FT RD.
- Financement : RNRT.
- Période : 24 mois à partir de juin 1999.

Le projet a pour objectif de fournir des solutions au problème de la gestion de la qualité de service dans des applications de transmission vidéo sur l'Internet. Il étudie différentes techniques de représentation scalable à grain fin des sources avec des mécanismes de régulation de débit appropriés, de contrôle de congestion et le contrôle d'erreur. Il étudie également les concepts architecturaux et systèmes liés à la répartition des sources entre plusieurs serveurs et les techniques de cache vidéo. L'objectif est en outre de développer des solutions aussi pérennes que possible, à savoir qui puissent d'une part s'appliquer à l'Internet *best-effort* d'aujourd'hui, mais être également compatibles avec une évolution – à l'étude au sein de l'IETF (*Internet Engineering Task Force*) – du réseau vers l'intégration et la différenciation de services.

## 7.3 Contrats européens

### 7.3.1 Esprit-Nemesis

**Participants** : Henri Nicolas, Lionel Oisel.

- Numéro de la convention : 196C4590031315006.
- Intitulé : Convention Esprit-Nemesis, *New Multimedia Services Using Analysis Synthesis*.
- Activité de recherche concernée : § 6.1.4.
- Période : 24 mois à partir du 01/09/1998.
- Partenaires : Thomson CSF Optronique, Inria, HHI-Berlin, ICS-Forth, Mikros-France.
- Financement : CEE-ESPRIT.

Le monde de la post-production vidéo est actuellement très demandeur d'outils permettant de manipuler le contenu des images et de créer de nouvelles séquences en utilisant la notion d'objets vidéo. Le projet Nemesis, organisé autour de trois partenaires académiques (Inria, Hhi (Allemagne), et Forth (Grèce)), et de deux partenaires industriels (Thomson Optronique et Mikros (France)), impliqués dans la post-production vidéo, a pour objectif le développement et la commercialisation d'un produit logiciel intégrant des fonctionnalités telles que la segmentation spatiale ou spatiotemporelle, la création d'images mosaïques ou le suivi 3D. Ce logiciel a été développé sur la base du logiciel *After Effect* utilisé dans le monde professionnel de la post-production. Plus spécifiquement, la contribution scientifique de Temics à ce projet concerne le développement d'outils de création d'images panoramiques, et l'utilisation de ces images pour la manipulation des séquences d'origines (modification du point de vue et de la focale de la caméra, modification de la trajectoire de certains objets dans la scène, ...).

### 7.3.2 IST-Openise

**Participants** : Christine Guillemot, Reda Hosny, Gagan Rath.

- Numéro de la convention : 1 99 A 1800 00 00MPR.
- Intitulé : Convention IST-OPENISE, *Open Platform for Enhanced Interactive Services*.
- Activité de recherche concernée : § 6.2.2.
- Partenaires : Italtel, France Télécom, DT-Berkom, FHW, Inria, Optibase, Orkit, Siemens, VideoTime.
- Période : 24 mois à partir de janvier 2000.
- Financement : CEE.

L'objectif du projet est le développement d'une plate-forme de communication sur Internet incorporant des applications interactives telles que mondes virtuels partagés, télévision interactive... Les nouvelles technologies développées dans le projet concernent le codage robuste, le contrôle de transmission des contenus dans un contexte de différenciation de services, les outils de modélisation et de partitionnement de mondes virtuels permettant de faciliter la navigation au travers d'un réseau de communication et qui permettent également d'adapter les contenus transmis à la navigation de l'utilisateur ainsi qu'à la bande passante disponible sur le réseau. Dans le cadre de ce projet, TEMICS se concentre sur les questions de différenciation de services.

### 7.3.3 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 : § 6.2.3.
- Partenaires : France Télécom, DT-Berkom, Blaxxun, Cselt, Enst-br, Epfl, GMD, Inria, RUS, Telenor, Univ. Genève.
- Période : 30 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 audio-visuels 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 contribue 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.

## 8 Actions régionales, nationales et internationales

### 8.1 Réseaux et groupes de travail internationaux

Le projet Temics participe à plusieurs projets et groupes de travail internationaux :

- participation à l'IETF - Internet Engineering Task Force -, et en particulier aux réunions des groupes AVT, RMTP et DiffServ,
- contribution à la normalisation MPEG-4 au travers de partenaires industriels.

## 8.2 Réseaux et groupes de travail 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'Unam-Mexico (professeur Garcia-Ugalde) dans le cadre d'une convention CNRS-Conacyt ;
- collaboration France-Egypte (participants : S. Pateux, C. Guillemot) : dans le cadre de cette collaboration, le projet Temics a assuré un module d'enseignement intitulé « communication vidéo » à l'ITI (Information Technology Institute), Le Caire.
- participation au colloque France-Taiwan (participants : L. Amsaleg et H. Nicolas), Paris, Avril 2000.
- participation au colloque France-Chine sous l'égide des académies des sciences française et chinoise (participant : C. Guillemot), Pékin, Octobre 2000.

## 8.3 Visites et invitations de chercheurs

Visite du professeur R. Ansari, UIC, university of Illinois à Chicago, pour un mois en mai 2000. La collaboration avec le professeur Ansari a porté sur les thèmes du codage conjoint source-canal et du codage par descriptions multiples pour la transmission vidéo sur des réseaux à qualité de service non garantie.

## 9 Diffusion de résultats

### 9.1 Animation de la communauté scientifique

- S. Pateux est coordinateur scientifique au niveau de l'IRISA du CNRT (Centre National de Recherche Technologique) TIM-Bretagne.
- C. Guillemot est nommée, pour deux années, éditeur associé du journal IEEE transactions on Image Processing;
- C. Guillemot est membre du comité de programme de la conférence Spie-VCIP 2000;
- C. Guillemot est membre de la commission 2 du RNRT;
- C. Guillemot a participé à l'évaluation des propositions de projets de recherche européens IST;
- L. Morin, H. Nicolas et S. Pateux ont participé activement (par le biais de présentations) aux réunions du GDR Isis.

## 9.2 Enseignement

- Diic-LSI, Ifsic, université de Rennes 1 (H. Nicolas : codage) ;
- Diic-INC, Ifsic, université de Rennes 1 (L. Morin, C. Labit, H. Nicolas, L. Bonnaud, F. Garcia Ugalde : traitement d'images, vision 3D, mouvement, codage, compression d'images, cryptographie) ;
- 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 (L. Morin : analyse du mouvement 3D), (C. Guillemot : codage) ;
- Enic, Villeneuve-d'Ascq, (C. Labit : Objets vidéo et communication d'images numériques) ; (C. Guillemot : Vidéo sur IP) ;
- Ensar Rennes, option informatique (L. Morin : fondements du traitement d'image, morphologie mathématique) ;
- ENSTBR (C. Labit : Objets vidéo et communication d'images numériques) ; (C. Guillemot : Les normes de codage vidéo de MPEG-1 à MPEG-4) ;
- ENS-Cachan (C. Labit : Objets vidéo et communication d'images numériques).

## 10 Bibliographie

### Ouvrages et articles de référence de l'équipe

- [1] C. GUILLEMOT, E. CETIN AND R. ANSARI, *Nonrectangular Wavelet Representation of 2-D Signals. Application to Image Coding*, Eds Elsevier Press 1994, September 1994, p. 27–64.
- [2] C. GUILLEMOT, R. ANSARI, «Layered Coding Schemes for Video Transmission on ATM Networks», *Journal of Visual Communication and Image Representation* 5, 1, March 1994, p. 62–74.
- [3] J. P. LEDUC, J. M. ODOBEZ, C. LABIT, «Adaptive motion-compensated wavelet filtering for image sequence coding», *IEEE Transactions on Image Processing* 6, 6, June 1997, p. 862–879.
- [4] R. MOHR, L. MORIN, E. GROSSO, «Relative positioning with uncalibrated cameras», *in: Geometric Invariance in Computer Vision*, Editors: A. Zisserman J.L. Mundy, MIT Press, 1992, p. 440–460.
- [5] E. NGUYEN, C. LABIT, «Selective coding by focus of attention: a new tool to achieve VLBR video coding (Invited Paper)», *in: Proceedings of Eusipco'96, Trieste, Italie, 1*, p. 431–434, September 1996.
- [6] H. NICOLAS, C. LABIT, «Motion and illumination variation estimation using a hierarchy of models: application to image sequence coding», *Journal of Visual Communication and Image Representation* 6, 4, December 1995, p. 303–316.

- [7] A. TAMTAOUI, C. LABIT, «Constrained disparity and motion estimators for 3DTV image sequence coding», *Signal Processing: Image Communication* 4, 1991, p. 45–54.
- [8] G. TZIRITAS, C. LABIT, *Motion Analysis for Image Sequence Coding*, Elsevier Science, Book in Series “Advances in Image Communication”, 390 pages, 1994.
- [9] D. WANG, C. LABIT, J. RONSIN, «Morphological spatio-temporal simplification for video image segmentation», *IEEE Trans on Circuits and Systems for Video Technology* 7, 3, June 1997, p. 549–554.

## Articles et chapitres de livre

- [10] L. AMSALEG, P. GROS, «Content-based Retrieval Using Local Descriptors: Problems and Issues from a Database Perspective», à paraître, *Pattern Analysis and Applications, Special Issue on Image Indexation*, 2001.
- [11] F. DENOVAL, H. NICOLAS, «Étude et modifications de trajectoires d’objets vidéo pour l’édition et la post-production vidéo», *accepté pour publication dans le numéro spécial image et vidéo de la revue Calculateur Parallèle*, 2000, à paraître.
- [12] F. GALPIN, L. MORIN, «Video coding using streamed 3D representation», *accepté pour publication dans le numéro spécial image et vidéo de la revue Calculateur Parallèle*, 2000, à paraître.
- [13] C. GUILLEMOT, T. TURLETTI, *Algorithmes de codage et de contrôle de transmission vidéo sur réseaux hétérogènes*, Éditions Hermes, 2000, ch. dans ouvrage intitulé «Systèmes Multimédia Communicants», à paraître.
- [14] X. HÉNOCCQ, C. GUILLEMOT, «Source Adaptive Error Control for Real-time Video over the Internet», *accepté pour publication dans le numéro spécial image et vidéo de la revue Calculateur Parallèle*, 2000, à paraître.
- [15] E. IBN-ELHAJ, D. ABOUTAJDINE, S. PATEUX, L. MORIN, «Robust HOS-region based motion estimation technique for noisy image sequences», *accepté pour publication dans le numéro spécial image et vidéo de la revue Calculateur Parallèle*, 2000, à paraître.
- [16] H. NICOLAS, S. PATEUX, D. LE GUEN, «Minimum Description Length Criterion and Segmentation Map Coding for Region-based Video Compression», *accepté pour publication in IEEE Transactions on Circuits and Systems for Video Technology*, 2000, à paraître.
- [17] H. NICOLAS, «New methods for Dynamic Mosaicking», *accepté pour publication in IEEE Transactions on Image Processing*, 2000, à paraître.
- [18] S. PATEUX, «Spatial segmentation of color images according to the MDL formalism», *accepté pour publication dans le numéro spécial image et vidéo de la revue Calculateur Parallèle*, 2000, à paraître.

## Communications à des congrès, colloques, etc.

- [19] A. DE SIMON, J. FIGUE, H. NICOLAS, «Video reframing relying on panoramic estimation based on a 3D representation of the scene», *in: Proceedings of VCIP (SPIE)*, p. 1508–1516, Perth, Australia, juin 2000.

- 
- [20] F. DENOVAL, H. NICOLAS, «Artificial Objects Trajectory Modifications for 2D Object-based Video Compositing», *in: Proceedings of VCIP (SPIE)*, p. 1499–1507, Perth, Australia, juin 2000.
- [21] F. DENOVAL, H. NICOLAS, «Édition vidéo par modification de trajectoires d’objets et de mouvement caméra», *in: Actes du colloque CORESA*, p. 55–60, Poitiers, France, septembre 2000.
- [22] F. DENOVAL, H. NICOLAS, «Étude et modifications de trajectoires d’objets vidéo pour l’édition et la post-production vidéo», *in: Workshop ISIVC, Rabat, Maroc*, p. 185–191, octobre 2000.
- [23] F. GALPIN, L. MORIN, S. PATEUX, «Sequences reconstruction from streamed 3D models», *in: Proceedings of the IEEE Conference on Image Processing, vol. III*, p. 636–639, Chicago, Illinois, USA, septembre 2000.
- [24] F. GALPIN, L. MORIN, «Construction de modèles 3D pour la transmission très bas débit», *in: Actes du colloque CORESA*, p. 319–327, Poitiers, France, octobre 2000.
- [25] F. GALPIN, L. MORIN, «Video coding using streamed 3D representation», *in: Proceedings of the International Symposium on Image/Video communications over Fixed and Mobile Networks, vol. 3*, p. 416–423, Rabat, Morocco, avril 2000.
- [26] C. GUILLEMOT, T. TURLETTI, «Joint Source and Channel coding for Video transmission over the Internet», *in: exposé invité à la conférence Networking 2000*, Paris, France, mai 2000.
- [27] X. HÉNOQCQ, C. GUILLEMOT, F. LE LÉANNEC, «Joint Source and Channel Rate Allocation for Video over Erasure Channels», *in: European Signal Processing Conference, EUSIPCO 2000*, Tampere, Finland, septembre 2000.
- [28] X. HÉNOQCQ, C. GUILLEMOT, «Source Adaptive Error Control for Real-time Video over the Internet», *in: International Symposium on Image/Video Communication over Fixed and Mobile Networks, ISIVC, 3*, p. 220–229, Rabat, Morocco, avril 2000.
- [29] X. HÉNOQCQ, F. LE LÉANNEC, C. GUILLEMOT, «Joint Source and Channel Rate Control in Multicast Layered Transmission», *in: International Conference on Visual Communications and Image Processing 2000, VCIP 2000*, p. 296–307, Perth, Australia, juin 2000.
- [30] E. IBN-ELHAJ, D. ABOUTAJDINE, S. PATEUX, L. MORIN, «Robust HOS-region based motion estimation technique for noisy image sequences», *in: International Symposium on Image/Video Communications over fixed and mobile networks, ISIVC’2000, 2*, p. 207–211, Rabat, Maroc, avril 2000.
- [31] D. LE GUEN, S. PATEUX, C. LABIT, G. MOURY, D. LEBEDEFF, «Bit rate and local quality control for on-board satellite images compression», *in: IEEE Data Compression Conference: DCC’2000*, p. 558–559, Snowbird, Utah, USA, mars 2000.
- [32] D. LE GUEN, S. PATEUX, C. LABIT, G. MOURY, D. LEBEDEFF, «Bit rate and local quality control using a rate-distorsion criterion for on-board satellite images compression», *in: 10th European Signal Processing Conference: EUSIPCO’2000, II*, Tampere, Finlande, septembre 2000.
- [33] D. LE GUEN, S. PATEUX, C. LABIT, G. MOURY, D. LEBEDEFF, «Optimal adapted weighting of wavelet coefficients for on-board coding of satellite images», *in: IEEE International Geoscience and Remote Sensing Symposium: IGARSS’2000*, Honolulu, Hawaï, USA, juillet 2000.

- 
- [34] F. LE LÉANNEC, C. GUILLEMOT, «Compatible extension of H.263+ for robust real-time video over the Internet», *in: International Symposium on Image/Video Communications over fixed and mobile networks, ISIVC'2000*, p. 230–235, Rabat, Maroc, April 2000.
- [35] F. LE LÉANNEC, C. GUILLEMOT, «Packet loss resilient H.263+ compliant video coding», *in: European Signal Processing Conference, EUSIPCO 2000*, Tampere, Finland, septembre 2000.
- [36] F. LE LÉANNEC, J. VIÉRON, X. HÉNOCQ, C. GUILLEMOT, «Hybrid Sender and receiver driven rate control in multicast layered video transmission», *in: International Conference on Image Processing, ICIP'2000, III*, p. 532–535, Vancouver, Canada, septembre 2000.
- [37] G. MARQUANT, S. PATEUX, C. LABIT, «Mesh and "Crack Lines": Application to Object-based Motion Estimation and Higher Scalability», *in: IEEE International Conference on Image Processing ICIP 2000, 2*, p. 554–557, Vancouver, BC, Canada, septembre 2000.
- [38] G. MARQUANT, S. PATEUX, C. LABIT, «Mesh-based scalable image coding with rate-distortion optimization», *in: Image and Video Communications and Processing 2000, 3974*, p. 101–110, San Jose, USA, janvier 2000.
- [39] G. MARQUANT, S. PATEUX, C. LABIT, «Mesh-based scalable image coding with rate-distortion optimization», *in: European Signal Processing Conference EUSIPCO 2000*, Tampere, Finland, septembre 2000.
- [40] G. MARQUANT, S. PATEUX, C. LABIT, «Mesh-Based Scalable Video Coding with Rate-Distortion Optimization», *in: Visual Communications and Image Processing 2000, 4067*, p. 967–976, Perth, Australia, juin 2000.
- [41] S. PATEUX, «Spatial segmentation of color images according to the MDL formalism», *in: International conference on Color in Graphics and Image Processing*, p. 89–93, St-Étienne, France, octobre 2000.
- [42] S. PATEUX, «Spatial segmentation of color images according to the MDL formalism», *in: International Symposium on Image/Video Communications over fixed and mobile networks, ISIVC'2000, 2*, p. 107–113, Rabat, Maroc, avril 2000.
- [43] S. PATEUX, «Tracking of video objects using a backward projection technique», *in: Visual Conference on Image Processing, VCIP2000, 4067*, p. 1107–1114, Perth, Australia, juin 2000.
- [44] J.-M. PINEL, H. NICOLAS, «Estimation de la direction d'illumination dans une séquence d'images», *in: Actes du colloque CORESA*, Poitiers, France, septembre 2000.
- [45] M. RZIZA, A. TAMTAOUI, L. MORIN, D. ABOUTAJDINE, «Estimation and segmentaion of a dense disparity map for 3D reconstruction», *in: Proceedings of the IEEE Conference on Acoustic, Speech, and Signal Processing, 4*, p. 2219–2222, Istambul, Turkey, juin 2000.
- [46] M. RZIZA, A. TAMTAOUI, L. MORIN, D. ABOUTAJDINE, «Une approche d'estimation multi-résolution et segmentation d'une carte de disparité dense en facettes planes pour la reconstruction 3D», *in: Proceedings of the Internationsl Symposium on Image/Video communications over Fixed and Mobile Networks, 3*, p. 368–376, Rabat, Marocco, avril 2000.

---

**Rapports de recherche et publications internes**

- [47] L. AMSALEG, P. GROS, R. MEZHOUD, « Mise en base d'images indexées par des descripteurs locaux: problèmes et perspectives », *rapport de recherche n° PI 1316*, Irisa, mars 2000, <http://www.irisa.fr/bibli/publi/pi/2000/1316/1316.html>.
- [48] L. AMSALEG, P. GROS, « A robust technique to recognize objects in images, and the DB problems it raises », *rapport de recherche n° PI 1371*, Irisa, novembre 2000, <http://www.irisa.fr/bibli/publi/pi/2000/1371/1371.html>.
- [49] J. STAUDER, H. NICOLAS, « Object Rotation Axis from Shading », *rapport de recherche n° RR 3865*, Inria, janvier 2000.