

# THESE

*Présentée*

DEVANT L'INSTITUT NATIONAL DES SCIENCES  
APPLIQUEES DE LYON

*Pour obtenir*

**LE GRADE DE DOCTEUR**

**FORMATION DOCTORALE : IMAGES ET SYSTEMES**

**ECOLE DOCTORALE : ELECTRONIQUE, ELECTROTECHNIQUE, AUTOMATIQUE**

PAR

**Sébastien VALETTE**

Ingénieur Génie Electrique – INSA Lyon 1998

**Modèles de maillages déformables 2D et multirésolution  
surfaceutes 3D sur une base d'ondelettes**

**Soutenue le 7 Janvier 2002 devant la Commission d'Examen**

**Jury :**

Michel BARLAUD  
Jean-Marc CHASSERY  
Ho Youl JUNG  
Isabelle MAGNIN  
Françoise PRETEUX  
Rémy PROST

Rapporteur  
Président  
Codirecteur de thèse  
Rapporteur  
Codirecteur de thèse



Directeur : A. STORCK

Professeurs :

AUDISIO S	PHYSICOCHIMIE INDUSTRIELLE
BABOUX JC	GEMPPM*
BALLAND B	PHYSIQUE DE LA MATIERE
BARBIER D	PHYSIQUE DE LA MATIERE
BASTIDE JP	THERMODYNAMIQUE APPLIQUEE
BAYADA G	MAPLY - MATHÉMATIQUES APPLIQUÉES DE LYON
BERGER C (Mlle)	PHYSIQUE DE LA MATIERE
BETEMPS M	AUTOMATIQUE INDUSTRIELLE
BLANCHARD JM	LAEPSI ***
BOISSON C	VIBRATIONS ACOUSTIQUES
BOIVIN M	MECANIQUE DES SOLIDES
BOTTA H	Equipe DEVELOPPEMENT URBAIN
BOTTA-ZIMMERMAN M (Mme)	Equipe DEVELOPPEMENT URBAIN
BOULAYE G (Prof. émérite)	INFORMATIQUE
BRAU J	CENTRE DE THERMIQUE DE LYON –Thermique du bâtiment
BRISAU M	GENIE ELECTRIQUE ET FERROELECTRIQUE
BRUNET M	MECANIQUE DES SOLIDES
BRUNIE L	INGENIERIE DES SYSTEMES D'INFORMATION
BUREAU JC	THERMODYNAMIQUE APPLIQUEE
CAVILLE JY	GEMPPM*
CHANTE JP	CEGELY**** - Composants de puissance et applications
CHOCAT B	UNITE DE RECHERCHE EN GENIE CIVIL – Hydrologie urbaine
COUSIN M	UNITE DE RECHERCHE EN GENIE CIVIL – Structures
DOUTHEAU A	CHIMIE ORGANIQUE
DUFOUR R	MECANIQUE DES STRUCTURES
DUPUY JC	PHYSIQUE DE LA MATIERE
EMPTOZ H	RECONNAISSANCE DES FORMES ET VISION
ESNOUF C	GEMPPM*
EYRAUD L (Prof. émérite)	GENIE ELECTRIQUE ET FERROELECTRIQUE
FANTOZZI G	GEMPPM*
FAVREL J	PRISMa – PProductique et Informatique des Systèmes Manufacturiers
FAYARD JM	BIOLOGIE FONCTIONNELLE, INSECTES ET INTERACTIONS
FAYET M	MECANIQUE DES SOLIDES
FERRARIS-BESSO G	MECANIQUE DES STRUCTURES
FLAMAND L	MECANIQUE DES CONTACTS
FLEISCHMANN P	GEMPPM*
FLORY A	INGENIERIE DES SYSTEMES D'INFORMATION
FOUGERES R	GEMPPM*
FOUQUET R	GEMPPM*
FRECON L	INFORMATIQUE
GERARD JF	MATERIAUX MACROMOLECULAIRES
GIMENEZ G	CREATIS**
GONNARD P	GENIE ELECTRIQUE ET FERROELECTRIQUE
GONTRAND M	GEGELY**** - Composants de puissance et applications
GOUTTE R (Prof. émérite)	CREATIS **
GRANGE G	GENIE ELECTRIQUE ET FERROELECTRIQUE
GUENIN G	GEMPPM*
GUICHARDANT M	BIOCHIMIE ET PHARMACOLOGIE
GUILLOT G	PHYSIQUE DE LA MATIERE
GUINET A	PRISMa – PProductique et Informatique des Systèmes Manufacturiers
GUYADER JL	VIBRATIONS ACOUSTIQUES
GUYOMAR D	GENIE ELECTRIQUE ET FERROELECTRIQUE
JACQUET-RICHARDET G	MECANIQUE DES STRUCTURES
JOLION JM	RECONNAISSANCE DES FORMES ET VISION
JULLIEN JF	UNITE DE RECHERCHE EN GENIE CIVIL – Structures
JUTARD A (Prof. émérite)	AUTOMATIQUE INDUSTRIELLE
KASTNER R	UNITE DE RECHERCHE EN GENIE CIVIL – Géotechnique
KOULOUMDJIAN J	INGENIERIE DES SYSTEMES D'INFORMATION
LAGARDE M	BIOCHIMIE ET PHARMACOLOGIE
LALANNE M (Prof. émérite)	MECANIQUE DES STRUCTURES
LALLEMAND A	CENTRE DE THERMIQUE DE LYON – Energétique et thermique
LALLEMAND M (Mme)	CENTRE DE THERMIQUE DE LYON – Energétique et thermique
LAREAL P	UNITE DE RECHERCHE EN GENIE CIVIL – Géotechnique
LAUGIER A	PHYSIQUE DE LA MATIERE
LAUGIER C	BIOCHIMIE ET PHARMACOLOGIE
LEJEUNE P	GENETIQUE MOLECULAIRE DES MICROORGANISMES
LUBRECHT A	MECANIQUE DES CONTACTS
MARTINEZ Y	INGENIERIE INFORMATIQUE ET INDUSTRIELLE
MAZILLE H	PHYSICOCHIMIE INDUSTRIELLE
MERLE P	GEMPPM*

MERLIN J	GEMPPM*
MILLET JP	PHYSICOCHIMIE INDUSTRIELLE
MIRAMOND M	UNITE DE RECHERCHE EN GENIE CIVIL – Hydrologie urbaine
MOREL R	MECANIQUE DES FLUIDES
MOSZKOWICZ P	LAEPSI***
NARDON P (Prof. émérite)	BIOLOGIE FONCTIONNELLE, INSECTES ET INTERACTIONS
NAVARRO A	LAEPSI***
NOURI A (Mme)	MAPLY - MATHÉMATIQUES APPLIQUÉES DE LYON
ODET C	CREATIS**
OTTERBEIN M (Prof. émérite)	LEAPSI***
PASCAULT JP	MATERIAUX MACROMOLECULAIRES
PAVIC G	VIBRATIONS ACOUSTIQUES
PELLETIER JM	GEMPPM*
PERA J	UNITE DE RECHERCHE EN GENIE CIVIL – Matériaux
PERACHON G	THERMODYNAMIQUE APPLIQUEE
PERRIAT P	GEMPPM*
PERRIN J	ESCHIL – Equipe Sciences Humaines de l’Insa de Lyon
PINARD P (Prof. émérite)	PHYSIQUE DE LA MATIERE
PINON JM	INGENIERIE DES SYSTEMES D’INFORMATION
PLAY D	CONCEPTION ET ANALYSE DES SYSTEMES MECANIQUES
POUSIN J	MAPLY - MATHÉMATIQUES APPLIQUÉES DE LYON
PREVOT P	GRACIMP – Groupe de Recherche en Apprentissage, Coopération et Interfaces Multimodales pour la Productique
PROST R	CREATIS**
RAYNAUD M	CENTRE DE THERMIQUE DE LYON –Transferts Interfaces et Matériaux
REDARCE H	AUTOMATIQUE INDUSTRIELLE
REYNOUARD JM	UNITE DE RECHERCHE EN GENIE CIVIL – Structures
RIGAL JF	CONCEPTION ET ANALYSE DES SYSTEMES MECANIQUES
RIEUTORD E (Prof. émérite)	MECANIQUE DES FLUIDES
ROBERT-BAUDOUY J (Mme) (Prof. émérite)	GENETIQUE MOLECULAIRE DES MICROORGANISMES
ROUBY D	GEMPPM*
ROUX JJ	CENTRE DE THERMIQUE DE LYON
RUBEL P	INGENIERIE DES SYSTEMES D’INFORMATION
RUMELHART C	MECANIQUE DES SOLIDES
SACADURA JF	CENTRE DE THERMIQUE DE LYON – Transferts Interfaces et Matériaux
SAUTERAU H	MATERIAUX MACROMOLECULAIRES
SCAVARDA S	AUTOMATIQUE INDUSTRIELLE
THOMASSET D	AUTOMATIQUE INDUSTRIELLE
TROCCAZ M	GENIE ELECTRIQUE ET FERROELECTRIQUE
UNTERREINER R	CREATIS**
VELEX P	MECANIQUE DES CONTACTS
VIGIER G	GEMPPM*
VINCENT A	GEMPPM*
VUILLERMOZ PL (Prof. émérite)	MATIERE
ZIMMERMANN M.(Mme)	Equipe Développement Urbain

Directeurs de recherche C.N.R.S. :

BERTHIER Y	MECANIQUE DES CONTACTS
COTTE-PATAT N (Mme)	UNITE MICROBIOLOGIE ET GENETIQUE
FRANCIOSI P	GEMPPM*
MANDRAND MA (Mme)	UNITE MICROBIOLOGIE GENETIQUE
QUINSON JL	GEMPPM*
ROCHE A	MATERIAUX MACROMOLECULAIRES
SEGUELA R	GEMPPM*

Directeurs de recherche I.N.R.A. :

FEBVAY G	BIOLOGIE FONCTIONNELLE, INSECTES ET INTERACTIONS
GRENIER S	BIOLOGIE FONCTIONNELLE, INSECTES ET INTERACTIONS

Directeurs de recherche I.N.S.E.R.M. :

PRINGENT AF (Mme)	BIOLOGIE ET PHARMACOLOGIE
MAGNIN I (Mme)	CREATIS**

\***GEMPPM** GROUPE D’ETUDE METALLURGIE PHYSIQUE ET PHYSIQUE DES MATERIAUX

\*\* **CREATIS** CENTRE DE RECHERCHE ET D’APPLICATIONS EN TRAITEMENT DE L’IMAGE ET DU SIGNAL

\*\*\***LAEPSI** LABORATOIRE d’ANALYSE ENVIRONNEMENTALE DES PROCEDES ET SYSTEMES INDUSTRIELS

\*\*\*\***CEGELY** CENTRE DE GENIE ELECTRIQUE DE LYON

**Ecoles Doctorales et Diplômes d'Etudes Approfondies**  
habilités pour la période 1999-2003

ECOLES DOCTORALES n° code national	RESPONSABLE PRINCIPAL	CORRESPONDANT INSA	DEA INSA n° code national	RESPONSABLE DEA INSA
<u>CHIMIE DE LYON</u> (Chimie, Procédés, Environnement) EDA206	M. D. SINOUC UCBL1 04.72.44.62.63 Sec 04.72.44.62.64 Fax 04.72.44.81.60	M. P. MOSZKOWICZ 83.45 Sec 84.30 Fax 87.17	Chimie Inorganique 910643	M. J.F. QUINSON Tél 83.51 Fax 85.28
			Sciences et Stratégies Analytiques 910634	
			Sciences et Techniques du Déchet 910675	M. P. MOSZKOWICZ Tél 83.45 Fax 87.17
<u>ECONOMIE, ESPACE ET MODELISATION DES COMPORTEMENTS</u> (E <sup>2</sup> MC) EDA417	M.A. BONNAFOUS LYON 2 04.72.72.64.38 Sec 04.72.72.64.03 Fax 04.72.72.64.48	Mme M. ZIMMERMANN 84.71 Fax 87.96	Villes et Sociétés 911218	Mme M. ZIMMERMANN Tél 84.71 Fax 87.96
			Dimensions Cognitives et Modélisation 992678	M. L. FRECON Tél 82.39 Fax 85.18
<u>ELECTRONIQUE, ELECTROTECHNIQUE, AUTOMATIQUE</u> (E.E.A.) EDA160	M. G. GIMENEZ INSA DE LYON 83.32 Fax 85.26		Automatique Industrielle 910676	M. M. BETEMPS Tél 85.59 Fax 85.35
			Dispositifs de l'Electronique Intégrée 910696	M. D. BARBIER Tél 85.47 Fax 60.81
			Génie Electrique de Lyon 910065	M. J.P. CHANTE Tél 87.26 Fax 85.30
			Images et Systèmes 992254	Mme I. MAGNIN Tél 85.63 Fax 85.26
<u>EVOLUTION, ECOSYSTEME, MICROBIOLOGIE, MODELISATION</u> (E2M2) EDA403	M. J.P. FLANDROIS UCBL1 04.78.86.31.50 Sec 04.78.86.31.52 Fax 04.78.86.31.49	M. S. GRENIER 79.88 Fax 85.34	Analyse et Modélisation des Systèmes Biologiques 910509	M. S. GRENIER Tél 79.88 Fax 85.34
<u>INFORMATIQUE ET INFORMATION POUR LA SOCIETE</u> (EDIIS) EDA 407	M. J.M. JOLION INSA DE LYON 87.59 Fax 80.97		Documents Multimédia, Images et Systèmes d'Information Communicants 992774	M. A. FLORY Tél 84.66 Fax 85.97
			Extraction des Connaissances à partir des Données 992099	M. J.F. BOULICAUT Tél 89.05 Fax 87.13
			Informatique et Systèmes Coopératifs pour l'Entreprise 950131	M. A. GUINET Tél 85.94 Fax 85.38
INTERDISCIPLINAIRE SCIENCES- SANTÉ (EDISS) EDA205	M. A.J. COZZONE UCBL1 04.72.72.26.72 Sec 04.72.72.26.75 Fax 04.72.72.26.01	M. M. LAGARDE 82.40 Fax 85.24	Biochimie 930032	M. M. LAGARDE Tél 82.40 Fax 85.24
<u>MATERIAUX DE LYON</u> UNIVERSITE LYON 1 EDA 034	M. J. JOSEPH ECL 04.72.18.62.44 Sec 04.72.18.62.51 Fax 04.72.18.60.90	M. J.M. PELLETIER 83.18 Fax 84.29	Génie des Matériaux : Microstructure, Comportement Mécanique, Durabilité 910527	M. J.M.PELLETIER Tél 83.18 Fax 85.28
			Matériaux Polymères et Composites 910607	M. H. SAUTEREAU Tél 81.78 Fax 85.27
			Matière Condensée, Surfaces et Interfaces 910577	M. G. GUILLOT Tél 81.61 Fax 85.31
<u>MATHEMATIQUES ET INFORMATIQUE FONDAMENTALE</u> (Math IF) DA 409	M. NICOLAS UCBL1 04.72.44.83.11 Fax 04.72.43.00.35	M. J. POUSIN 88.36 Fax 85.29	Analyse Numérique, Equations aux dérivées partielles et Calcul Scientifique 910281	M. G. BAYADA Tél 83.12 Fax 85.29
<u>MECANIQUE, ENERGETIQUE, GENIE CIVIL, ACOUSTIQUE</u> (MEGA) EDA162	M. J. BATAILLE ECL 04.72.18.61.56 Sec 04.72.18.61.60 Fax 04.78.64.71.45	M. G.DALMAZ 83.03 Fax 04.72.89.09.80	Acoustique 910016	M. J.L. GUYADER Tél 80.80 Fax 87.12
			Génie Civil 992610	M. J.J.ROUX Tél 84.60 Fax 85.22
			Génie Mécanique 992111	M. G. DALMAZ Tél 83.03 Fax 04.78.89.09.80
			Thermique et Energétique 910018	Mme. M. LALLEMAND Tél 81.54 Fax 60.10

En Grisé : Les Ecoles Doctorales et DEA dont l'INSA est établissement principal



# REMERCIEMENTS

Ce travail de thèse a été réalisé au sein du Centre de Recherche Et d'Application en Traitement de l'Image et du Signal (CREATIS), unité mixte (UMR CNRS 5515), Institut National des Sciences Appliquées de Lyon (INSA Lyon), Université Claude Bernard Lyon I (UCBL), affiliée à l'INSERM. Je remercie Messieurs les professeurs Gérard GIMENEZ et Didier REVEL, codirecteurs de CREATIS lors de mon arrivée, pour l'accueil qu'ils m'ont réservé dans le laboratoire.

Que Monsieur Rémy PROST, codirecteur (principal) de thèse trouve ici mes plus vifs remerciements pour son aide précieuse, ses conseils éclairés et pour toute la confiance qu'il m'a accordée pendant ces trois années.

Que Madame Isabelle MAGNIN, désormais codirecteur de CREATIS avec Didier REVEL, et codirecteur de thèse reçoive toute ma gratitude pour ses remarques précieuses et son soutien manifesté au laboratoire.

Mes remerciements s'adressent aussi à Monsieur Jean-Marc CHASSERY de m'avoir fait l'honneur de présider le jury de thèse, Madame Françoise PRETEUX et Monsieur Michel BARLAUD d'avoir accepté d'être rapporteurs de cette thèse, et Monsieur Ho Youl JUNG d'avoir accepté de participer au jury.

Je tiens aussi à remercier tous les membres du laboratoire CREATIS pour leurs conseils et aides multiples qui m'ont été très précieux durant ces trois années.

Enfin, je dédie ce mémoire à mes proches, mes amis, pour leur soutien et tous les moments partagés.



*A ma famille,  
à mes amis.*



# Table des matières

<b>INTRODUCTION</b> .....	<b>15</b>
<b>MAILLAGES SURFACIQUES 3D MULTIRÉSOLUTION POUR LA COMPRESSION OU LA TRANSMISSION PROGRESSIVE</b> .....	<b>17</b>
<b>1 Motivations</b> .....	<b>18</b>
<b>2 Etat de l'art</b> .....	<b>22</b>
2.1 Analyse multirésolution .....	22
2.1.1 Approches « fin à grossier ».....	23
2.1.2 Approches « grossier à fin ».....	24
a) Principe de la décomposition en ondelettes : .....	25
b) Choix des ondelettes : .....	27
c) Construction de la meilleure approximation .....	29
d) Application numérique .....	33
2.1.3 Méthodes de remaillage .....	36
2.2 Compression de maillages.....	39
2.2.1 Compression de maillages à résolution unique .....	40
a) Approches basées faces .....	40
b) Approches basées sommets .....	41
2.2.2 Compression multirésolution des maillages.....	41
<b>3 Notre proposition</b> .....	<b>42</b>
3.1 Objectif.....	42
3.2 Proposition d'une subdivision irrégulière .....	43
3.3 Le problème direct (subdivision) .....	44
3.4 Le problème inverse (fusion) .....	45
3.5 Algorithme de fusion de triangles .....	48
3.6 Définition d'un nouveau produit scalaire pour la construction du banc de filtres	54
3.7 Schéma multirésolution en arithmétique entière .....	58
3.7.1 Analyse - synthèse exacte avec le banc de filtres « lazy ».....	58
3.7.2 Analyse - synthèse exacte avec le lifting .....	59
3.8 Compression sans perte .....	61
3.8.1 Compression de la connectivité du maillage.....	61
3.8.2 Compression de la géométrie du maillage .....	64
3.9 Résultats .....	65
3.10 Conclusion et perspectives .....	71

# MAILLAGES ACTIF POUR LA SEGMENTATION ET LE SUIVI D'OBJETS VIDEO

---

<b>1</b>	<b>Introduction .....</b>	<b>74</b>
1.1	Historique .....	74
1.2	Le standard MPEG-4 .....	74
1.3	Le standard MPEG-7 .....	75
1.4	Objectif du travail.....	76
<b>2</b>	<b>Etat de l'art .....</b>	<b>79</b>
2.1	Segmentation spatiale.....	79
2.2	Estimation de mouvement .....	80
2.2.1	<i>Approche différentielle : flux optique</i> .....	80
a)	Principe.....	80
b)	Problème d'ouverture .....	81
c)	Limitations de l'approche différentielle.....	81
2.2.2	Approche par mise en correspondance.....	82
2.2.3	Mise en correspondance de blocs.....	85
2.2.4	Maillages dynamiques.....	86
a)	Grilles quadrangulaires .....	86
b)	Maillages triangulaires .....	87
c)	Maillages adaptés au contenu ou « maillages actifs ».....	88
d)	Maillage « objet ».....	92
e)	Approche par « lignes de rupture » .....	94
f)	Méthodes numériques d'estimation du mouvement par maillage.....	95
g)	Problèmes liés à l'estimation de mouvement .....	99
2.2.5	Analyse et segmentation de séquences vidéo.....	101
a)	Segmentation par le mouvement .....	101
b)	Segmentation selon un modèle prédéfini .....	101
c)	Segmentation par détection de changement .....	102
d)	Suivi d'objets vidéo.....	103
<b>3</b>	<b>Notre proposition .....</b>	<b>105</b>
3.1	Stratégie de construction .....	105
3.2	Déformation du maillage.....	108
3.2.1	Principe.....	108
3.2.2	Critère de mouvement .....	111
3.2.3	Critère spatial .....	117
3.2.4	Critère de régularisation temporelle .....	120
3.2.5	Minimisation du critère d'énergie .....	122
3.2.6	Normalisation et paramètres de pondération.....	124
3.3	Stratégie de découpage.....	124

3.4	Méthodes de remaillage.....	125
3.4.1	Critère de contraction d'arête.....	125
3.4.2	Critère de permutation d'arête.....	126
3.5	Prédiction du maillage d'une image à l'autre.....	128
<b>4</b>	<b>Résultats .....</b>	<b>129</b>
4.1	Minimisation d'énergie .....	129
4.2	Séquence synthétique .....	132
4.3	Séquences réelles.....	138
<b>5</b>	<b>Conclusion et perspectives .....</b>	<b>144</b>
<b>BILAN ET PERSPECTIVES.....</b>		<b>146</b>
<b>1</b>	<b>Bilan et perspectives .....</b>	<b>147</b>
<b>2</b>	<b>Vers une représentation multirésolution des objets vidéo : méthode proposée.....</b>	<b>147</b>
<b>3</b>	<b>Résultats expérimentaux .....</b>	<b>150</b>
<b>4</b>	<b>Conclusion et perspectives .....</b>	<b>150</b>
<b>BIBLIOGRAPHIE .....</b>		<b>152</b>



## Introduction

Dès la fin du 20<sup>e</sup> Siècle, les progrès conjoints de l'informatique et des télécommunications ont véritablement changé les habitudes de l'être humain. Force est de constater que la vision de l'an 2000 qu'avait en tête la population du début du siècle ne s'est pas concrétisée : il n'y a pas encore de voitures volantes, ni de robots véritablement intelligents. Mais l'émergence d'outils tels que le téléphone sans fil et le réseau mondial Internet ont véritablement révolutionné les comportements. Actuellement, les informations communiquées en temps réel sont pour la plupart vocales ou écrites, et l'augmentation du débit des canaux de transmission et des capacités de traitement des processeurs de signaux numériques permet l'enrichissement du contenu. La notion de *contenu multimédia* fait référence à un document englobant plusieurs types d'informations : sons, images fixes ou animées, réelles ou synthétiques, voix, données. La nécessité d'analyser, gérer et transmettre ces informations a donné naissance au traitement numérique des signaux, dont les applications ont dépassé le cadre des télécommunications. Les images ne sont plus uniquement stockées ou transmises, mais aussi analysées. Ainsi, le traitement d'images apporte une aide dans de nombreux domaines tels que les prévisions météorologiques, le diagnostic médical, les systèmes automatisés, la représentation des informations.

Ce mémoire de recherche porte sur l'utilisation des structures à éléments finis, ou maillages, pour le traitement des images tridimensionnelles et des images animées 2D.

### **Structure du document :**

Dans le chapitre II, nous abordons l'analyse multirésolution des surfaces 3D représentées par des maillages triangulaires. Les maillages 3D sont employés dans l'infographie, les simulations mécaniques, l'imagerie médicale, et dans beaucoup d'applications nécessitant une visualisation tridimensionnelle des données. Lounsbery a proposé en 1994 la décomposition en ondelettes des maillages pour l'analyse multirésolution, mais une contrainte forte limitait son approche aux maillages possédant une connectivité particulière : connectivité de subdivision régulière. Nous proposons une extension du formalisme existant, augmentant ainsi le rayon d'application de la décomposition des maillages 3D sur une base d'ondelette aux maillages à connectivité quelconque. Notre innovation porte sur l'introduction d'un schéma de subdivision irrégulier qui permet la création, à partir d'un maillage quelconque, d'une hiérarchie de maillages de résolution inférieure, approchant le maillage selon plusieurs niveaux de détails, par décomposition des coordonnées des sommets sur une base d'ondelettes. La construction d'une telle hiérarchie prend la forme d'un problème inverse, dont nous proposons la résolution par un algorithme déterministe de conquête. Une telle représentation multirésolution permet une transmission progressive des maillages, de l'approximation grossière à la reconstruction parfaite, et le codage compact des informations de reconstruction conduit à une compression efficace du volume des données. L'efficacité de

l'algorithme est démontrée sur plusieurs exemples, allant du maillage synthétique aux triangulations construites sur des données médicales réelles.

Dans le chapitre III, nous développons une approche maillage 2D pour la segmentation et le suivi d'objets dans des séquences vidéo, dans le but d'enrichir la décomposition sémantique des séquences, connaissance nécessaire pour l'utilisation du standard audio-visuel émergent : MPEG-4. Ce nouveau format introduit la notion d'objet vidéo dans les séquences, dont la forme et la localisation est fournie par une carte de segmentation (un masque) appelée plan alpha. Cependant la segmentation automatique des objets reste à l'heure actuelle un problème ouvert. Nous proposons dans cette section l'utilisation de maillages triangulaires 2D dynamiques pour la segmentation et le suivi des objets, réalisant ainsi une segmentation automatique ou semi-automatique. Notre algorithme construit un maillage actif, dont les arêtes sont placées sur les discontinuités de mouvement et les discontinuités spatiales, pour séparer au mieux les objets présents dans la scène. Une estimation de mouvement bidirectionnelle réduit les difficultés rencontrées lors des occultations et les découvements des différentes régions de l'image. Une contrainte de régularité temporelle, imposée à la segmentation, stabilise l'évolution des contours entre les objets. Une évaluation quantitative de la qualité de la segmentation sur des séquences synthétiques est effectuée afin de démontrer l'importance du choix de la modélisation du mouvement introduite dans notre algorithme. Notre algorithme est évalué à l'aide de séquences d'images par résonance magnétique nucléaire et de séquences vidéo diverses.

Le chapitre IV apporte une conclusion et ouvre des perspectives sur la représentation multirésolution des contours des objets vidéo par des maillages 2D triangulés est abordée. Une adaptation de l'approche proposée dans le chapitre II y est proposée et illustrée par un exemple concret.

**Chapitre II**

**Maillages surfaciques 3D multirésolution  
pour la compression ou la transmission  
progressive**

## ***1*** **MOTIVATIONS**

Cette partie traite d'un domaine encore récent du traitement numérique des images : les surfaces tridimensionnelles représentées par des maillages. En effet, la quantité d'information contenue dans un volume formé d'objets de densité homogène (représenté par un ensemble de voxels) peut être énormément réduite en modélisant ce volume par un ensemble de surfaces représentant les frontières entre les différents milieux présents dans le volume. A titre d'exemple, le volume de la figure 1.1 représentant une partie de vertèbre est de dimension  $411*180*511$  peut être mémorisé sous deux formes différentes :

- une description volumique sous forme de voxels pesant  $411*180*511=37\ 803\ 780$  bits soit 4,51 Mo
- une description surfacique de la frontière entre l'os et l'extérieur sous forme de maillage triangulaire. Le maillage équivalent, construit à l'aide de l'algorithme 'marching cubes' [LORE87] possède 9478 faces et 4741 sommets. La quantité d'information nécessaire pour représenter ce maillage est de 492 908 bits soit 60,2 Ko.



Figure 1 : partie d'un os humain.

Le même volume peut donc être représenté très efficacement sous forme de maillage, ce qui diminue la place qu'il prendrait s'il devait être stocké, mais peut aussi être plus facilement manipulé et affiché à l'écran, du fait du nombre d'opération nécessaires considérablement réduit.

Un maillage est un ensemble de points sommets reliés entre eux par des arêtes. Plusieurs caractéristiques les distinguent :

- Si les sommets sont tous situés dans un même plan, le maillage sera 2D. Dans le cas contraire, ce sera un maillage 3D.
- Les polyèdres formés peuvent être de différents ordres. Si tous les polyèdres sont des triangles, on parlera alors de maillage triangulaire (Figure 1.2.a). Si tous les polyèdres sont des quadrangles, le maillage est dit quadrangulaire (figure 1.2.b). Dans [DELI99] Delingette introduit la notion de maillage simplexe, où tous les sommets possèdent 3 sommets voisins, à l'exception des sommets situés aux bords. Nos travaux ont principalement été effectués sur les maillages triangulaires, qui sont les plus répandus actuellement.
- Un maillage est dit conforme (le terme anglais est 'manifold') si aucun des sommets qui le composent n'est situé sur une arête du même maillage (voir figure 1.3)

Les maillages triangulaires sont parfois construits en respectant le critère de Delaunay [FREY99], qui veut que tous les cercles circonscrits aux triangles formant le maillage considéré ne contiennent aucun sommet.

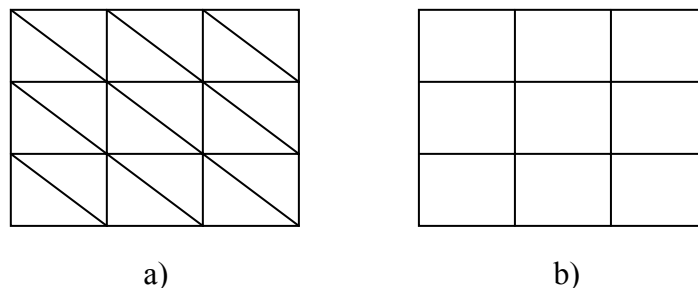


Figure 1.2 : maillages 2D triangulaires (a) et quadrangulaires (b)

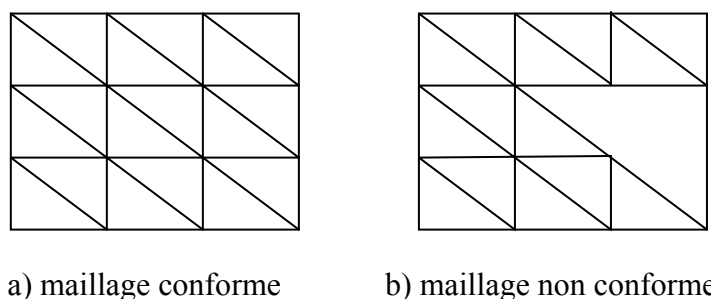


Figure 1.3 : conformité des maillages

La représentation d'un maillage contient deux types d'information :

- des informations géométriques : les coordonnées des sommets
- des informations topologiques concernant la connectivité du maillage c'est à dire la description des faces du maillages, d'où peuvent être déduites toutes les relations de voisinage entre les sommets.

Pour un maillage 3D triangulaire  $M$  comportant  $f$  faces et  $s$  sommets, dont les coordonnées sont inscrites dans volume de taille  $L^3$  et représentées par des entiers , il faudra :

- $3s \log_2(L)$  bits pour stocker les coordonnées des sommets (3 entiers par sommet)
- $3f \log_2(s)$  bits pour stocker la connectivité du maillage (3 indices de sommets par face).

Il faudra donc au total  $C(M) = 3s \log_2(L) + 3f \log_2(s)$  bits pour stocker le maillage.

En utilisant l'équation d'Euler [FREY99] qui, pour un maillage  $M$ , triangulaire, conforme, ayant  $s$  sommets,  $f$  faces et  $a$  arêtes, donne la propriété suivante:

$$f - a + s = 1 \quad \text{II-1}$$

lorsque le maillage a peu de bords, on a l'approximation suivante [ROSS99] :

$$f \approx \frac{2}{3}a \approx 2s \quad \text{II-2}$$

En utilisant ce résultat, on peut estimer le coût global de codage sans compression pour un maillage triangulaire 3D qui vaut alors :

$$\begin{aligned} C(M) &= 3s \log_2(L) + 3f \log_2(s) = \\ C(M) &= f \left( \frac{3}{2} \log_2(L) + 3 \log_2\left(\frac{f}{2}\right) \right) \end{aligned} \quad \text{II-3}$$

On peut alors en déduire que le coût de codage d'un maillage n'est pas linéaire en fonction du nombre de faces, ce qui revient à dire que le coût de codage par face n'est pas constant (voir figure 1) , contrairement aux images 2D ou 3D, où le coût de codage par pixel ou voxel est, lui, constant, quelle que soit la taille de l'image ou du volume.

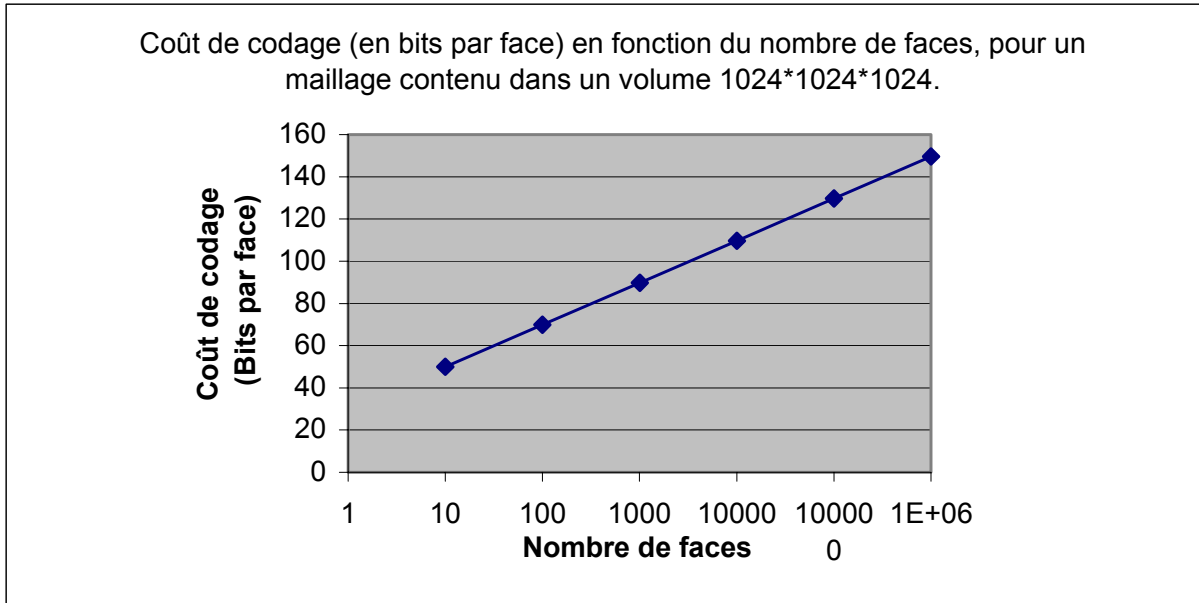


Figure 1 : coût de codage par face

Avec les progrès de l'informatique et les besoins grandissant dans les domaines de la modélisation, de l'infographie, ou de l'imagerie médicale, les maillages tendent à contenir de plus en plus de faces et de sommets. Il n'est plus rare maintenant de rencontrer des maillages contenant plusieurs centaines de milliers de faces, et cette tendance va en s'amplifiant.

Deux axes de recherche apparaissent alors :

- permettre une compression efficace du volume de données nécessaire à la description des maillages, garantissant un coût de codage (en bits par face) le plus faible possible, et constant avec le nombre de faces.
- à partir d'un maillage original existant, générer plusieurs maillages à divers niveaux de résolution (nombre de faces réduit), permettant une représentation simplifiée du maillage original, afin d'accélérer certaines tâches comme le rendu graphique, la transmission sur un canal de communication ou la modélisation par éléments finis.

## 2 ETAT DE L'ART

### 2.1 Analyse multirésolution

L'analyse multirésolution appliquée sur les maillages peut se résumer ainsi : soit un maillage  $M^n$  de résolution élevée, l'algorithme utilisé doit pouvoir générer plusieurs maillages  $M^j$  ayant des niveaux de résolution de plus en plus bas, comme indiqué sur la figure 2.1.1 .

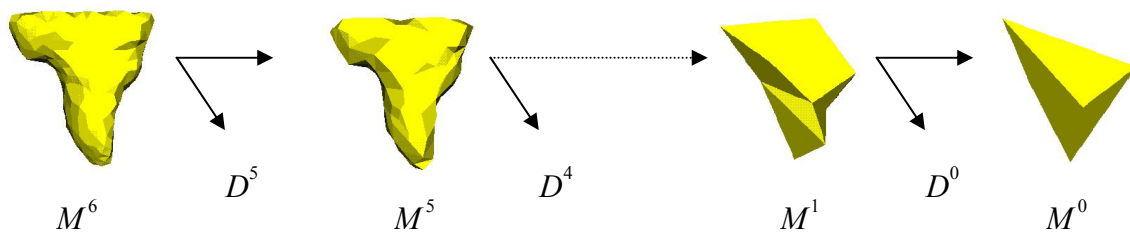


Figure 2.1.1 : simplification de maillages

Réciproquement, la méthode utilisée se doit d'être réversible, ceci permettant, à partir du maillage de résolution la plus basse  $M^0$  de reconstruire tous les niveaux de résolution, jusqu'au maillage de résolution la plus haute  $M^n$ , par la connaissance des détails perdus  $D^i$  lors des simplifications successives, comme le montre la figure 2.1.2 .

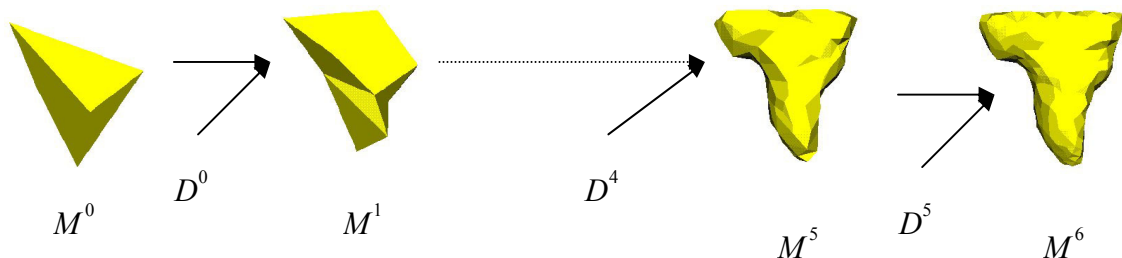


Figure 2.1.2 : reconstruction des différents niveaux de résolution

Pour une bibliographie exhaustive des algorithmes existant, le lecteur pourra se référer à [HECK97]. On peut distinguer deux types d'analyse multirésolution sur les maillages 3D triangulaires :

- celles utilisant l'approche « fin à grossier » (« fine to coarse » en anglais) pouvant être appliquées sur tous les maillages existant, quelle que soit leur origine
- celles utilisant l'approche « grossier à fin » (« coarse to fine » en anglais) imposant des contraintes de connectivité aux maillages existants, et nécessitant des outils appropriés pour créer des maillages adaptés à ces algorithmes.

### 2.1.1 APPROCHES « FIN A GROSSIER »

Les algorithmes employant une approche fin à grossier simplifient récursivement le maillage haute résolution de manière à faire diminuer le nombre de faces et de sommets le composant. On procède alors à une décimation de maillage. Plusieurs opérations peuvent être envisagées, telles que:

- Des suppressions de sommets, suivies d'un remaillage des zones altérées comme décrit dans [SOUC96] (figure 2.1.3). Les sommets supprimés sont choisis de manière à minimiser l'écart entre le maillage simplifié et le maillage original.
- Des contractions d'arête successives, comme utilisées dans l'algorithme « Progressive meshes » [HOPP96]. Dans cette opération, deux faces voisines d'un maillage sont supprimées, et deux sommets sont fusionnés (figure 2.1.4). Partant de ce type de décimation, Guskov [GUSK99] propose une méthode de traitement multirésolution pour les maillages, permettant un filtrage des maillages que l'on peut apparenter aux méthodes de traitement multirésolution connues dans les domaines du traitement du signal et de l'image.
- Des suppressions de faces ou de groupe de faces. On peut citer par exemple l'algorithme « Progressive Forest Split » [TAUB98b], qui à chaque opération élémentaire supprime 4 faces et fusionne trois sommets en un (figure 2.1.5).

Pour les méthodes citées ci-dessus, les simplifications sont généralement effectuées de manière à minimiser l'erreur de simplification entre le maillage original et le maillage simplifié, selon une métrique choisie, ce qui conduit généralement à simplifier les zones de maillage possédant la courbure la plus faible possible.

Il faut remarquer que les algorithmes de décimation de maillages ne sont pas toujours facilement inversibles, ce qui limite leur intérêt dans une utilisation multirésolution où le maillage haute résolution doit être reconstruit [SOUC96].

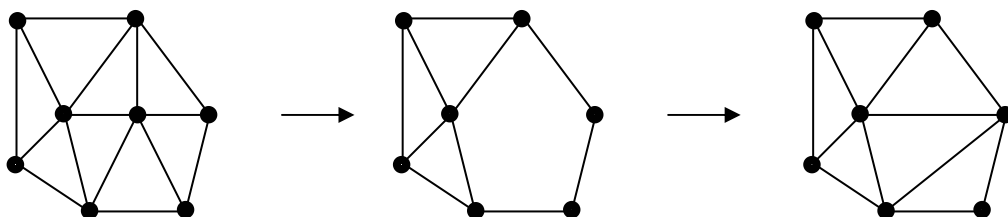


Figure 2.1.3 : suppression de sommet suivie d'un remaillage

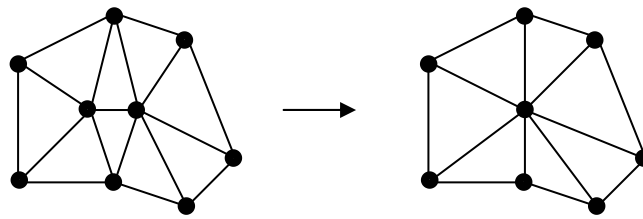


Figure 2.1.4 : contraction d'arête

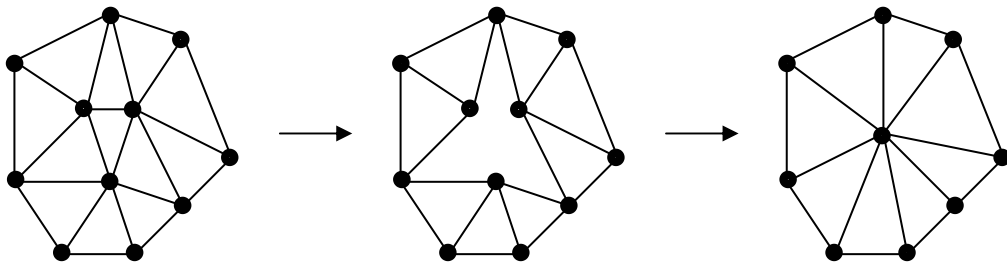


Figure 2.1.5 : suppression de 4 faces et fusion de trois sommets

### 2.1.2 APPROCHES « GROSSIER A FIN »

Avec ces approches, les opérations élémentaires utilisées augmentent le nombre de sommets et de faces d'un maillage, augmentant par conséquent la qualité de celui-ci. Cette notion est étroitement liée à la subdivision de surfaces, qui permet d'accroître la résolution des maillages.

La subdivision de surfaces a été initialement utilisée pour lisser l'aspect visuel des maillages [CATM78] [DOO78] [LOOP87] [DYN90] [ZORI97] : Un maillage de topologie arbitraire  $M^0$  est subdivisé, ce qui augmente son nombre de sommets. Les sommets subissent ensuite un post-traitement dépendant du type de subdivision, lissant le maillage (d'où le terme anglais « smoothing »), comme le montre l'exemple de la figure 2.1.6 .

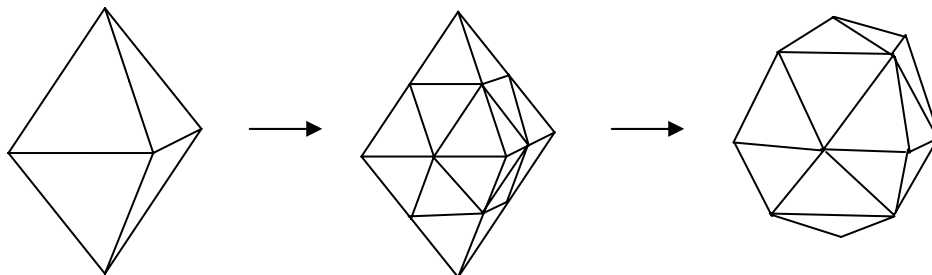


Figure 2.1.6 : exemple de subdivision de maillage suivie d'un lissage.

De plus, les subdivisions de surfaces permettent la représentation du maillage sous forme de fonctions d'échelles susceptibles d'être raffinées [STOL96], propices à une analyse multirésolution par décomposition en ondelettes, introduite initialement par Meyer et Mallat [MEYE93] [MALL89]. Lounsbery a ainsi formalisé la décomposition en ondelettes appliquée aux maillages triangulaires, permettant une analyse multirésolution sur des maillages triangulaires [LOUN94]. Nous allons par la suite décrire cette approche plus en détail, car notre contribution s'appuie principalement sur celle-ci.

a) Principe de la décomposition en ondelettes :

On considère tout d'abord un ensemble emboîté d'espaces vectoriels :

$$V^0 \subset V^1 \subset V^2 \subset \dots \subset V^{n-2} \subset V^{n-1} \subset V^n \quad \text{II-4}$$

La résolution des fonctions de l'espace vectoriel  $V^j$  croît avec  $j$ . Les fonctions de base  $\varphi_i^j(x)$  pour  $V^j$  sont appelées fonctions d'échelle. On définit ensuite pour chaque espace vectoriel  $V^j$  l'espace vectoriel  $W^j$  complément de  $V^j$  dans  $V^{j+1}$  les relations suivantes apparaissent :

$$V^j \oplus W^j = V^{j+1} \quad \text{II-5}$$

$$V^j \cap W^j = \emptyset \quad \text{II-6}$$

Les fonctions de base  $\Psi_i^j(x)$  pour  $W^j$  sont appelées ondelettes. Pour qu'une fonction de l'espace  $V^j$  soit la meilleure approximation d'une fonction de l'espace  $V^{j+1}$ , les ondelettes doivent toutes être orthogonales à toutes les fonctions d'échelle de  $V^j$ , selon un produit scalaire choisi :

$$\langle \Psi_i^j(x), \varphi_k^j(x) \rangle = 0 \quad \forall i, k \quad \text{II-7}$$

Lounsbery a choisi de représenter la surface d'un maillage triangulaire sous forme de B-splines d'ordre 1 : les fonctions d'échelles  $\Phi_i^j(x)$  sont des fonctions chapeau, définies sur un espace fixé par la topologie du maillage considéré. Pour un maillage  $M^j$  comportant  $s^j$  sommets, représenté par une fonction  $S^j$ , l'espace vectoriel  $V^j$  correspondant aura  $s^j$  fonctions d'échelles, prenant chacune la valeur 1 sur un sommet du maillage, et 0 sur tous les autres. La figure 2.1.7 montre un exemple de fonction d'échelle définie sur un maillage.

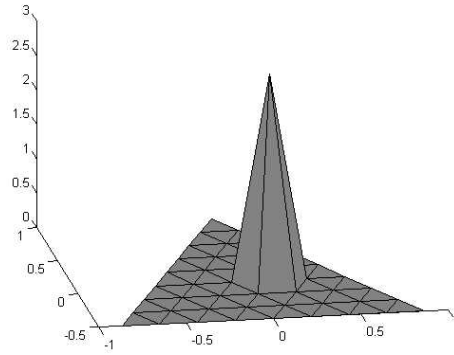


Figure 2.1.7 : une fonction d'échelle (B-spline d'ordre 1)

Le choix d'une telle représentation est tout à fait en adéquation avec la représentation par maillage, puisque la surface des maillages triangulaires est totalement linéaire entre les sommets qui les constituent.

L'utilisation de telles fonctions d'échelle permet une représentation mathématique de la surface  $S^j(x)$  d'un maillage  $M^j$  (hormis sa connectivité) sur cette base par une matrice  $C^j$  de taille  $(s^j \times 3)$  contenant les coordonnées des sommets du maillage:

$$S^j(x) = \Phi^j(x)C^j \quad \text{II-8}$$

avec :

$$\Phi^j(x) = [\varphi_0^j(x) \dots \varphi_{s^j}^j(x)] \quad \text{II-9}$$

$$C^j = \begin{bmatrix} x_0 & y_0 & z_0 \\ x_1 & y_1 & z_1 \\ \dots & \dots & \dots \\ x_{s^{j-1}} & y_{s^{j-1}} & z_{s^{j-1}} \\ x_{s^j} & y_{s^j} & z_{s^j} \end{bmatrix} \quad \text{II-10}$$

L'imbrication des espaces vectoriels ( $V^j \subset V^{j+1}$ ) implique que les fonctions d'échelle de  $V^{j-1}$  peuvent être construites par combinaison linéaire des fonctions d'échelle de  $V^j$ . Il existe donc une matrice  $P^j$  de taille  $(s^j \times s^{j-1})$  satisfaisant :

$$\Phi^{j-1}(x) = \Phi^j(x)P^j \quad \text{II-11}$$

Par un raisonnement similaire sur les ondelettes ( $W^j \subset W^{j+1}$ ), les équations suivantes apparaissent :

$$\Psi^j(x) = [\psi_0^j(x) \dots \psi_{s^{j+1}-s^j}^j(x)] \quad \text{II-12}$$

$$\Psi^{j-1}(x) = \Phi^j(x)Q^j \quad \text{II-13}$$

où  $Q^j$  est une matrice de taille  $(s^j \times (s^j - s^{j-1}))$ .

Les équations précédentes peuvent être exprimées sous une seule expression avec des matrices blocs [LOUN97]:

$$\left[ \Phi^{j-1} | \Psi^{j-1} \right] = \Phi^j \left[ P^j | Q^j \right] \quad \text{II-14}$$

En décomposant  $S^j(x)$  sur les espaces  $V^j$  puis  $V^{j-1}$  et  $W^{j-1}$  il en résulte :

$$\begin{aligned} S^j(x) &= \Phi^j(x)C^j \\ S^j(x) &= \Phi^{j-1}(x)C^{j-1} + \Psi^{j-1}(x)D^{j-1} \end{aligned} \quad \text{II-15}$$

$C^{j-1}$  et  $D^{j-1}$  représentent les coefficients de décomposition du maillage  $M^j$ .  $C^{j-1}$  est la matrice des coefficients concernant les fonctions d'échelle de  $V^{j-1}$  et  $D^{j-1}$  représente les détails perdus lors du passage de l'espace  $V^j$  à  $V^{j-1}$ .

Il en résulte :

$$\begin{aligned} S^j(x) &= \Phi^j(x)C^j \\ S^j(x) &= \Phi^j(x)P^jC^{j-1} + \Phi^j(x)Q^jD^{j-1} \end{aligned} \quad \text{II-16}$$

b) Choix des ondelettes :

La subdivision du maillage augmente la résolution de celui-ci, mais permet aussi de définir la famille d'ondelettes utilisée.

Lounsbery a utilisé une subdivision de type « 1 vers 4 » : chaque face du maillage est subdivisée en quatre faces, en créant un nouveau sommet au milieu de chaque arête, comme le montre la figure 2.1.8 .

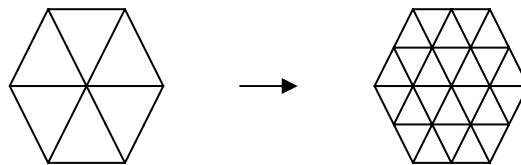


Figure 2.1.8 : subdivision de type « 1 vers 4 »

Ainsi l'augmentation de la résolution s'effectue ainsi : le maillage est subdivisé, et les sommets créés sont déplacés pour améliorer la précision du maillage, comme le montre l'exemple appliqué à une pyramide dans la figure 2.1.9 .

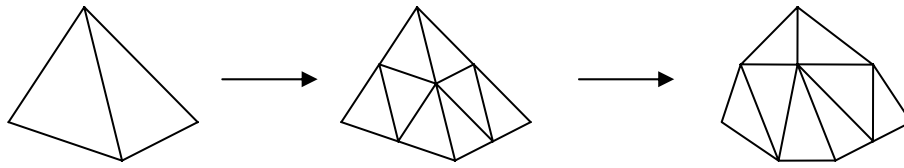


Figure 2.1.9 : augmentation de la résolution par subdivision

Ce choix nous permet de définir une famille d'ondelettes pour l'espace  $V^{j-1}$  du même type que les fonctions d'échelle pour le niveau  $V^j$  : des B-spline d'ordre 1, avec un support 4 fois plus étroit que les fonctions d'échelle. La figure 2.1.10 montre les fonctions d'échelle et les ondelettes pour le cas unidimensionnel d'un espace  $V^2$  possédant 4 fonctions d'échelles, définies sur l'intervalle  $[0 ; 1]$ .

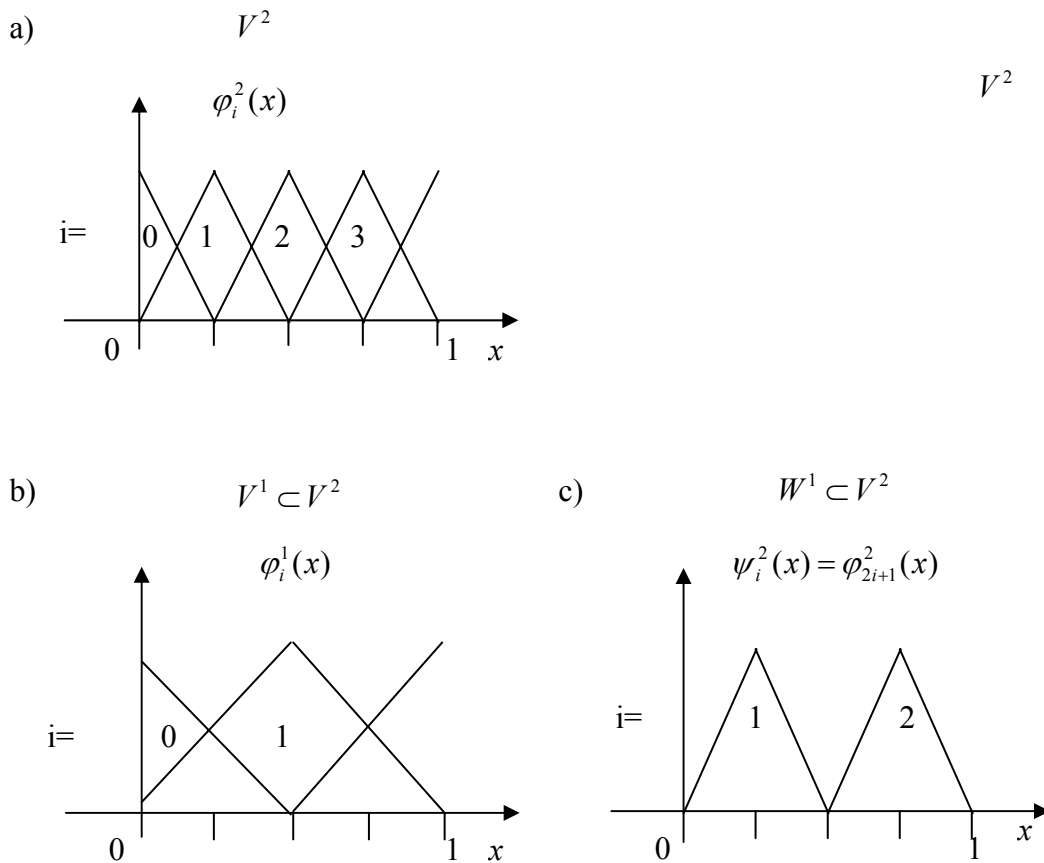


Figure 2.1.10 : fonctions d'échelle et ondelettes (cas 1D) :

- a) Fonctions d'échelles de l'espace  $V^2$
- b) Fonctions d'échelles de l'espace  $V^1$
- c) Ondelettes (fonctions de base de  $W^1$ )

Par identification dans l'équation II-16 il vient :

$$C^j = P^j C^{j-1} + Q^j D^{j-1} \quad \text{II-17}$$

La complémentarité des espaces  $V^{j-1}$  et  $W^{j-1}$  implique que les fonctions d'échelle à la résolution  $j$  sont une combinaison linéaire unique des fonctions d'échelle à la résolution  $j-1$  et des ondelettes, que l'on écrira sous la forme :

$$\Phi^j(x) = \Phi^{j-1} A^j + \Psi^{j-1}(x) B^j \quad \text{II-18}$$

Ainsi nous avons :

$$S^j(x) = \Phi^j(x) C^j = \Phi^{j-1} A^j C^j + \Psi^{j-1}(x) B^j C^j \quad \text{II-19}$$

Par identification avec l'équation II-16

$$C^{j-1} = A^j C^j \quad \text{II-20}$$

$$D^{j-1} = B^j C^j \quad \text{II-21}$$

Ainsi

$$\begin{bmatrix} A^j \\ B^j \end{bmatrix} = [P^j | Q^j]^{-1} \quad \text{II-22}$$

$A^j$  et  $B^j$  sont appelées filtres d'analyse,  $P^j$  et  $Q^j$  sont les filtres de synthèse.

Par le choix d'une subdivision de type 1:4, les matrices  $A^j$ ,  $B^j$ ,  $P^j$  et  $Q^j$  sont aisément explicables. La figure 2.1.11 illustre le cas de la subdivision appliquée sur une pyramide. Il faut remarquer que ces matrices dépendent de la connectivité des maillages traités. Il sera donc nécessaire de les construire pour chaque nouveau maillage.

### c) Construction de la meilleure approximation

L'obtention de la meilleure approximation au sens des moindres carrés de  $M^j$  par  $M^{j-1}$  est garantie si les ondelettes  $\Psi_i^{j-1}(x)$  sont orthogonales aux fonctions d'échelle  $\Phi_i^{j-1}(x)$  selon un produit scalaire choisi. Lounsbery a défini le produit scalaire entre deux fonctions  $f$  et  $g$  définies sur la surface d'un maillage  $M$  ainsi :

$$\langle f, g \rangle = \sum_{\tau \in \Delta(M)} \left( \frac{K^j}{\text{aire}(\tau)} \int_{s \in \tau} f(s) \cdot g(s) \cdot ds \right) \quad \text{II-23}$$

où  $\tau$  est un triangle,  $\Delta(M)$  l'ensemble des triangles formant le maillage  $M$  et  $K^j$  une constante pour un niveau de résolution  $j$  donné ( $K^j = 4^{-j}$ ). Il faut noter que cette définition considère que les triangles du maillage ont tous la même surface. L'algorithme utilisé sera donc d'autant plus performant que les triangles du maillage auront une surface proche.

En décomposant les fonctions  $f$  et  $g$  sur la base des fonctions d'échelle  $\Phi_i^j(x)$  :

$$f(x) = \sum_i f_i^j \varphi_i^j(x) = \Phi^j(x) F^j \quad \text{II-24}$$

$$g(x) = \sum_i g_i^j \varphi_i^j(x) = \Phi^j(x) G^j \quad \text{II-25}$$

le produit scalaire entre deux fonctions pourra s'exprimer sous la forme bilinéaire:

$$\langle f, g \rangle = (g^j)^T I^j f^j \quad \text{II-26}$$

$I^j$  est une matrice de produits scalaires, que l'on peut calculer au préalable telle que :

$$(I^j)_{i,i'} = \langle \varphi_i^j, \varphi_{i'}^j \rangle \quad \text{II-27}$$

Le calcul des coefficients de  $I^j$  est abordé plus en détail dans [VALE98]. En utilisant les relations de linéarités entre les fonctions d'échelles à deux niveaux de résolution consécutifs, la propriété suivante apparaît [LOUN94]:

$$I^{j-1} = (P^j)^T I^j P^j \quad \text{II-28}$$

Si on conserve les fonctions  $\varphi_i^{j-1}(x)$  et  $\psi_i^{j-1}(x)$  telles qu'elles sont définies ici, la condition d'orthogonalité entre ondelettes et fonctions d'échelle se traduit par :

$$(P^j)^T I^j Q^j = 0 \quad \text{II-29}$$

On vérifie facilement (figure 2.1.10) que cette contrainte n'est pas respectée. Ainsi, la famille d'ondelettes choisie n'est pas optimale. Ce résultat n'est cependant pas surprenant, car si on observe les matrices de filtrage  $A^j$  et  $B^j$ , elles ne traduisent qu'un sous-échantillonnage de la surface du maillage (en supprimant certains sommets et en gardant les coordonnées des sommets restants inchangés). Cette famille d'ondelettes est appelée « Lazy wavelets ».

Pour assurer une approximation optimale, Lounsbery a utilisé l'opération de « lifting » [SWEL96] : la base d'ondelette est modifiée de manière à rendre les ondelettes orthogonales aux fonctions d'échelle. Ceci s'effectue en soustrayant à chaque ondelette une fraction pondérée de toutes les fonctions d'échelles, pour créer une nouvelle base d'ondelettes  $\Psi_{orth}^{j-1}(x)$  telle que :

$$\Psi_{orth}^{j-1}(x) = \Psi^{j-1}(x) - \Phi^{j-1}(x).\alpha^j \quad \text{II-30}$$

où  $\alpha^j$  est la matrice des coefficients de pondération pour les fonctions d'échelle, choisie pour rendre la base d'ondelettes orthogonale à la base des fonctions d'échelle. En insérant l'équation II-39 dans l'équation d'orthogonalité (équation II-7), la matrice  $\alpha^j$  peut être explicitée :

$$\langle \varphi_i^{j-1}(x), \psi_{orth_k}^{j-1}(x) \rangle = 0 \quad \forall i, k$$

ou sous forme matricielle :

$$\Leftrightarrow \langle \Phi^{j-1T}(x) \Psi_{orth}^{j-1}(x) \rangle = 0$$

ici les éléments de la matrice  $\langle \times \times \rangle$  sont des produits scalaires entre les fonctions d'échelles  $\varphi_i^{j-1}(x)$  et les ondelettes  $\psi_k^{j-1}(x)$ . On montre facilement que

$$\alpha^j = (I^{j-1})^{-1} (P^j)^T I^j Q^j \quad \text{II-31}$$

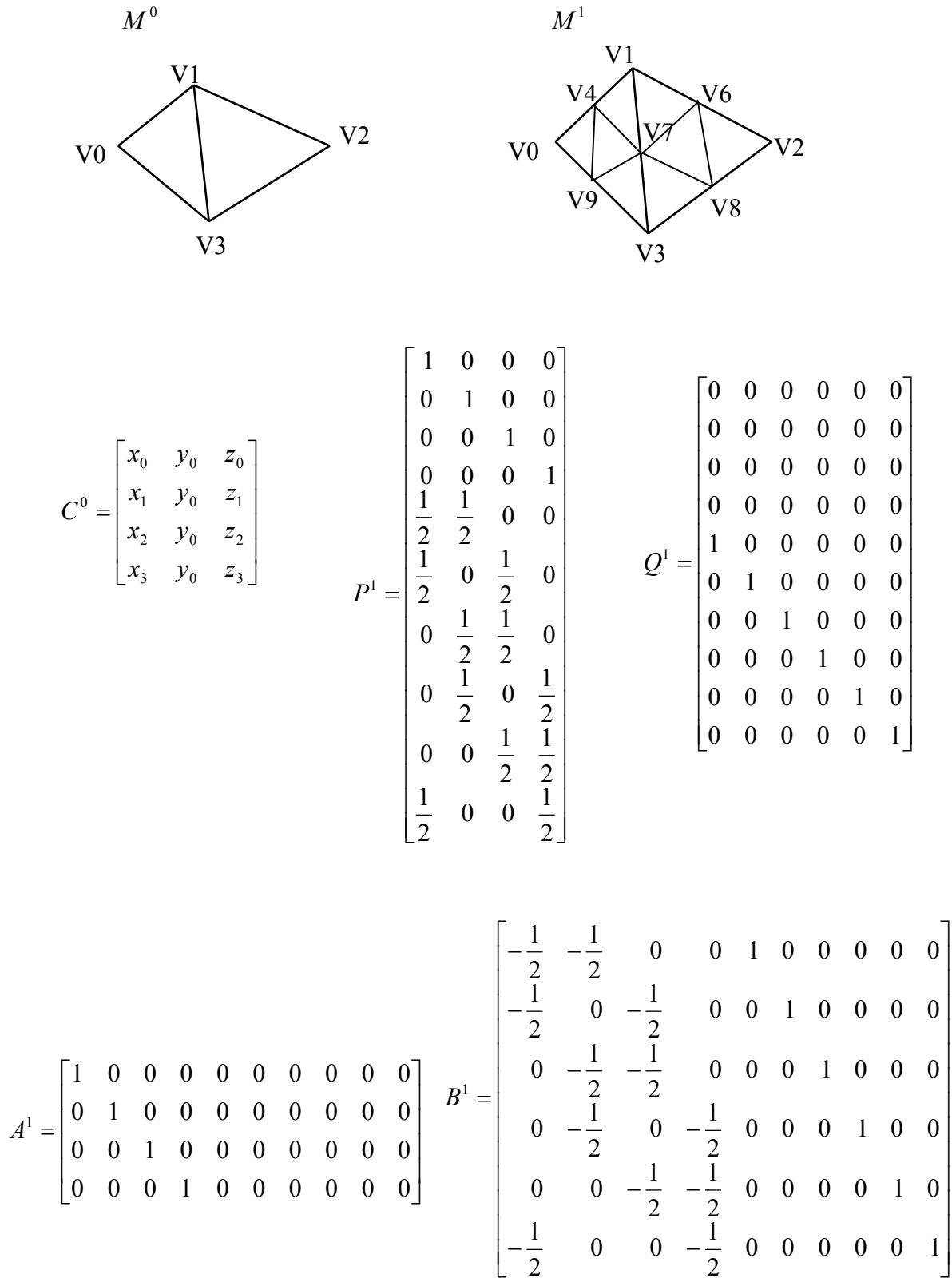


Figure 2.1.11 : filtres de synthèse et d'analyse pour une pyramide

Le « lifting » conduit alors à la création d'un nouveau banc de filtres d'analyse et de synthèse :

$$\begin{aligned}
 A_{orth}^j &= A^j + \alpha^j B^j \\
 B_{orth}^j &= B^j \\
 P_{orth}^j &= P^j \\
 Q_{orth}^j &= Q^j - P^j \alpha^j
 \end{aligned}
 \tag{II-32}$$

La figure 2.1.12 illustre l'application du lifting sur une ondelette « paresseuse » (« lazy » en anglais)

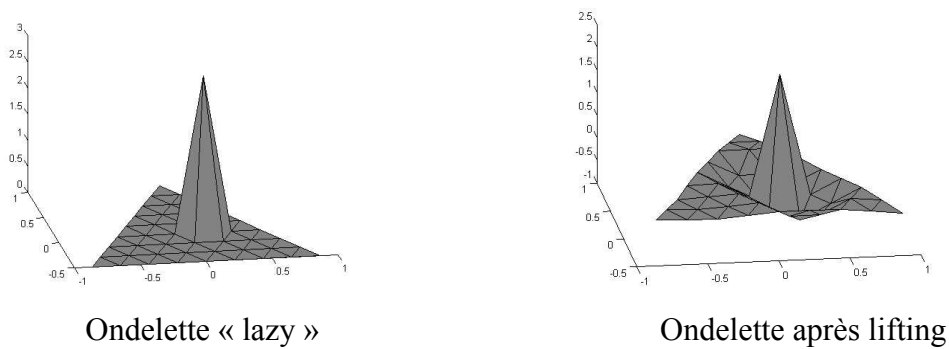


Figure 2.1.12 : application du lifting

d) Application numérique

Voici le nouveau banc de filtres orthogonalisé pour le cas de la subdivision de la pyramide :

$$\alpha^1 = \begin{bmatrix} \frac{3}{8} & \frac{3}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & \frac{3}{8} \\ \frac{3}{8} & -\frac{1}{8} & \frac{3}{8} & \frac{3}{8} & -\frac{1}{8} & -\frac{1}{8} \\ \frac{8}{8} & \frac{8}{8} & \frac{8}{8} & \frac{8}{8} & \frac{8}{8} & \frac{8}{8} \\ -\frac{1}{8} & \frac{3}{8} & \frac{3}{8} & -\frac{1}{8} & \frac{3}{8} & -\frac{1}{8} \\ -\frac{8}{8} & \frac{8}{8} & \frac{8}{8} & -\frac{8}{8} & \frac{8}{8} & \frac{8}{8} \\ -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & \frac{3}{8} & \frac{3}{8} & \frac{3}{8} \\ -\frac{8}{8} & \frac{8}{8} & \frac{8}{8} & \frac{8}{8} & \frac{8}{8} & \frac{8}{8} \end{bmatrix}$$

$$A_{orth}^1 = \begin{bmatrix} \frac{7}{16} & -\frac{1}{16} & -\frac{1}{16} & -\frac{1}{16} & \frac{3}{8} & \frac{3}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & \frac{3}{8} \\ \frac{1}{16} & \frac{7}{16} & -\frac{1}{16} & -\frac{1}{16} & \frac{3}{8} & -\frac{1}{8} & \frac{3}{8} & \frac{3}{8} & -\frac{1}{8} & -\frac{1}{8} \\ \frac{16}{16} & \frac{16}{16} & \frac{16}{16} & \frac{16}{16} & \frac{8}{8} & -\frac{8}{8} & \frac{8}{8} & \frac{8}{8} & -\frac{8}{8} & \frac{8}{8} \\ -\frac{1}{16} & \frac{1}{16} & \frac{7}{16} & -\frac{1}{16} & \frac{1}{8} & \frac{3}{8} & \frac{3}{8} & -\frac{1}{8} & \frac{3}{8} & -\frac{1}{8} \\ \frac{16}{16} & \frac{16}{16} & \frac{16}{16} & \frac{16}{16} & \frac{8}{8} & \frac{8}{8} & \frac{8}{8} & -\frac{8}{8} & \frac{8}{8} & -\frac{8}{8} \\ -\frac{1}{16} & -\frac{1}{16} & -\frac{1}{16} & \frac{7}{16} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & \frac{3}{8} & \frac{3}{8} & \frac{3}{8} \\ -\frac{16}{16} & \frac{16}{16} & -\frac{16}{16} & \frac{16}{16} & -\frac{8}{8} & -\frac{8}{8} & -\frac{8}{8} & \frac{8}{8} & \frac{8}{8} & \frac{8}{8} \end{bmatrix}$$

$$Q_{orth}^1 = \begin{bmatrix} \frac{3}{8} & \frac{3}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{3}{8} \\ \frac{3}{8} & \frac{1}{8} & \frac{3}{8} & \frac{3}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{3}{8} & \frac{3}{8} & \frac{1}{8} & \frac{3}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{3}{8} & \frac{3}{8} & \frac{3}{8} \\ \frac{5}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{5}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{5}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{5}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{5}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{5}{8} \end{bmatrix}$$

Deux points importants apparaissent :

- Les nouveaux filtres ainsi construits sont alors représentés par des matrices pleines. Ceci pose un problème d'implémentation, car la dimension de ces matrices augmente avec le nombre de sommets du maillage. Par exemple, la matrice  $A_{orth}^j$  construite pour approcher un maillage possédant 20 000 sommets par un maillage possédant 5000 sommets, possède  $20\ 000 \times 5000 = 100\ 000\ 000$  coefficients réels.
- Pour expliciter le banc de filtres entre les niveaux de résolution  $j$  et  $j-1$ , le calcul de  $\alpha^j$  nécessite l'inversion de la matrice  $I^{j-1}$ , qui pour un maillage possédant 5000 sommets est une matrice  $(5000 \times 5000)$

Par conséquent, et comme dans les domaines plus classiques des ondelettes 1D et 2D, le support de chaque ondelette est réduit à un certain voisinage autour du sommet sur lequel elle est centrée. Ceci se traduit mathématiquement par le calcul d'une matrice  $\alpha^j$  dont une partie seulement des coefficients est non nulle. Une implémentation par gestion des matrices creuses [VALE98] permet de garantir que la complexité calculatoire de l'algorithme varie linéairement avec le nombre de sommets des maillages utilisés. Durant nos travaux, nous avons restreint le support de chaque ondelette à un disque d'ordre 2 (voir figure 2.1.13) autour du sommet considéré.

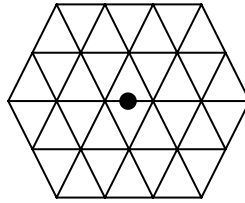


Figure 2.1.13 : disque d'ordre 2 autour d'un sommet crée lors le la subdivision

La méthode de décomposition en ondelettes proposée par Lounsbery est séduisante, mais présente un défaut majeur : elle ne s'applique qu'aux maillages ayant été subdivisés itérativement par la méthode 1:4, qui permet d'établir une hiérarchie entre les triangles à différents niveaux de résolution, comme le montre la figure 2.1.14

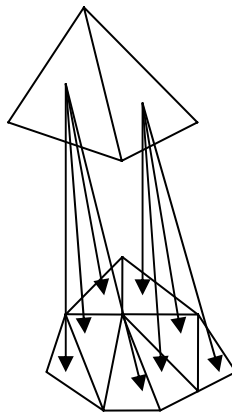


Figure 2.1.14 : relations hiérarchiques entre deux niveaux de résolution

Cette contrainte de connectivité n'est jamais respectée pour un maillage généré par une autre méthode que la subdivision répétée d'un maillage basse résolution, ce qui limite le champ d'application de cet algorithme. Bonnaud [BONN98] propose une extension de la décomposition en ondelettes pour les maillages irréguliers, mais l'algorithme proposé n'est applicable qu'aux maillages planaires ou sphériques.

Pour pallier ce problème, des algorithmes de remaillage ont été mis en œuvre, comme décrit dans la section suivante.

### 2.1.3 METHODES DE REMAILLAGE

Pour surmonter la contrainte de connectivité de subdivision, plusieurs méthodes de remaillage (« remeshing » en anglais) existent : ce sont des algorithmes de pré-traitement appliqués sur des maillages quelconques, qui changent leur connectivité tout en conservant leur géométrie. Le maillage fourni en sortie de l’algorithme de remaillage possède alors une connectivité de subdivision, le rendant propice à une analyse multirésolution par décomposition en ondelettes, comme le montre la figure 2.1.15

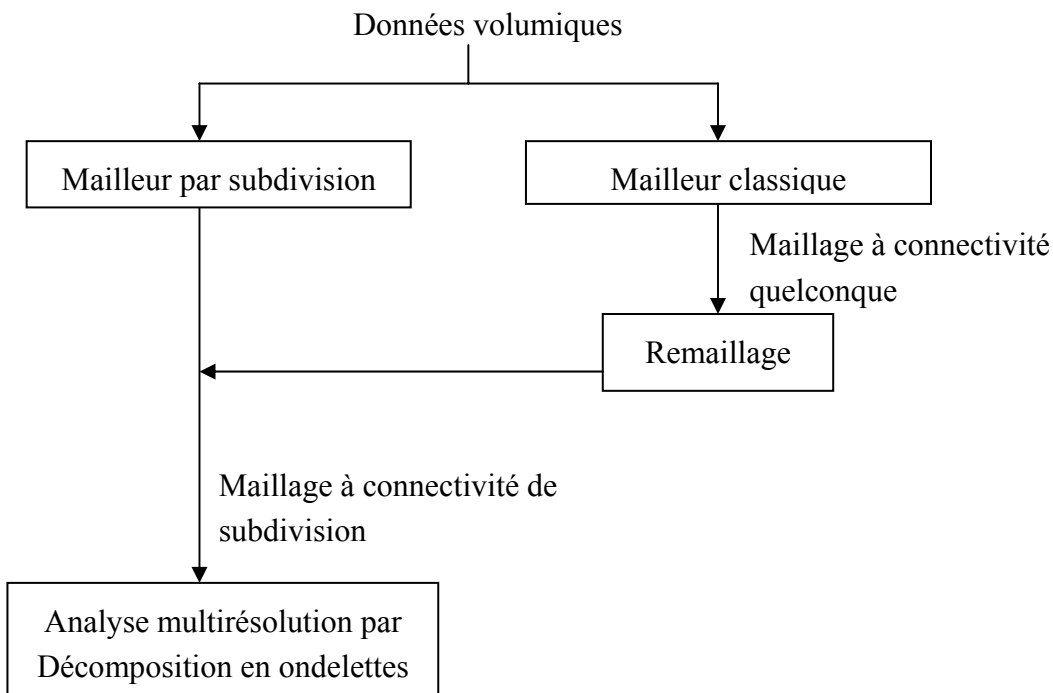


Figure 2.1.15 : enchaînement des opérations nécessaires pour la décomposition en ondelettes des maillages 3D

L’approche proposée dans [ECK95] consiste à approcher un maillage quelconque  $M$  (de connectivité irrégulière) par un maillage subdivisé  $M^j$  (de connectivité régulière) avec une erreur inférieure à une tolérance prescrite  $\varepsilon$ . La première étape consiste à créer un maillage basse résolution homéomorphe au maillage d’origine. L’approximation est réalisée à l’aide de la transformée harmonique. Le maillage basse résolution est subdivisé itérativement jusqu’à atteindre la qualité d’approximation requise. La figure 2.1.16 montre l’effet d’un tel remaillage sur le maillage du ventricule gauche d’un cœur humain, remaillé sur la base d’un octaèdre.

Un algorithme rapide de paramétrisation des surfaces à topologie arbitraire est proposé dans [LEE98], conduisant ainsi à un algorithme de remaillage hiérarchique.

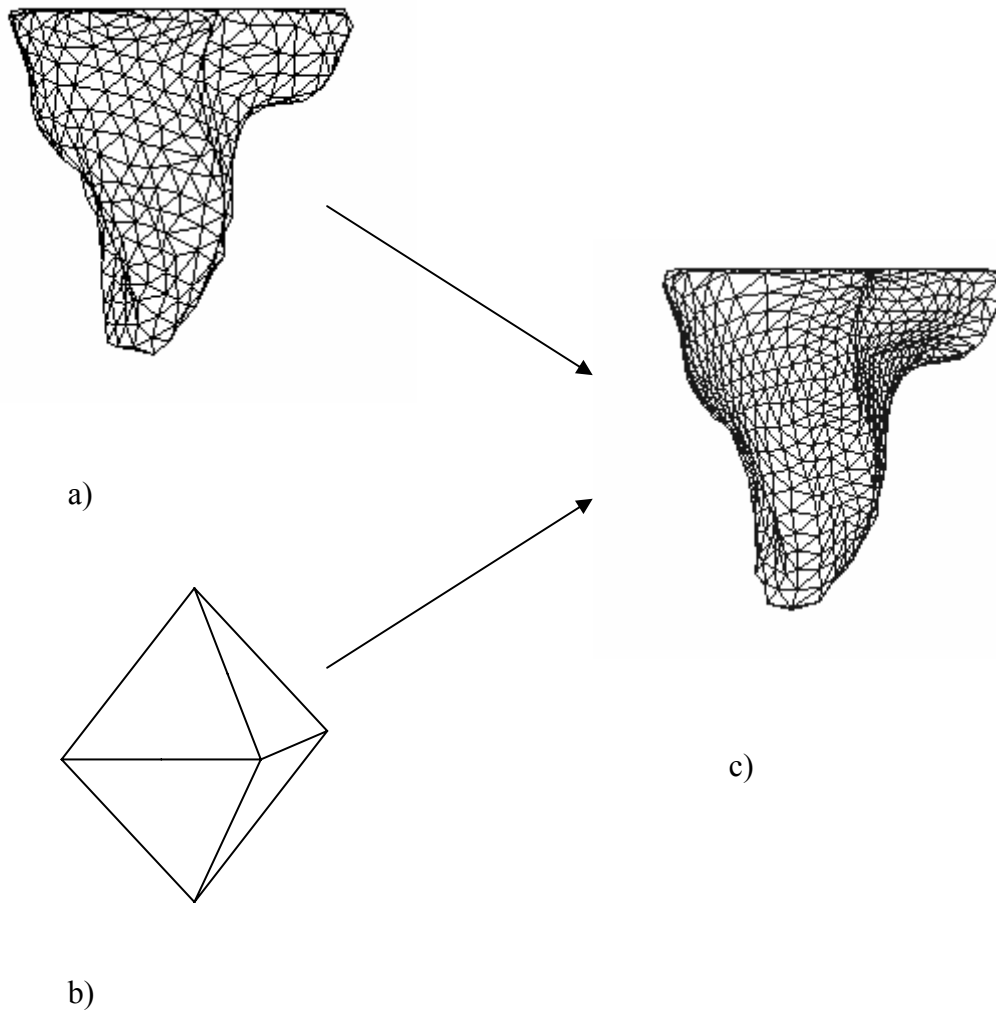


Figure 2.1.16 : remaillage pour décomposition en ondelettes

- a) Maillage à connectivité quelconque
- b) Maillage basse résolution (octaèdre)
- c) Maillage à connectivité de subdivision

Cette technique de remaillage a cependant quelques inconvénients : les calculs de transformation harmoniques sont complexes, et la subdivision inconditionnelle de toutes les faces du maillage impose un nombre de faces constant, autant dans les zones à faible qu'à forte courbure.

Une méthode de remaillage adapté à la courbure locale des maillages est proposée par [KIM01a]. Pour cela, une proposition de décomposition en ondelettes des maillages

triangulaires par subdivision locale est proposée, par une formulation de matrices de produits scalaires locaux [KIM99]. Un algorithme utilisant la transformation conforme et la transformation harmonique permet ensuite d'approcher le maillage d'entrée par un maillage dont la subdivision est adaptée à la courbure locale de la surface (voir figure 2.1.17).

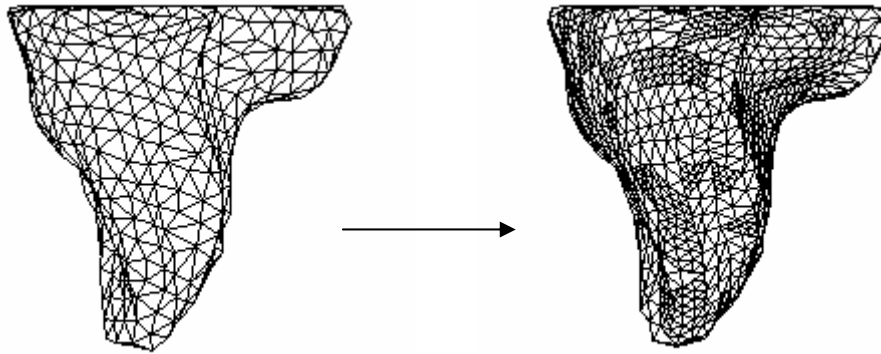


Figure 2.1.17 : remaillage adapté à la courbure locale :

- a) maillage original irrégulier
- b) maillage dont la subdivision est adaptée à la courbure locale

## 2.2 Compression de maillages

La compression de données concernant les maillages triangulaires est apparue relativement récemment. Une revue des algorithmes existants est proposée dans [ROSS99]. Les différentes approches ont toutes un point commun pour rendre le codage efficace : elles couplent la description de la connectivité du maillage avec la description de la géométrie. En effet, ces deux informations sont fortement corrélées : un sommet  $v_1$  ayant pour voisins un ensemble de sommets  $\{v_2, v_3, v_4, v_5, \dots\}$ , verra sa localisation dans l'espace 3D proche de ces mêmes voisins. La description géométrique du maillage (les coordonnées des sommets) doit donc être transmise en fonction de la description topologique (la connectivité du maillage). On peut classer les algorithmes de compression de maillages 3D en deux grandes catégories :

- Les approches à résolution unique : la connectivité du maillage est codée en une fois, par énumération de la connectivité du maillage selon un code compact.
- Les approches multirésolution : le maillage original est simplifié itérativement jusqu'à un maillage de résolution minimale. Un code compact permet la reconstruction du maillage original à partir du maillage de basse résolution

Dans la majorité des cas, au niveau du décodeur, les coordonnées des nouveaux sommets dans la description sont prédites linéairement à partir :

- Des sommets voisins déjà transmis pour les approches à résolution unique
- Des sommets voisins présents au niveau de résolution inférieure par rapport à la résolution en cours pour les approches multirésolution

Il suffit ensuite de transmettre la différence entre les coordonnées réelles et les coordonnées prédites. On réalise donc un codage différentiel. Les différences peuvent être quantifiées pour avoir un meilleur rapport débit/distorsion ou transmises telles quelles, dans un soucis de codage sans perte.

Les sections suivantes décrivent les deux types d'approches utilisées pour coder la connectivité des maillages. Dans les de sections suivantes nous appellerons  $s$  le nombre de sommets et  $f$  le nombre de faces du maillage considéré, et les coûts de codage indiqués sont les coûts de codage pour la connectivité uniquement

### 2.2.1 COMPRESSION DE MAILLAGES A RESOLUTION UNIQUE

Les algorithmes de compression à résolution unique proposent une énumération de la connectivité du maillage selon un ordre déterminé, fournissant un codage compact de la connectivité. Deux approches sont possibles : une description basée sur les faces du maillage et une description basée sur les sommets.

#### a) Approches basées faces

Les approches basées sur la description des faces proposent un découpage du maillage en bandes de triangles (« triangles strips ») dont la description est itérative [DEER95]: chaque nouveau triangle est créé à l'aide de deux des trois sommets formant le triangle décrit à l'étape précédente, un bit est transmis pour indiquer la direction dans laquelle le triangle doit être construit (droite ou gauche, voir figure 2.2.1). Cette technique nécessite un découpage adéquat du maillage afin de réaliser une compression efficace

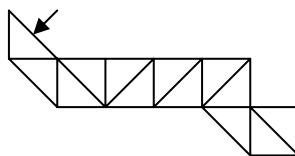


Figure 2.2.1 : codage par bandes de triangles : le premier triangle codé est pointé par la flèche. Les triangles suivants sont codés par un flux binaire indiquant la direction de propagation de la bande (0 pour gauche, 1 pour droite). Pour cet exemple, le flux binaire est :

0001101011000.

Mais ce flux binaire n'est pas suffisant pour transmettre correctement la connectivité du maillage. En effet, chaque sommet est rencontré deux fois dans la description des bandes. On doit donc, pour chaque nouveau triangle, savoir si le nouveau sommet a déjà été rencontré (1 bit par triangle) et le cas échéant, donner son indice dans la table des triangles (1 fois sur deux, et  $\log_2(v)$  bits par sommet, soit  $0.5\log_2(v)$  bits). Le coût de codage par face s'élève donc à  $2 + 0.5\log_2(v)$  bits.

De nombreux travaux dérivent de cette approche [TAUB98a], essayant de garantir un coût de codage par face constant avec le nombre de faces. Finalement, Rossignac [ROSS98] a proposé un algorithme, « edgebreaker » garantissant un coût de codage maximal de 2 bits par face, toujours basé sur le découpage du maillage en bandes de triangles, mais utilisant cette fois un alphabet à 5 codes au lieu d'un train binaire pour décrire les bandes de triangles.

## b) Approches basées sommets

Les approches basées sur les sommets énumèrent itérativement les sommets en indiquant leur valence (c'est-à-dire leur nombre de sommets voisins) [TOUM98][ISEN99][ALLI01a], à partir d'un algorithme de conquête. Ici encore, des informations supplémentaires doivent être transmises pour gérer les maillages non réguliers et les bords. Associée à un codage entropique (le sommets de valence égale à 6 sont les plus fréquemment rencontrés), Cette approche propose un coût de codage de la connectivité des maillages compris entre 1 et 2 bits par face.

### **2.2.2 COMPRESSION MULTIREOLUTION DES MAILLAGES**

La compression multirésolution des maillages 3D permet une transmission progressive, par raffinements successifs du maillage. Le choix judicieux des méthodes de raffinement permet une description compacte des relations entre les niveaux de résolutions obtenus. Plusieurs schémas de simplification-raffinement ont été utilisés :

- Une implémentation de l'algorithme « progressive meshes » est proposée dans [PARA00]. Le coût de codage obtenu est d'environ 3.6 bits par face,
- Le schéma de raffinement « Progressive Forest Split » décrit dans [TAUB98b] fournit un coût de codage de 3.5 bits par face, en moyenne.
- Dans [COHE99], des sommets sont supprimés, et leur voisinage est remaillé. Une technique de remaillage déterministe permet une reconstruction du maillage avec un faible coût de codage, qui se situe aux alentours de 3 bits par face.
- L'algorithme précédent a récemment été amélioré dans [ALLI01b], où l'auteur utilise deux phases distinctes de suppression de sommets. La première phase supprime des sommets dont la valence est comprise entre 3 et 6, et la deuxième supprime uniquement les sommets de valence égale à 3. La combinaison de ces deux étapes permet de garder un maillage régulier au cours des simplifications du maillage, et présente un coût de codage moyen de 1.85 bits/face.

### 3 NOTRE PROPOSITION

#### 3.1 Objectif

Dans cette section , nous proposons un nouvel algorithme de décomposition en ondelettes applicable aux maillages 3D triangulaires quelconques [VALE99a]. Cette approche présente deux avantages :

- L'étape de remaillage est ici évitée, car la décomposition en ondelettes s'applique directement sur le maillage d'origine haute résolution (figure 3.1). Ainsi, la connectivité du maillage original est conservée, ce qui est un avantage majeur pour des maillages sur lesquels on applique des textures lors du rendu, car on évite un ré-échantillonnage des textures. D'autre part, lorsque le changement de la connectivité du maillage est à proscrire, comme par exemple en modélisation par éléments finis, notre proposition présente l'avantage indispensable de ne pas influencer les résultats des simulations.
- Le schéma multirésolution proposé permet une représentation machine efficace de la connectivité du maillage, et la décomposition en ondelettes appliquée à la géométrie transforme les coordonnées des sommets du maillage en coefficients d'ondelettes, propices à la compression. L'algorithme proposé permet donc une compression avec ou sans pertes des maillages triangulaires [VALE99b].

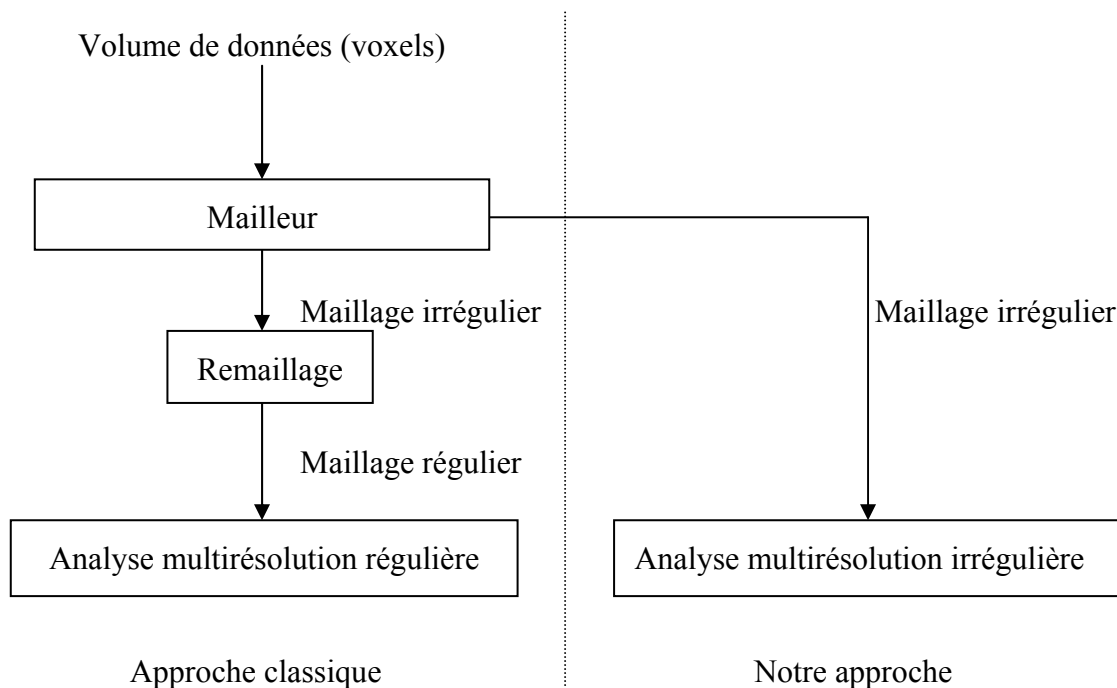


Figure 3.1 : Proposition d'une analyse multirésolution irrégulière

Une telle approche nécessite deux modifications dans l’algorithme classique de décomposition en ondelettes des maillages triangulaires [LOUN94] :

- L’étape de subdivision, qui conduit à des relations hiérarchiques entre les faces et les nœuds des maillages à différents niveaux de résolution
- La construction des bancs de filtres, qui doit prendre en compte le nouveau type de subdivision.

L’utilisation d’une arithmétique entière permet de garantir un codage sans pertes des coordonnées des sommets du maillage.

### 3.2 Proposition d’une subdivision irrégulière

Dans le schéma de décomposition en ondelettes, la connectivité de tous les maillages dépend de la connectivité du maillage à la plus basse résolution.

Ainsi, seuls les maillages possédant une connectivité de subdivision peuvent être traités. Pour résoudre ce problème, nous proposons un nouveau type de subdivision.

Dans notre approche, une face peut être subdivisée en 4 faces (le cas 1:4), en 3 faces, en 2 faces, ou rester inchangée. La figure 3.2.1 montre les 11 cas de subdivisions envisagés pour une face donnée.













Face originale	Face subdivisée (inchangée)	Face subdivisée (1:2)	Face subdivisée (1:3)	
	 1	 2	 5	 6
	Face subdivisée (1:4)	 3	 7	 8
	 11 (approche Lounsbery)	 4	 9	 10

Figure 3.2.1 : Liste des cas de subdivision possibles

### 3.3 Le problème direct (subdivision)

Soit un maillage triangulaire  $M^j$  ayant  $n^j$  faces et  $v^j$  sommets,  $S^j$  est un schéma de subdivision irrégulière appliqué à  $M^j$ .

$S^j$  est représenté par un vecteur  $s^j$  contenant  $n^j$  éléments (des entiers entre 1 et 11, voir figure 3.2.1). Chacun des éléments de  $s^j$  décrit la subdivision employée pour une face du maillage :

$$s^j = [q_1^j \ q_2^j \ \dots \ q_n^j] \quad \text{II-33}$$

$M^{j+1}$  est le résultat de la subdivision de  $M^j$  par  $S^j$  :

$$S^j(M^j) = M^{j+1} \quad \text{II-34}$$

La figure 3.3.1 illustre 4 cas de subdivision possibles pour un maillage constitué de 6 faces.

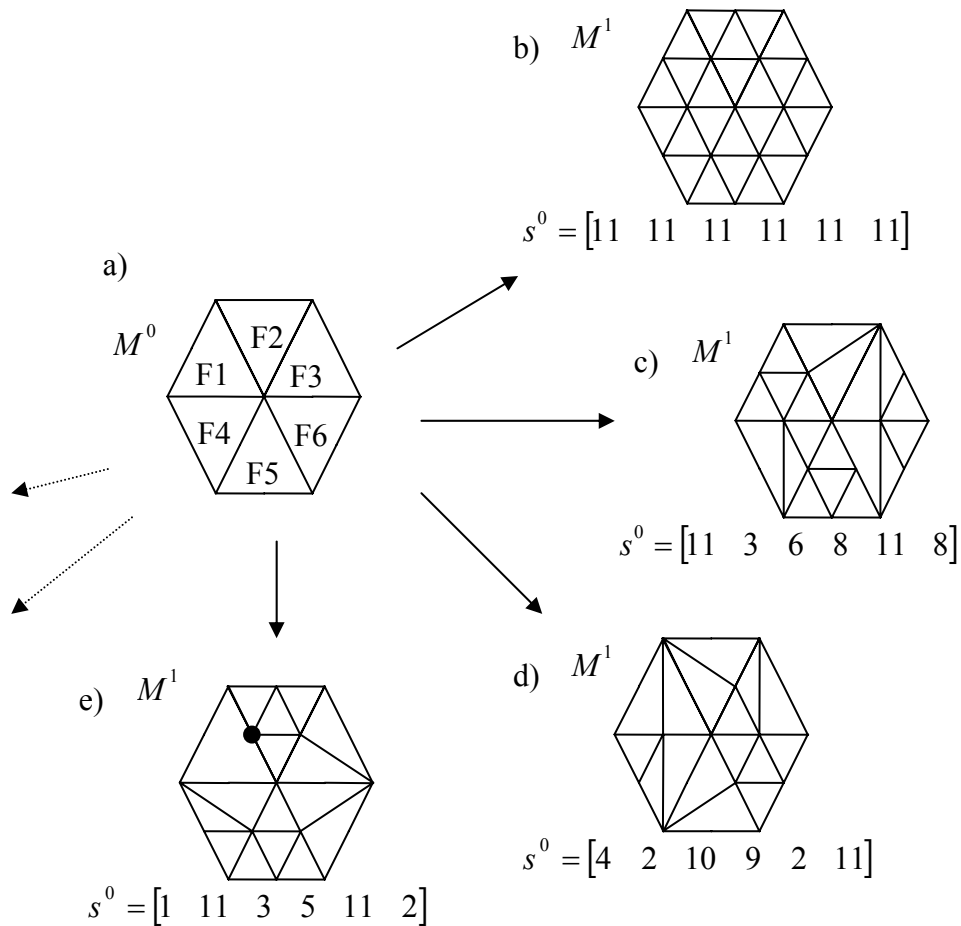


Figure 3.3.1 : Un maillage peut subir plusieurs types de subdivisions :

- a) Maillage original (4 faces)
- b) Maillage subdivisé (approche de Lounsbery)
- c) Maillage subdivisé (approche irrégulière)
- d) Maillage subdivisé (approche irrégulière)
- e) Maillage subdivisé (approche irrégulière) non conforme

Nous définissons ensuite le taux de fusion pour une approche fin à grossier  $r^j$  :

$$r^j = \frac{n^{j+1}}{n^j} \quad \text{II-35}$$

Il vient alors :

$$1 \leq r^j \leq 4 \quad \text{II-36}$$

Il existe  $11^i$  cas de subdivision pour un maillage donné, mais tous ne mènent pas à un maillage conforme, comme le montre la figure 3.3.1 : La subdivision des six faces du maillage en 3.3.2.a par un certain schéma de subdivision donne un maillage non conforme (figure 3.2.2.c, sommet marqué).

### 3.4 Le problème inverse (fusion)

Pour appliquer l'algorithme d'analyse multirésolution sur un maillage haute résolution  $M^j$ , il faut construire un maillage  $M^{j-1}$  et un schéma de subdivision  $S^{j-1}$  satisfaisant :

$$S^{j-1}(M^{j-1}) = M^j \quad \text{II-37}$$

Il s'agit donc de résoudre un problème inverse. Dans une optique de compression efficace, il faut trouver le schéma de subdivision tel que  $r^{j-1}$  soit le plus grand possible ( $r^{j-1}=4$  au maximum). Cela consiste à fusionner les faces du maillage  $M^j$ , et construire un maillage  $M^{j-1}$  aillant le nombre de faces le plus faible possible.

La figure 3.4.1 montre un exemple de simplification, où un ensemble de 15 faces est réduit à un ensemble de 6 faces, résultant de la fusion simultanée de 4 faces pour G1, 3 faces pour G3, 2 faces pour G1 et G4 (G5 est une face inchangée). Pour ce schéma de fusion - subdivision,

$$r^{j-1} = \frac{15}{6} = 2.5.$$

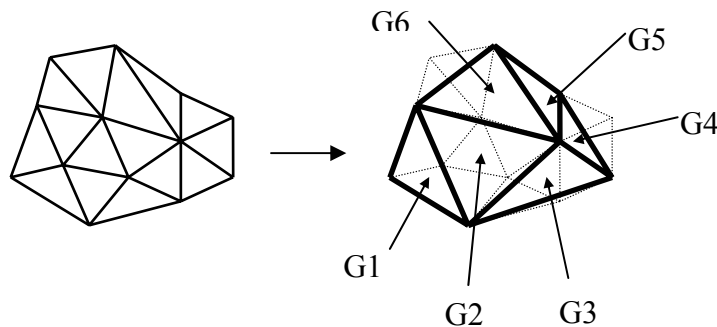


Figure 3.4.1 : un exemple de simplification de maillage

Il a donc fallu réaliser un algorithme de fusion de triangles, selon les cas de subdivision envisagés. Notre choix s'est porté sur un algorithme de *conquête*. Tout d'abord, l'algorithme commence par la sélection d'une seule face, constituant ainsi un domaine « fusionné », dont la taille augmente avec le nombre de triangles fusionnés. La figure 3.4.2. montre le début de l'expansion du domaine fusionné (figure 3.4.2.a en gris), par la fusion d'un groupe de 2 faces (figure 3.4.2.b), 4 faces (figure 3.4.2.c) et 3 faces (figure (3.4.2.d).

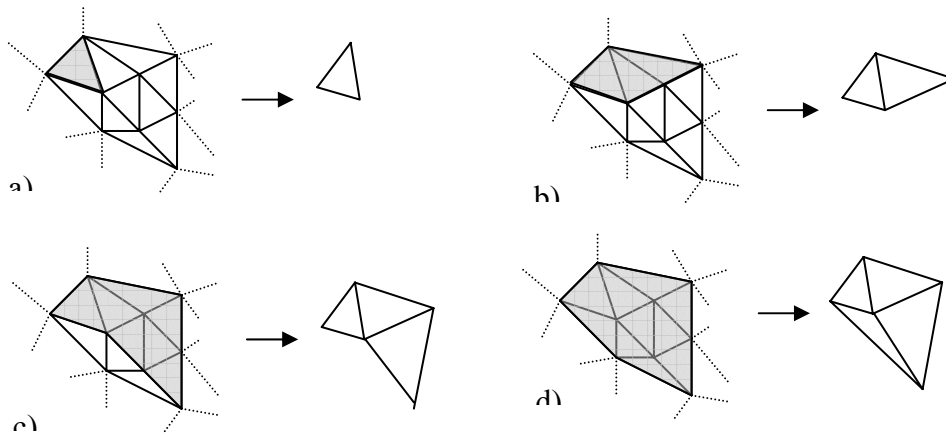


Figure 3.4.2 : expansion du domaine fusionné sur le maillage

La conception d'un tel algorithme est fastidieuse, et à l'heure actuelle, aucune justification théorique n'a pu être avancée quant à la découverte d'un schéma de fusion - subdivision optimal.

Pour éviter un blocage de l'algorithme dans certaines zones du maillage fortement irrégulières, une modification du maillage d'entrée est autorisée. Si une configuration bloquante est rencontrée, une permutation d'arête entre 2 faces voisines est réalisée, comme le montre la figure 3.4.3. Cette modification devra être mémorisée pour assurer une reconstruction parfaite lors du processus de synthèse.

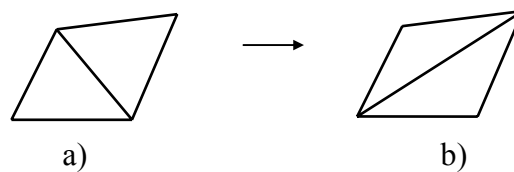


Figure 3.4.3 : permutation d'arête :

a) 2 faces voisines

b) permutation entre les deux faces

Un exemple de configuration bloquante est donné par la figure 3.4.5. Il faut remarquer que ces modifications vont introduire une perte de qualité dans les approximations successives. Les permutations d'arête devront donc être employées le moins souvent possible.

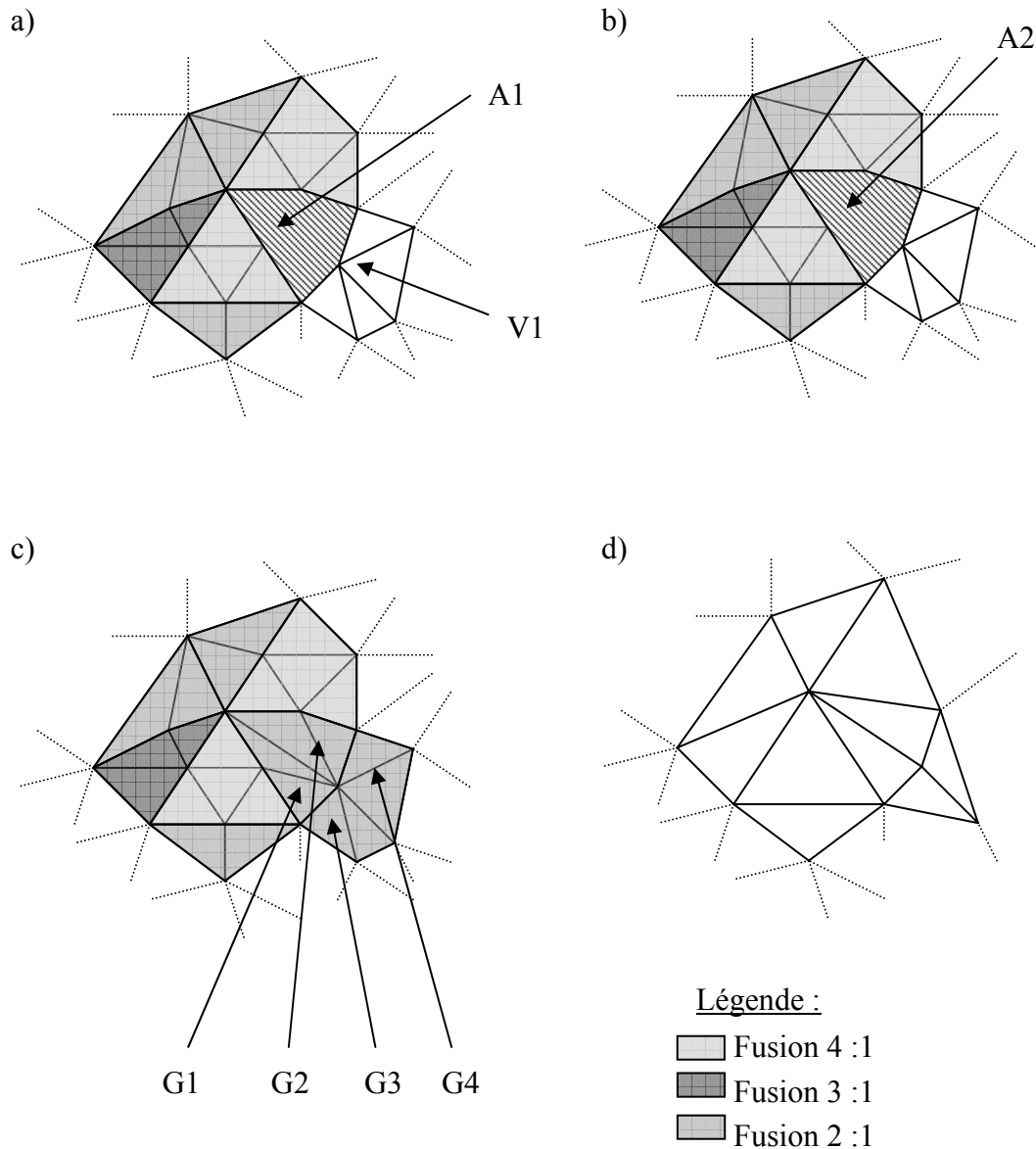


Figure 3.4.5 : exemple de configuration bloquante résolue par permutation d'arête :

- a) L'algorithme a fusionné toute une zone du maillage (faces colorées en gris) mais un groupe de 4 faces (zone hachurée) ne peut être fusionné car le sommet V1 possède 7 sommets voisins et ne peut donc pas être supprimé lors de la simplification (s'il avait eu 6 voisins, les quatre faces hachurées auraient été fusionnées ensemble).
- b) Pour résoudre ce problème, on modifie le maillage en supprimant l'arête A1 et en créant l'arête A2.
- c) Après la permutation d'arête, l'algorithme peut continuer sa progression dans la simplification du maillage, en fusionnant quatre groupes de deux faces (G1, G2, G3 et G4) autour du sommet V1
- d) Maillage après simplification.

### 3.5 Algorithme de fusion de triangles

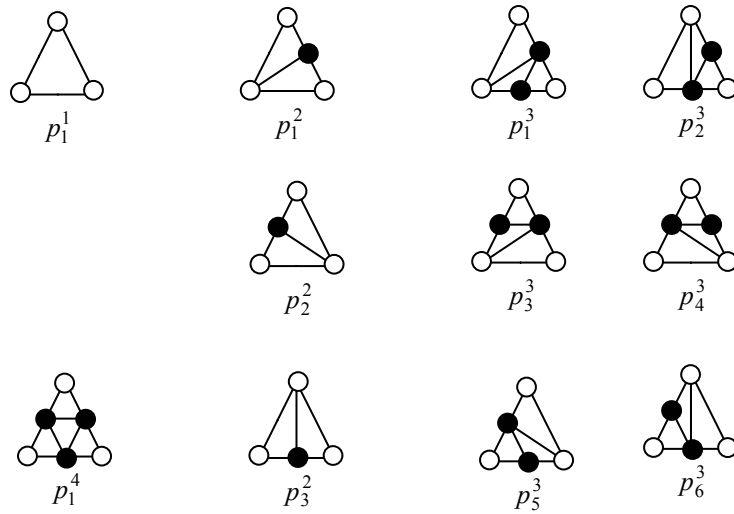
Nous allons ici décrire l'algorithme de fusion des triangles d'un maillage. Il faut tout d'abord définir 3 dictionnaires, décrivant les cas de fusion-subdivision possibles :

- Le dictionnaire des ondelettes  $W = \{p_i^u\}$ . La figure 3.5.1.a décrit les cas de subdivision irrégulière (1:1, 1:2, 1:3, 1:4). L'exposant  $u$  indique l'ordre de la subdivision du motif décrit (1: $u$ ).  $u$  est aussi le nombre de faces contenues dans le motif considéré.
- Le dictionnaire des incidents  $I = \{f_e^v\}$ . La figure 3.5.1.b décrit les cas possibles de subdivisions combinées avec une permutation d'arête. L'exposant  $v$  indique le nombre de faces dans chaque motif.
- Le dictionnaire des fusions  $M = \{g_i^w\}$ . La figure 3.5.1.c indique comment les triangles sont fusionnés : chaque motif du dictionnaire des ondelettes et du dictionnaire des incidents correspond à un motif de fusion dans le dictionnaire des fusions (tous les motifs du dictionnaire des ondelettes correspondent au motif  $g_1^1$  du dictionnaire des fusions).

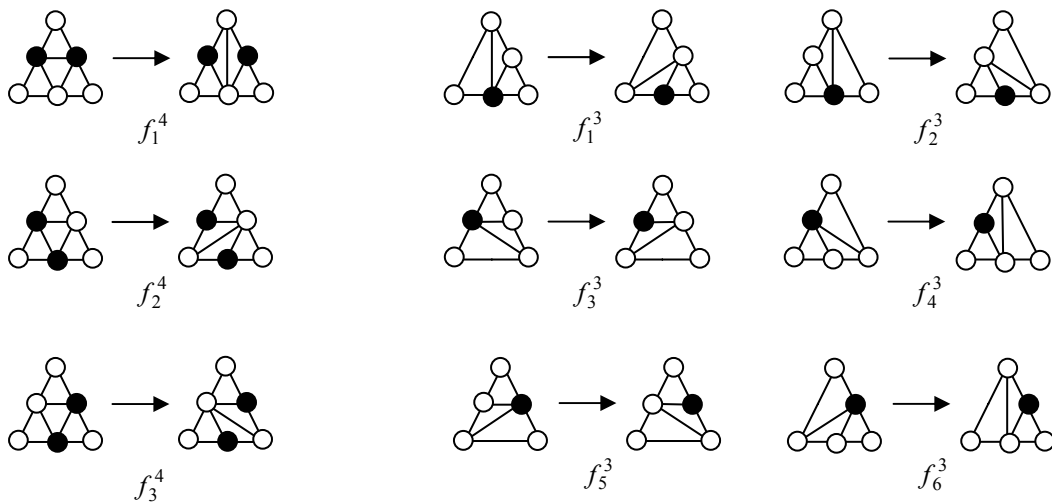
Dans ces trois dictionnaires les sommets marqués en blanc sont appelés sommets *parents*. Les sommets marqués en noir sont appelés sommets  *fils* . Les sommets fils disparaissent lors de la fusion des faces, les sommets pères sont conservés. Un sommet fils possède toujours deux sommets parents. Certaines règles de fusion doivent être respectées, pour que le processus de fusion soit réversible :

- Un sommet peut être marqué comme sommet fils uniquement si sa valence (le nombre de sommets directement voisins de celui-ci) est égale à 4, 5 ou 6.
- Plusieurs sommets fils peuvent avoir un sommet père en commun, mais deux sommets fils ne peuvent pas avoir les deux mêmes sommets parents.

a) Dictionnaire des ondelettes  $W = \{p_i^u\}$  :



b) Dictionnaire des incidents  $I = \{f_e^v\}$  :



c) Dictionnaire des fusions  $M = \{g_i^w\}$  :

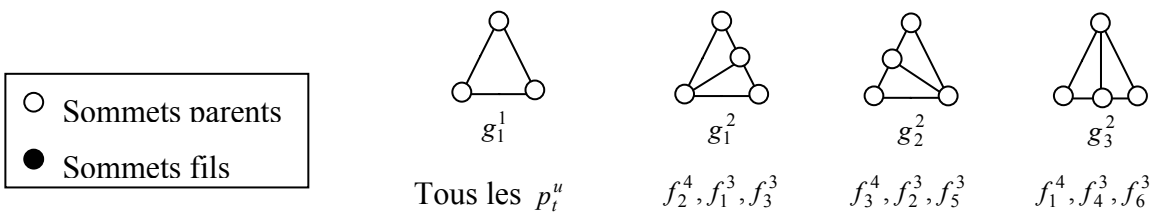


Figure 3.5.1 : Dictionnaires de subdivision - fusion

Cet algorithme permet donc de fusionner de manière réversible un maillage donné  $M^j$  en un maillage de résolution inférieure  $M^{j-1}$ .

Nous dirons qu'un motif de fusion « correspond » lorsqu'il est le motif le plus approprié à la configuration locale du maillage, et lorsqu'un test sur le nombre d'arêtes à l'extérieur du domaine fusionné est positif pour tout sommet devant être marqué en tant que sommet fils. Ce nombre d'arêtes doit être égal à 1 ou 2 pour éviter que l'algorithme de fusion soit bloqué dans une configuration bloquante (voir exemple sur la figure 3.5.2.c). L'algorithme consiste en l'enchaînement des opérations suivantes :

Choisir un triangle "germe"  $s$ . Le triangle germe définit le domaine simplifié  $F_0 = \{s\}$  et le domaine complémentaire à simplifier  $\overline{F_0}$  avec  $F_0 \cup \overline{F_0} = M^j$ . Elle définit aussi l'ensemble d'arêtes  $E_0 = \{e_{0,k}\}_{k=0\dots N_0}$  tel que  $E_0 = F_0 \cap \overline{F_0}$ .

$i = 0$

$N_i$  est défini par le nombre d'arêtes au bord du domaine  $F_i$  moins 1.

Tant que  $N_i \geq 0$  faire

{

Pour  $k=0$  à  $N_i$  :

{

Sélectionner le motif dans le dictionnaire des ondelettes  $p_m^u = (p_n^v \text{ correspondant à } \overline{F_i} \text{ en } e_{i,k}) / u = \max(v)$

Si aucun motif ne correspond : sélectionner dans le dictionnaire des incidents  $f_l^w = (f_n^v \text{ correspondant à } \overline{F_i} \text{ en } e_{i,k}) / w = \max(v)$

}

Créer le nouveau domaine simplifié étendu  $F_{i+1} = \bigcup_m p_m^u \text{ sélectionnés} + \bigcup_l f_l^w \text{ sélectionnés} \cup F_i$

Créer le maillage  $M_{i+1}^{j-1}$  en fusionnant les motifs sélectionnés  $p_m^u$  et  $f_l^w$  selon leur motif de fusion du dictionnaire des fusions  $M = \{g_l^w\}$

$i = i + 1$

Définir le domaine à simplifier  $\overline{F_i}$  et l'ensemble des arêtes frontières  $E_i = \{e_{i,k}\}_{k=0\dots N_i}$  tel que  $F_i \cup \overline{F_i} = M^j$  et  $E_i = F_i \cap \overline{F_i}$

}

$M^{j+1} = M_i^{j-1}$  est le nouveau maillage simplifié à la résolution  $j-1$

La figure 3.5.2 montre un exemple détaillé de la progression de l'algorithme proposé pour un maillage comportant peu de faces :

- Le triangle germe  $s$  choisi est donné sur la figure 3.5.2.a. Le domaine simplifié  $F_0$  est alors égal au triangle germe,  $E_0 = \{e_{0,0}, e_{0,1}, e_{0,2}\}$  est l'ensemble des trois arêtes autour de  $s$ .
- Les figures 3.5.2.b, 3.5.2.c, 3.5.2.d montrent les motifs correspondants à  $\overline{F_0}$  en  $e_{0,0}$ . Il faut remarquer que sur la figure 3.5.2.c, le motif  $p_3^2$  ne peut être considéré comme valide, car 3 arêtes partent du sommet marqué en noir vers le domaine non simplifié. Si ce motif était sélectionné pour la fusion, l'algorithme serait alors bloqué dans cette configuration, car aucun motif de fusion - subdivision correspondant ne pourrait être trouvé dans les dictionnaires. Finalement, le motif sélectionné pour l'arête  $e_{0,0}$  est  $p_6^3$ , pour des raisons d'optimalité (c'est le motif fusionnant le plus grand nombre de faces possible).
- La figure 3.5.2.e montre la configuration de l'algorithme après la première boucle. Trois motifs d'ondelette ont été sélectionnés :  $p_6^3$  pour  $e_{0,0}$ ,  $p_1^3$  pour  $e_{0,1}$  et  $p_2^2$  pour  $e_{0,2}$ . Le nouveau domaine simplifié  $F_1$  est le domaine grisé. Le maillage fusionné résultant  $M_1^{j-1}$  est montré sur la figure 3.5.2.
- Après la seconde boucle,  $F_2$  couvre l'ensemble des faces du maillage et le processus de fusion est terminé (figure 3.5.2.g) . Le maillage final fusionné  $M^{j-1}$  est montré dans la figure 3.5.2.h .

Dans cet exemple, 22 faces ont été fusionnées en 10 faces.

#### *Discussion sur les conditions initiales :*

Le choix du triangle germe influe directement le résultat de la simplification, car c'est lui qui définit le premier domaine simplifié. Ce choix est arbitraire (on prend toujours la première face du maillage).

Pour créer le premier domaine simplifié, l'algorithme pourrait aussi choisir n'importe quel autre motif du dictionnaire comme le montre la figure 3.5.3 . Nous avons choisi de toujours commencer par sélectionner un triangle unique, par souci de simplicité.

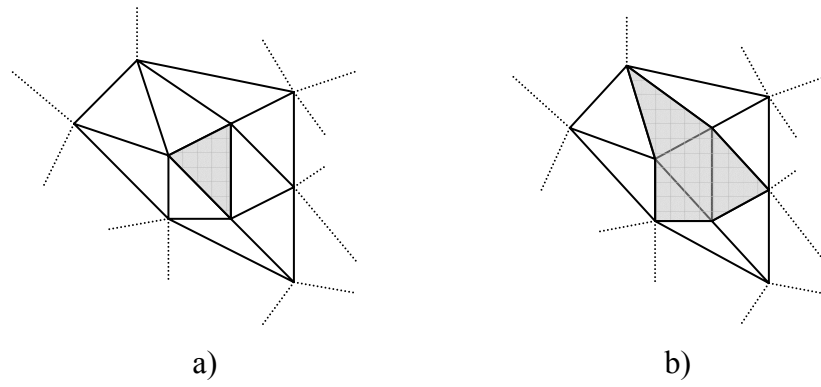


Figure 3.5.3 : choix du premier domaine simplifié :  
a) domaine constitué d'une seule face (motif  $p_1^1$ )  
b) domaine construit par fusion de 4 faces (motif  $p_1^4$ )

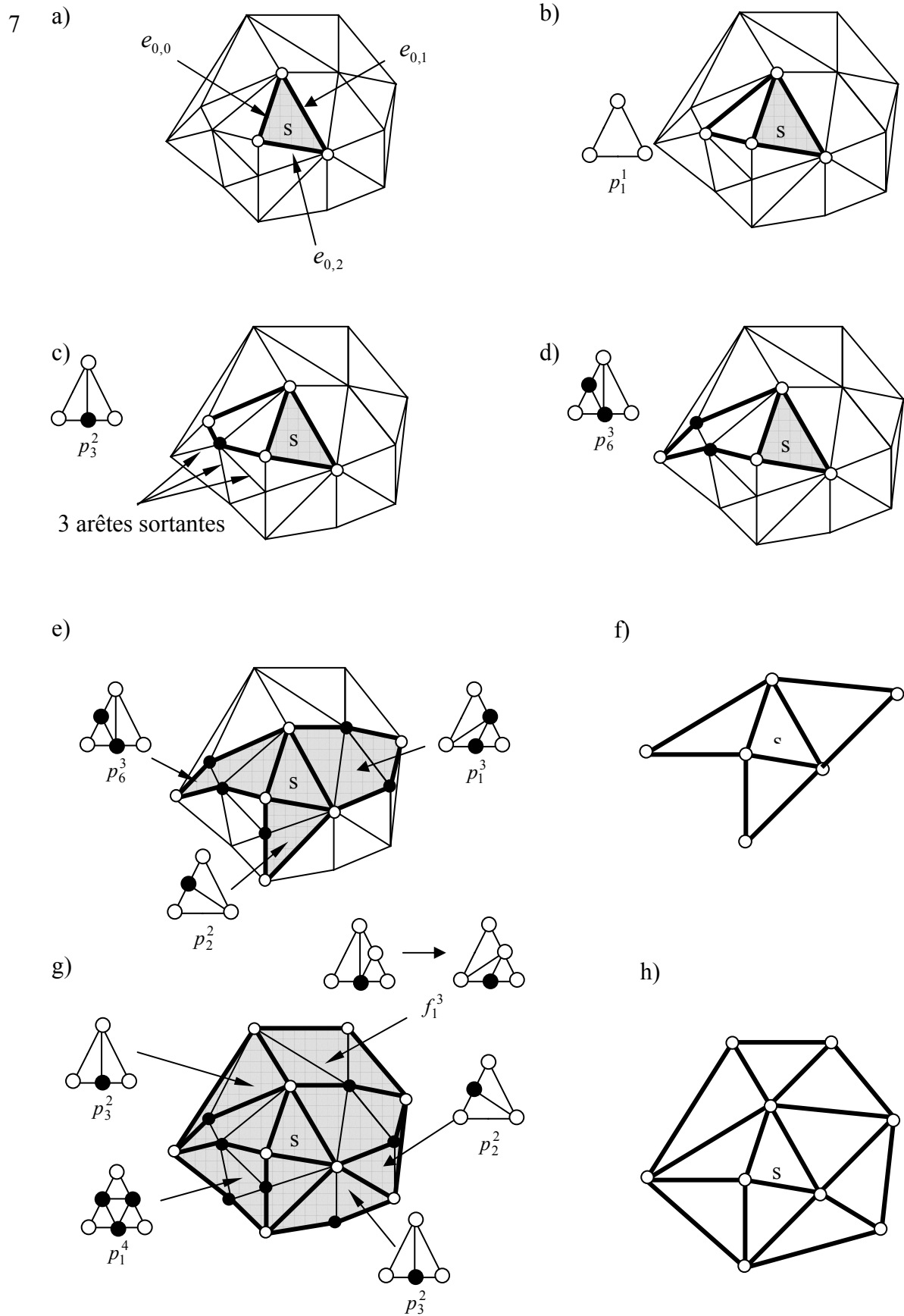


Figure 3.5.2 : déroulement de l'algorithme de fusion

3.6 Définition d'un nouveau produit scalaire pour la construction du banc de filtres

Une fois le schéma de subdivision choisi et le nouveau maillage fusionné construit, le banc de filtres d'analyses peut être construit. La différence avec la méthode classique de Lounsbery réside dans la formulation d'un nouveau produit scalaire, reflétant le nouveau mode de subdivision employée. Nous proposons le nouveau produit scalaire suivant :

$$\langle f, g \rangle = \sum_{\tau \in \Delta(M^j)} \left( \frac{K_j(\tau)}{\text{aire}(\tau)} \int_{s \in \tau} f(s).g(s).ds \right) \quad \text{II-38}$$

Où  $K_j(\tau)$  n'est plus constant avec le niveau de résolution mais change d'une face à l'autre. Par exemple, pour une face  $\tau$  du maillage basse résolution qui se subdivise en trois (figure 3.6.1.a),  $K_j(\tau) = 3$  et pour les trois faces résultant de cette subdivision (figure 3.6.1.b),  $K_{j+1}(\tau) = 1$ , ceci pour prendre en compte les différences entre les surfaces des triangles : la surface de la première face sera approximativement 3 fois plus grande que celle des trois dernières.

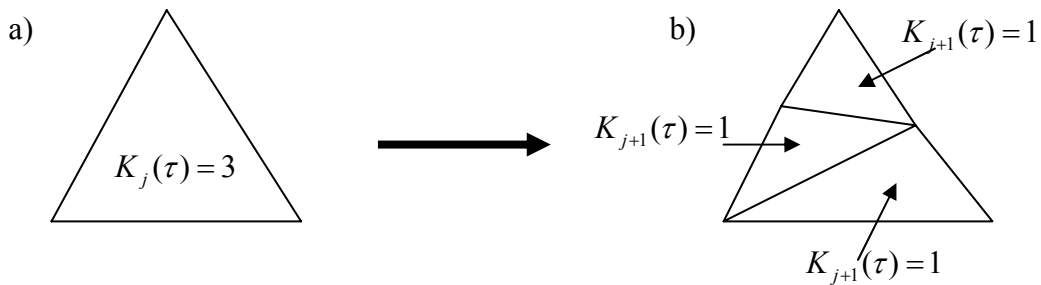


Figure 3.6.1 : coefficients de pondération  $K_j(\tau)$  pour une face se subdivisant en 3

En utilisant cette nouvelle formulation du produit scalaire, le banc de filtres d'analyse et de synthèse peut être construit, d'une manière similaire à l'approche régulière. La figure 3.6.2 donne un exemple de subdivision irrégulière d'une pyramide vue de haut: la face (V1,V2,V3) est subdivisée en 4, la face (V1,V3,V4) est subdivisée en deux et les faces (V1,V2,V4) et (V2,V3,V4) sont subdivisées en 3.

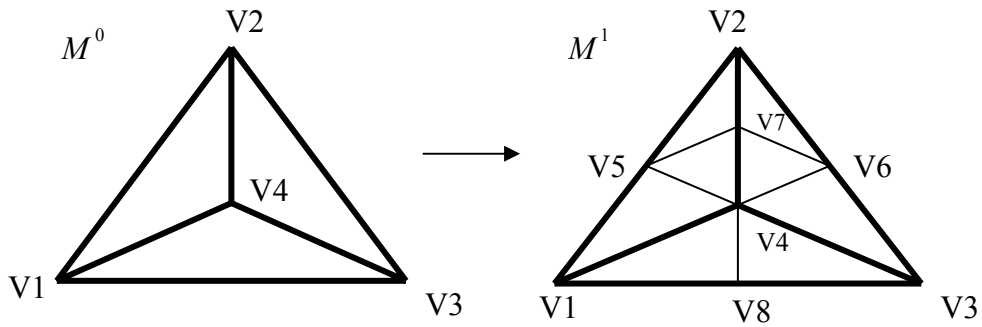


Figure 3.6.2 : exemple de subdivision irrégulière d'une pyramide

Le banc de filtres « lazy » vaut alors :

$$P_{lazy}^1 = \frac{1}{2} \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix} \quad Q_{lazy}^1 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$A_{lazy}^1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} \quad B_{lazy}^1 = \frac{1}{2} \begin{bmatrix} -1 & -1 & 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & -1 & -1 & 0 & 0 & 2 & 0 & 0 \\ 0 & -1 & 0 & -1 & 0 & 0 & 2 & 0 \\ -1 & 0 & -1 & 0 & 0 & 0 & 0 & 2 \end{bmatrix}$$

Les matrices de produit scalaires, selon notre définition valent:

$$I^0 = \frac{1}{2} \begin{bmatrix} 18 & 7 & 6 & 5 \\ 7 & 20 & 7 & 6 \\ 6 & 7 & 18 & 5 \\ 5 & 6 & 5 & 16 \end{bmatrix} \quad I^1 = \begin{bmatrix} 3 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 3 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 3 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 6 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 6 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 6 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 & 4 & 0 \\ 1 & 0 & 1 & 1 & 1 & 1 & 0 & 5 \end{bmatrix}$$

D'après ce banc de filtre, on peut calculer la matrice de « lifting »  $\alpha^1$  :

$$\alpha^1 = (I^0)^{-1} (P_{lazy}^1)^T I^1 Q_{lazy}^1 = \begin{bmatrix} 0.4033 & -0.1800 & -0.1415 & 0.3915 \\ 0.4355 & 0.4355 & 0.4057 & -0.1557 \\ -0.1800 & 0.4033 & -0.1415 & 0.3915 \\ -0.0456 & -0.0456 & 0.3113 & -0.0613 \end{bmatrix}$$

Il faut remarquer que les matrices de produit scalaires sont définies à un facteur près (la surface des triangles, supposée égale dans le maillage haute résolution), mais ce facteur n'intervient pas dans la détermination de la matrice de « lifting ». En effet, si on multiplie les deux matrices de produits scalaires par une constante  $s$ , on peut construire deux nouvelles matrices  $\tilde{I}^0 = sI^0$  et  $\tilde{I}^1 = sI^1$ , mais le calcul de la matrice de « lifting » reste inchangé :

$$\begin{aligned} \tilde{\alpha}^1 &= (\tilde{I}^0)^{-1} (P_{lazy}^1)^T \tilde{I}^1 Q_{lazy}^1 \\ &= (sI^0)^{-1} (P_{lazy}^1)^T (sI^1) Q_{lazy}^1 \\ &= \frac{1}{s} (I^0)^{-1} (P_{lazy}^1)^T sI^1 Q_{lazy}^1 \\ &= (I^0)^{-1} (P_{lazy}^1)^T I^1 Q_{lazy}^1 \\ &= \alpha^1 \end{aligned}$$

Finalement, on obtient le nouveau banc de filtres orthogonalisés :

$$A^1 = \begin{bmatrix} 0.6026 & -0.0409 & -0.1057 & 0.0708 & 0.4033 & -0.1800 & -0.1415 & 0.3915 \\ -0.1399 & 0.3616 & -0.21399 & -0.2028 & 0.4355 & 0.4355 & 0.4057 & -0.1557 \\ -0.1057 & -0.0409 & 0.6026 & 0.0708 & -0.1800 & 0.4033 & -0.1415 & 0.3915 \\ 0.0535 & -0.1101 & 0.0535 & 0.8443 & -0.0456 & -0.456 & 0.3113 & -0.0613 \end{bmatrix}$$

$$Q^1 = \begin{bmatrix} -0.4033 & 0.1800 & 0.1415 & -0.3915 \\ -0.4355 & -0.4355 & -0.4057 & 0.1557 \\ 0.1800 & -0.4033 & 0.1415 & -0.3915 \\ 0.0456 & 0.0456 & -0.3113 & 0.0613 \\ 0.5806 & -0.1278 & -0.1321 & -0.1179 \\ -0.1278 & 0.5806 & -0.1321 & -0.1179 \\ -0.1950 & -0.1950 & 0.6415 & 0.1085 \\ -0.1116 & -0.1116 & 0.1415 & 0.6085 \end{bmatrix}$$

La figure 3.6.3 illustre les différents niveaux de résolutions obtenus à partir du maillage du ventricule gauche. L'algorithme a ici été capable de simplifier le maillage original jusqu'au maillage ultime : une pyramide (4 faces et 4 sommets).





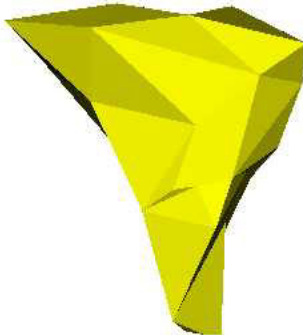

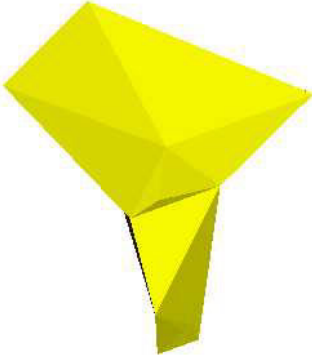


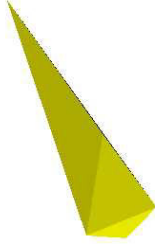
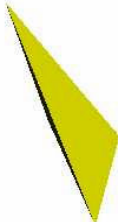
<p>Maillage original <math>M^{10}</math></p>  <p>1008 faces et 506 sommets</p>	<p>Approximation 1 <math>M^9</math></p>  <p>482 faces et 243 sommets</p>	<p>Approximation 2 <math>M^8</math></p>  <p>256 faces et 130 sommets</p>
<p>Approximation 3 <math>M^7</math></p>  <p>140 faces et 72 sommets</p>	<p>Approximation 4 <math>M^6</math></p>  <p>64 faces et 34 sommets</p>	<p>Approximation 5 <math>M^5</math></p>  <p>38 faces et 21 sommets</p>
<p>Approximation 6 <math>M^4</math></p>  <p>22 faces et 13 sommets</p>	<p>Approximation 7 <math>M^3</math></p>  <p>14 faces et 9 sommets</p>	<p>Approximation 8 <math>M^2</math></p>  <p>10 faces et 7 sommets</p>
<p>Approximation 9 <math>M^1</math></p>  <p>8 faces et 6 sommets</p>	<p>Approximation 10 <math>M^0</math></p>  <p>4 faces et 4 sommets</p>	<p>Bilan : 11 niveaux de résolution</p>

Figure 3.6.3 : analyse multirésolution sur le maillage d'un ventricule gauche

### 3.7 *Schéma multirésolution en arithmétique entière*

L'élaboration d'un schéma de codage sans pertes nécessite l'emploi d'un arithmétique basée sur les entiers. Ainsi nous considérons ici que les coordonnées des sommets des maillages traités sont entières. De plus, des opérations d'arrondi doivent être introduites dans le schéma multirésolution d'analyse et de synthèse. Pour une démonstration, il faut remonter à la définition des matrices de filtrage  $A^j$ ,  $B^j$ ,  $P^j$  et  $Q^j$  employées.

#### 3.7.1 ANALYSE - SYNTHÈSE EXACTE AVEC LE BANC DE FILTRES « LAZY »

Tout d'abord, nous introduisons le banc de filtres "Lazy":  $A_{lazy}^j$ ,  $B_{lazy}^j$ ,  $P_{lazy}^j$  et  $Q_{lazy}^j$ . Ces filtres n'exécutent aucune approximation, car durant l'analyse (où des sommets sont supprimés), les coordonnées des sommets restants restent inchangées. La structure simple de ces matrices permet la construction d'un schéma d'analyse – synthèse exact en utilisant la transformée en arrondis [JUNG98a] [JUNG98b], en considérant que les éléments de la matrice de coordonnées  $C^j$  sont entiers. La reconstruction exacte est possible, car  $A_{lazy}^j$  ne contient que des 1 et des 0, et  $B_{lazy}^j$  ne contient que des 0, 1, et -0.5. Les transformées en arrondis directes et inverses sont définies ainsi:

$$\begin{bmatrix} C_{lazy}^{j-1} \\ D_{lazy}^{j-1} \end{bmatrix} = \begin{bmatrix} A_{lazy}^j \\ B_{lazy}^j \end{bmatrix} \cdot C^j \quad \text{II-39}$$

$$C^j = \begin{bmatrix} P_{lazy}^j & Q_{lazy}^j \end{bmatrix} \begin{bmatrix} C_{lazy}^{j-1} \\ D_{lazy}^{j-1} \end{bmatrix} \quad \text{II-40}$$

où  $\lfloor \cdot \rfloor$  et  $\lceil \cdot \rceil$  sont respectivement l'opérateur d'arrondi à l'entier inférieur et l'opérateur d'arrondi à l'entier supérieur. L'opérateur d'arrondi à l'entier inférieur peut être employé dans la transformation inverse si l'opérateur d'arrondi à l'entier supérieur est utilisé dans la transformation directe.

Comme les éléments de  $A_{lazy}^j$  sont tous des entiers, il en résulte:

$$C_{lazy}^{j-1} = \lfloor A_{lazy}^j \cdot C^j \rfloor = A_{lazy}^j \cdot C^j \quad \text{II-41}$$

Par contre, les éléments de  $B_{lazy}^j$  ne sont pas tous des entiers, donc:

$$D_{lazy}^{j-1} = \lfloor B_{lazy}^j \cdot C^j \rfloor \quad \text{II-42}$$

D'après la transformée en arrondi inverse, la reconstruction est effectuée ainsi:

$$C^j = \begin{bmatrix} P_{lazy}^j \cdot C_{lazy}^{j-1} + Q_{lazy}^j \cdot D_{lazy}^{j-1} \end{bmatrix} \quad \text{II-43}$$

Ce schéma bidirectionnel est illustré par la figure 3.7.

### 3.7.2 ANALYSE - SYNTHESE EXACTE AVEC LE LIFTING

Le banc de filtres réalisant une approximation optimale est directement dérivé du banc de filtres « Lazy », sur lequel on applique le processus de lifting [SWEL96] :

$$A^j = A_{lazy}^j + \alpha^j . B_{lazy}^j \quad \text{II-44}$$

$$B^j = B_{lazy}^j \quad \text{II-45}$$

$$P^j = P_{lazy}^j \quad \text{II-46}$$

$$Q^j = Q_{lazy}^j - P_{lazy}^j . \alpha^j \quad \text{II-47}$$

où  $\alpha^j$  est une matrice  $v^{j-1} \times (v^j - v^{j-1})$  choisie pour assurer l'approximation la meilleure possible, i.e. construite afin que la base d'ondelettes soit orthogonale à la base des fonction d'échelles.

En remplaçant  $A^j$  par sa définition dans le processus de « lifting », il vient:

$$C^{j-1} = \left[ A_{lazy}^j . C^j + \alpha^j . B_{lazy}^j . C^j \right] \quad \text{II-48}$$

Sachant que tout les éléments de  $A_{lazy}^j . C^j$  sont des entiers :

$$C^{j-1} = A_{lazy}^j . C^j + \left[ \alpha^j . B_{lazy}^j . C^j \right] \quad \text{II-49}$$

D'après II-39 et II-45 :

$$D^{j-1} = \left[ B_{lazy}^j . C^j \right] \quad \text{II-50}$$

En accord avec [CALD98], nous modifions le traitement parallèle défini par II-49 et II-50 en un traitement séquentiel :

$$\begin{cases} D^{j-1} = \left[ B_{lazy}^j . C^j \right] \\ C^{j-1} = A_{lazy}^j . C^j + \left[ \alpha^j . D^{j-1} \right] \end{cases} \quad \text{II-51}$$

Le traitement inverse correspondant est alors:

$$C^j = \left[ P_{lazy}^j . \left( C^{j-1} - \left[ \alpha^j . D^{j-1} \right] \right) + Q_{lazy}^j . D^{j-1} \right] \quad \text{II-52}$$

Les équations II-51 et II-52 nous donnent maintenant la version entière du schéma classique d'analyse multirésolution par ondelettes, comme le montre la figure 3.7.1.b et 3.7.1.c .

Ce schéma peut être employé dans une chaîne de compression sans pertes, comme le décrit la section suivante.

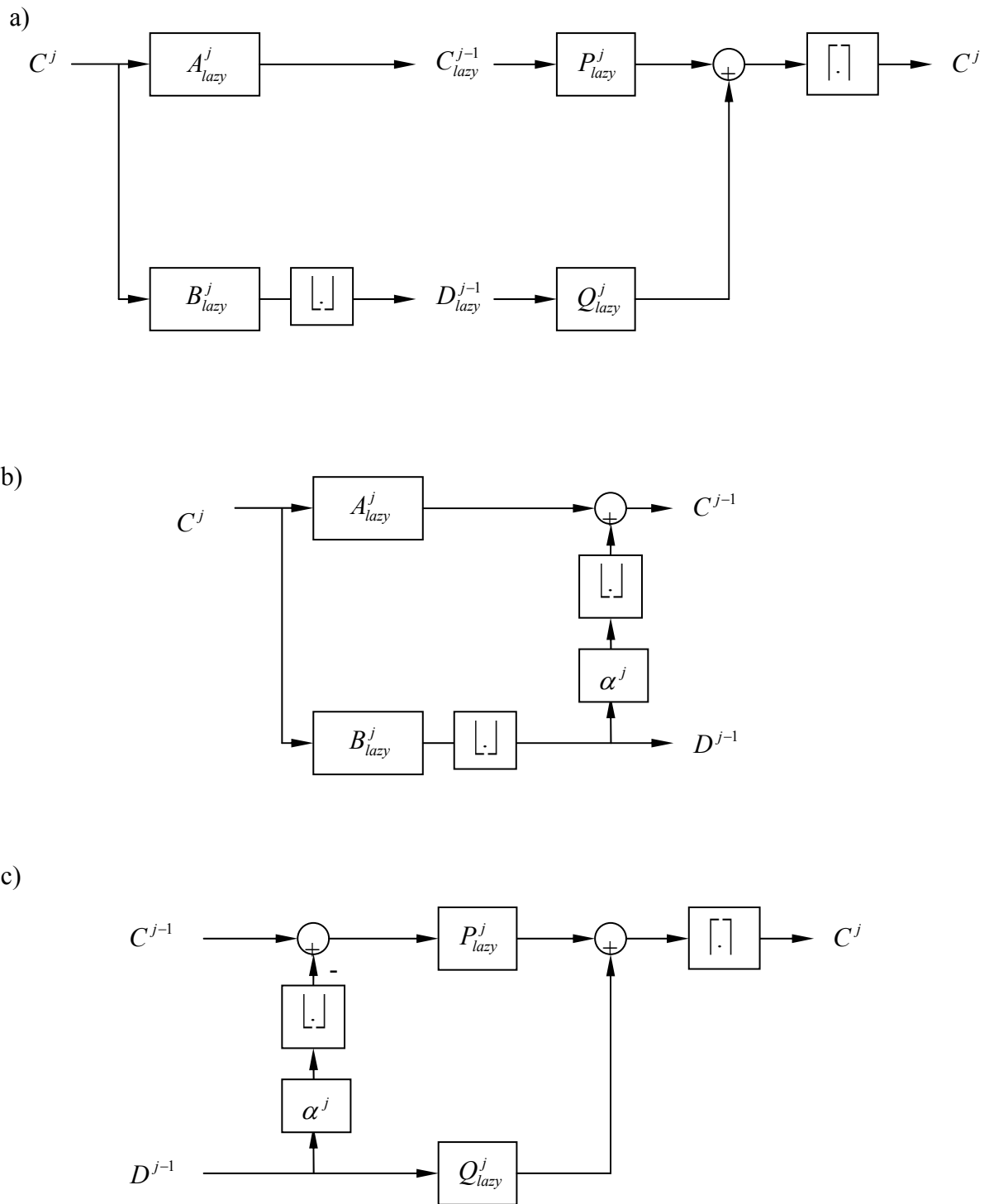


Figure 3.7.1 : Transformée en ondelettes utilisant l'arithmétique entière :

- a) Synthèse – analyse “lazy”
- b) Analyse (avec lifting)
- c) Synthèse (avec lifting)

### 3.8 Compression sans perte

#### 3.8.1 COMPRESSION DE LA CONNECTIVITE DU MAILLAGE

En considérant un maillage  $M^j$  possédant  $v^j$  sommets et  $n^j$  faces, la quantité d'information en bits par face représentant la connectivité du maillage est égale à :

$$B = 3 \cdot \log_2(v^j) \quad \text{II-53}$$

L' algorithme proposé permet une compression efficace de ce volume de données, car les seules informations nécessaires à la reconstruction d'un maillage à la plus haute résolution sont :

- la connectivité du maillage de résolution la plus basse (comportant peu de faces)
- les descriptions des différentes étapes de subdivision et de permutation d'arêtes.

On vérifie facilement que la quantité d'information nécessaire varie entre 2 et 4 bits par face :

Pour un maillage donné  $M^j$ , le codage des informations de subdivision pour reconstruire le maillage  $M^{j+1}$  nécessite 3 types d' informations:

- Les données relatives aux subdivisions, soit 1 bit par arête de  $M^j$ , pour savoir si, pour une arête donnée, un nouveau sommet doit y être créé.
- Des données supplémentaires, pour chaque face qui se subdivise en 3. En effet, il existe 2 moyens de subdiviser une face en trois si on connaît les arêtes où sont créés les 2 nouveaux sommets, comme le montre la figure 3.8.1.1.a . Ce volume de données sera donc de 1 bit pour chaque face subdivisée en trois
- Des données supplémentaires pour indiquer où des permutation d'arêtes ont été effectuées : 1 ou 2 bits par face subdivisée en 2, selon qu'il y ait ou non permutation d'arête. Si pour la face considérée il n'y a pas de permutation d'arête, 0 est transmis. Sinon, 1 est transmis et le bit transmis ensuite indique l'arête qui doit être permutée (voir figure 3.8.1.1.b).

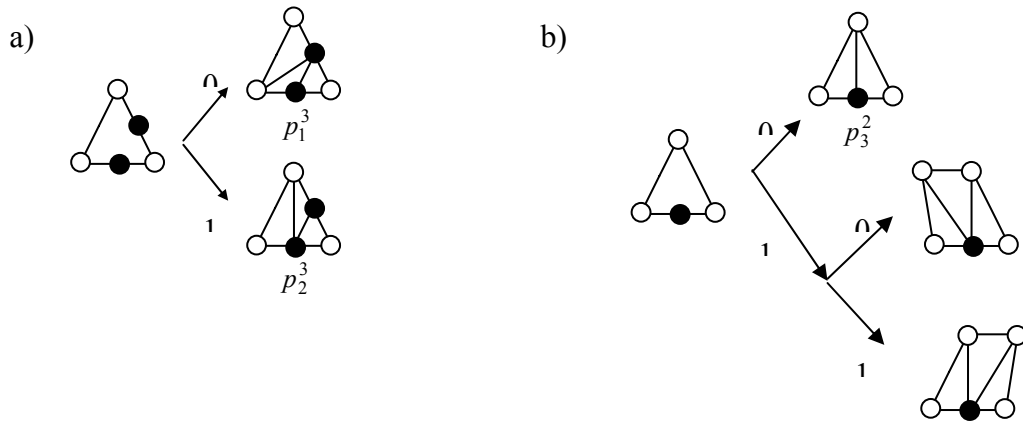


Figure 3.8.1.1 : Signification des données supplémentaires

Le nombre total de bits  $B_s^j$  nécessaire à la description du processus de subdivision peut donc être majoré ainsi :

$$B_s^j < e^j + 2.n^j \quad \text{II-54}$$

En utilisant l'équation d'Euler pour les maillages triangulaires conformes, on peut réaliser l'approximation suivante: le nombre de sommets  $v^j$ , le nombre d'arêtes  $e^j$  et le nombre de faces  $n^j$  d'un maillage donné  $M^j$ , sont liés par (cette propriété est valable uniquement lorsque le maillage possède peu de bords, voir [ROSS99]):

$$n^j \approx \frac{2}{3}e^j \approx 2v^j \quad \text{II-55}$$

L'inéquation II-54 devient alors :

$$B_s^j < 3.5n^j \quad \text{II-56}$$

Avec ces informations,  $n^{j+1} - n^j$  faces et  $v^{j+1} - v^j$  sommets ont été créés durant le processus de subdivision . La quantité moyenne d'information  $\Delta F^j$  (en bits) nécessaire pour créer une face vaut donc:

$$\Delta F^j = \frac{B_s^j}{n^{j+1} - n^j} = \frac{B_s^j}{(r^j - 1)n^j} \quad \text{II-57}$$

La combinaison des équations II-56 et II-57 permet d'établir la relation finale:

$$\Delta F^j < \frac{3.5}{r^j - 1} \quad \text{II-58}$$

L'efficacité du schéma de codage dépend donc essentiellement de la capacité de l'algorithme à fusionner les faces du maillage considéré, déterminé par le taux de fusion  $r^j$ .

En prenant l'exemple du maillage simplifié  $M^{j-1}$  de la figure in figure 3.5.2.h, La quantité totale d'information nécessaire pour la reconstruction de la connectivité du maillage de résolution supérieure  $M^j$  se décompose en:

- 18 bits pour la création des sommets fils (1 bit pour chaque arête). Après cette étape, La construction des 5 faces colorées en gris dans la figure 3.8.1.2.a est directe.
- 2 bits pour les deux faces subdivisées en 3 (figure 3.8.1.2.b)
- 5 bits pour les 5 faces subdivisées en deux (un seul de ces cinq bits vaudra 1 car une seule face nécessite une permutation d'arête, figure 3.8.1.2.c)
- 1 bit pour la permutation d'arête (figure 3.8.1.2.d).

Finalement, il faut 26 bits pour reconstruire la connectivité de  $M^j$  ( figure 3.8.1.2.d), i.e. pour ajouter 12 faces au maillage  $M^{j-1}$ . Ces 12 faces supplémentaires auront donc été codées avec une moyenne de  $\frac{26}{12} \approx 2.17$  bits par face.

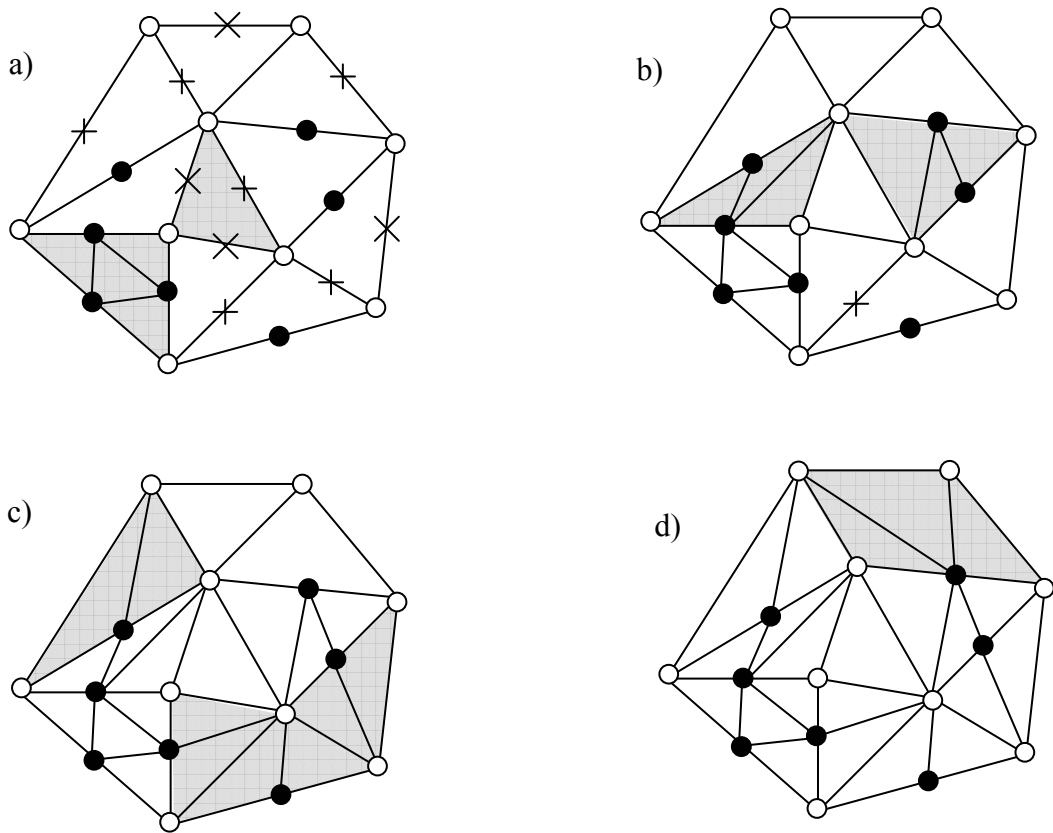


Figure 3.8.1.2 : reconstruction de la connectivité du maillage

### *Cas optimal*

Une compression optimale est obtenue lorsque toutes les faces sont fusionnées par groupe de 4 :

$$r_{opt}^j = 4 \quad \text{II-59}$$

Le volume d'information pour la reconstruction se limite alors à 1 bit par arête du maillage, car aucune face n'y est subdivisée en 3 ni en 2 :

$$B_{opt}^j = e^j \quad \text{II-60}$$

Le coût de codage optimal par face vaut donc :

$$\Delta F_{opt}^j = \frac{B_{opt}^j}{n^{j+1} - n^j} = \frac{B_{opt}^j}{(r_{opt}^j - 1)n^j} = \frac{e^j}{(4-1)n^j} = \frac{1}{3} \frac{e^j}{n^j} \approx \frac{1}{3} \frac{3}{2} \quad \text{II-61}$$

$$\Delta F_{opt}^j \approx 0.5 \text{ bits / face}$$

### **3.8.2 COMPRESSION DE LA GEOMETRIE DU MAILLAGE**

La méthode proposée est efficace pour la compression de la géométrie des maillages : la décomposition en ondelette, utilisée pour réaliser les approximations, transforme les coordonnées des sommets du maillage en coefficients d'ondelettes dont l'histogramme est concentré autour de zéro. Cette distribution des coefficients est donc propice à l'emploi d'un codeur entropique.

Dans nos travaux, nous avons utilisé un codeur arithmétique adaptatif pour comprimer les coefficients d'ondelette.

Dans les résultats qui suivent, tous les coefficients d'ondelettes sont transmis lors d'un passage d'un niveau de résolution à un autre. Pour une transmission progressive, on peut ne transmettre que les coefficients d'ondelettes les plus significatifs, afin d'améliorer l'efficacité de la compression (qui est alors une compression avec pertes). Les coefficients d'ondelettes non transmis peuvent toutefois être transmis ultérieurement, si on désire une reconstruction exacte. Un tel schéma de codage utilisé dans le cas de la subdivision régulière est décrit dans [BELZ98].

### 3.9 Résultats

Cette section montre les résultats obtenus sur divers maillages :

- Figure 3.9.1 : Une sphère synthétique, dont la connectivité est issue d'une série de subdivisions appliquées sur un octaèdre.
- Figure 3.9.2 : Une partie de la structure interne d'une vertèbre humaine, dont la forme est complexe et fine. Ce maillage a été créé à l'aide de l'algorithme décrit dans [LOTJ98]
- Figure 3.9.3 : Un maillage extrait d'un volume d'os par l'algorithme « marching cubes » [LORE87]. Sa connectivité est fortement irrégulière.
- Figure 3.9.4 : Le maillage « blob.wrl » disponible sur le réseau mondial.

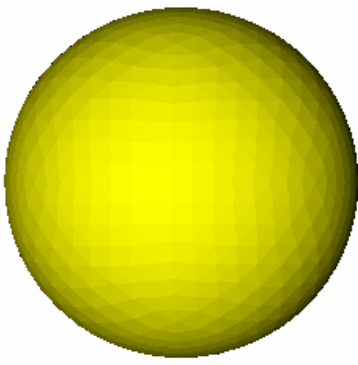
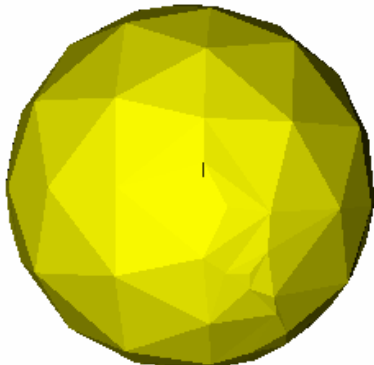
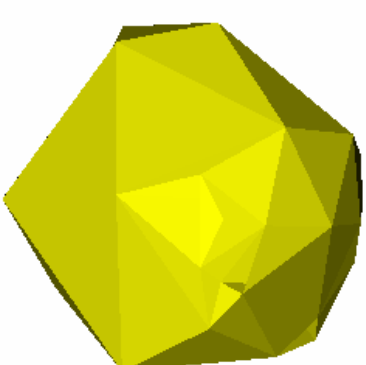



		
Maillage original (24 niveaux) 2048 faces	Approximation au niveau 20 254 faces	Approximation au niveau 17 90 faces
		
Approximation au niveau 13 60 faces	Approximation au niveau 6 32 faces	Approximation au niveau 0 12 faces

Figure 3.9.1 : résultat sur la maillage d'une sphère (connectivité de subdivision)

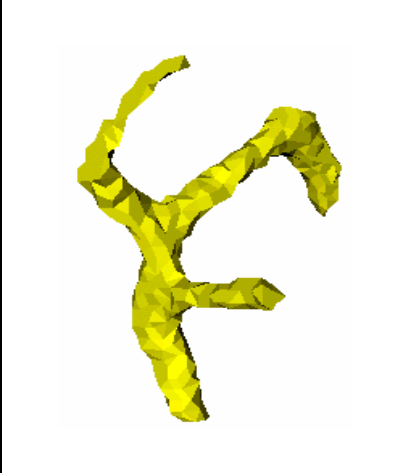

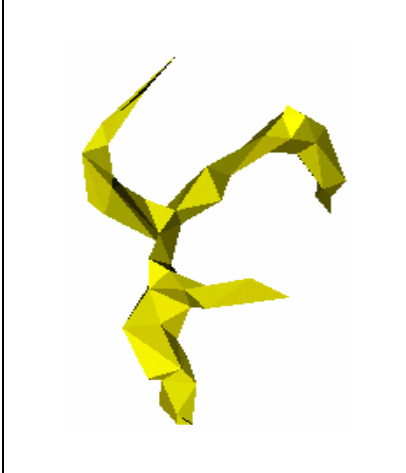

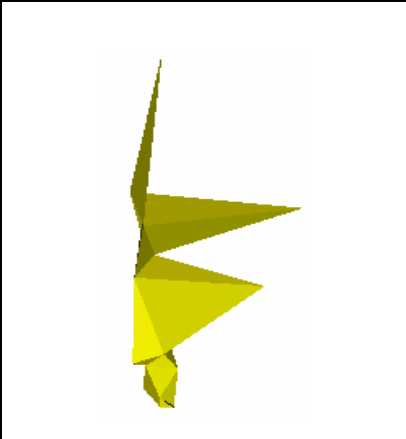
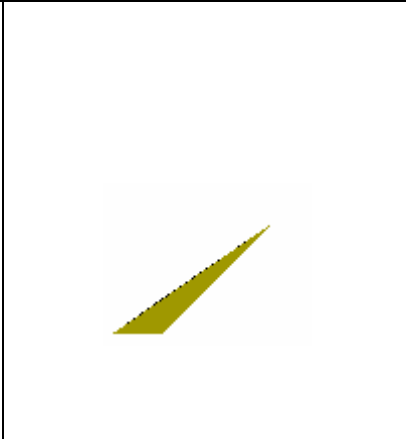
		
Maillage original (niveau 23) 1084 faces	Approximation au niveau 21 330 faces	Approximation au niveau 19 132 faces
		
Approximation au niveau 15 62 faces	Approximation au niveau 11 38 faces	Approximation au niveau 0 4 faces

Figure 3.9.2 : résultat sur une partie de la structure interne d'une vertèbre

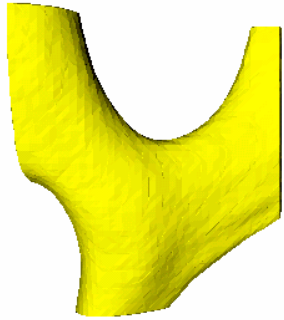
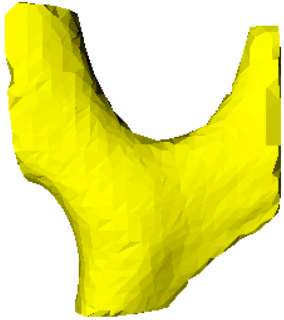

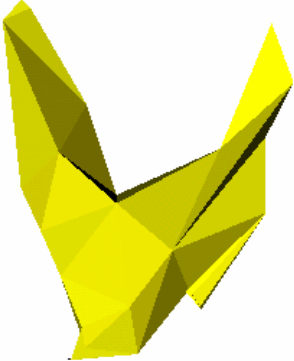
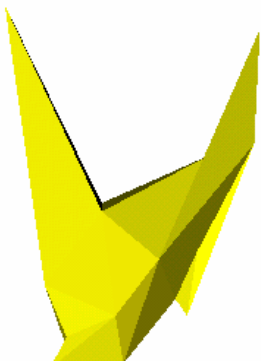

		
Maillage original (niveau 25) 9478 faces	Approximation au niveau 22 2170 faces	Approximation au niveau 20 660 faces
		
Approximation au niveau 15 58 faces	Approximation au niveau 10 28 faces	Approximation au niveau 0 4 faces

Figure 3.9.3 : résultat sur un maillage « os » crée par « marching cubes »

Maillage	Nombre de niveaux de résolution	Taille originale (bits)	Taille compressée (bits)	Taux de compression (sans perte)	Coût de codage moyen (bits/face)	Coût moyen (connectivité en bits/face)	Coût moyen (géométrie, en bits/face)	Dynamique des coordonnées de sommets
Sphère	23	113754	34764	3.3	16.75	<b>2.06</b>	14.69	15 bits
Vertèbre	24	41768	7016	5.9	6.47	<b>2.62</b>	3.85	6 bits
Os	25	492908	65660	7.5	6.92	<b>1.91</b>	5.10	10 bits
Blob	23	835588	112755	7.4	7.02	<b>2.24</b>	4.78	8 bits

Tableau 3.9 : résultats de compression

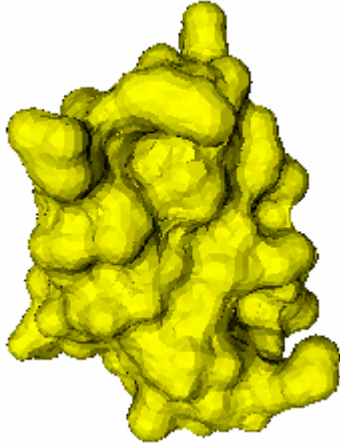
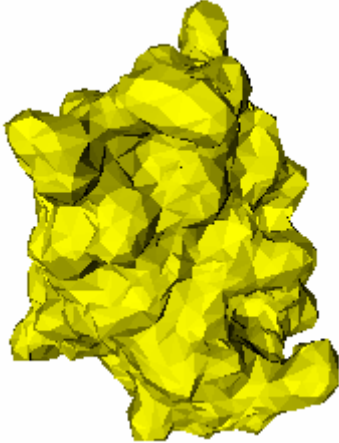


	
Maillage initial (niveau 22) 16068 faces	Approximation au niveau 20 442 faces
	
Approximation au niveau 17 776 faces	Approximation au niveau 0 90 faces

Figure 3.9.4 : résultat sur le maillage “blob”

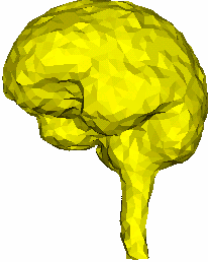

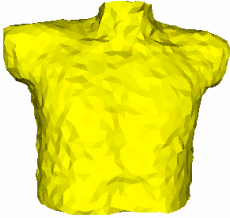
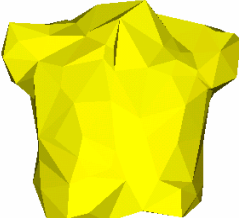


<p>Maillage original</p>  <p>3584 faces 8 niveaux</p>	<p>Niveau 4</p>  <p>524 faces</p>	<p>Fichier original: 19493 octets</p> <p>Fichier compressé : 3246 bytes</p> <p>Taux de compression: 6.1 : 1</p>
<p>Maillage original</p>  <p>4454 faces 10 niveaux</p>	<p>Niveau 5</p>  <p>306 faces</p>	<p>Fichier original : 25894 octets</p> <p>Fichier compressé: 4062 octets</p> <p>Taux de compression : 6.4 : 1</p>
<p>Maillage original</p>  <p>1916 faces 10 niveaux</p>	<p>Niveau 5</p>  <p>168 faces</p>	<p>Fichier original: 9705 octets</p> <p>Fichier compressé: 1801 octets</p> <p>Taux de compression 5.4 : 1</p>

Figure 3.9.5 : résultats supplémentaires

La figure 3.9.5 montre des résultats supplémentaires obtenus avec des maillages représentant un cerveau, un poumon et un thorax, avec un niveau de représentation intermédiaire, montrant la qualité de l'approximation.

### *Discussion sur les résultats*

Le tableau 3.9 montre les résultats obtenus en compression sans pertes. Il faut remarquer que pour un maillage donné, le fichier compressé permet la représentation de tous les niveaux de résolution construits, jusqu'au niveau le plus haut.

- La sphère de la figure 3.9.1 possède une connectivité de subdivision et 2048 faces. Les résultats optimaux seraient donc obtenus en créant 4 nouveaux niveaux de résolution, avec des maillages possédant 512, 128, 32, et 8 faces. L'algorithme de fusion de triangles est donc relativement loin de la solution optimale, car il a créé 23 niveaux de résolution inférieurs. Une des raisons de l'incapacité à atteindre la configuration optimale vient du fait que l'algorithme de fusion commence toujours par garder une face du maillage original intacte, alors que pour la configuration optimale, toutes les faces sont fusionnées 4 par 4. De plus, l'algorithme n'a pas été capable de simplifier le maillage jusqu'à la forme ultime : une pyramide.
- La partie de vertèbre de la figure 3.9.2 possède une forme arborescente très fine. L'algorithme a néanmoins été capable de simplifier le maillage jusqu'au cas le plus simple : une pyramide. Le coût moyen de codage de la connectivité de ce maillage est de 2.62 bits/face, ce qui est le maximum obtenu sur les maillages testés.
- Le maillage fabriqué avec l'algorithme « marching cubes », figure 3.9.3 a été simplifié jusqu'à la forme ultime: une pyramide.
- Le maillage « blob » a été simplifié jusqu'à un maillage possédant 90 faces.

### *Codage de la géométrie*

La taille nécessaire au codage de la géométrie dépend de la dynamique des coordonnées des sommets (tableau 3.9). Le taux de compression sans perte pour la géométrie varie entre 3 et 7.

### *Codage de la connectivité*

Comme nous l'avons abordé précédemment, l'efficacité de la compression de la connectivité du maillage dépend de la faculté de l'algorithme de fusion à fusionner les faces du maillage en créant un minimum de niveaux de résolution (pour maximiser les taux de fusion  $r^j$ ). Les maillages les plus propices à ces fusions (ceux dont la connectivité est la plus régulière) seront donc les mieux compressés. Pour l'ensemble des maillages testés, le coût de codage de la connectivité varie entre 1.9 et 2.62 bits par face. Le maillage le mieux compressé est le maillage Blob (figure 3.9.4), et le moins bon résultat est obtenu sur le maillage de la structure interne d'une vertèbre (figure 3.9.3).

### 3.10 *Conclusion et perspectives*

Dans cette partie, nous avons proposé un nouvel algorithme d'analyse multirésolution par décomposition en ondelettes. Par extension du cas classique des maillages à connectivité de subdivision [LOUN97], notre approche peut s'appliquer sur tous types de maillages.

Cette extension a nécessité l'introduction d'un nouveau schéma de subdivision, permettant la simplification de maillages quelconques, par la résolution d'un problème inverse discret, dont nous proposons une solution par l'utilisation d'un algorithme de conquête fusionnant itérativement les faces du maillage.

Une réflexion mathématique sur la décomposition en ondelettes des coordonnées des sommets représentables par des entiers nous permet de garantir une reconstruction exacte. La décomposition en ondelettes couplée à un codage compact des subdivisions à appliquer sur le maillage constitue un outil efficace de transmission progressive, allant jusqu'à la compression sans perte, si tous les niveaux de résolution sont transmis.

Plusieurs axes de recherche apparaissent :

- Une étude poussée du problème inverse de fusion des faces des maillages est à envisager. En effet, l'efficacité de compression de la méthode proposée dépend principalement de la capacité de l'algorithme à fusionner les faces entre-elles. En opposition à l'approche actuelle utilisant un algorithme de conquête, une approche plus globale de fusion des faces du maillage pourrait donner des résultats plus proches du cas optimal (fusions 4 :1).
- Les permutations d'arête effectuées pour « débloquer » l'algorithme dans certaines configurations diminuent la qualité de l'approximation et obligent un codage supplémentaire non négligeable, ce qui pénalise l'efficacité de compression de notre algorithme. Un algorithme de fusion de triangles n'utilisant pas ces permutations (ou en nombre très limité) améliorerait l'efficacité de l'approche proposée.
- Actuellement, le codage des coefficients d'ondelettes représentant la géométrie du maillage est réalisé par un codeur arithmétique adaptatif du fait de sa simplicité d'utilisation. Ce type de codeur propose une gestion transparente de la table de probabilités utilisée pour le codage entropique, car la table de probabilités est construite automatiquement par le codeur et le décodeur. La figure 3.10 montre l'histogramme des coefficients d'ondelettes entre -50 et +50 pour le maillage « os » (plus de 99% des coefficients sont contenus dans cet intervalle), qui est fortement concentré autour de 0, malgré l'aspect accidenté de la surface de ce maillage. Un codage utilisant une modélisation statistique des coefficients (par une fonction gaussienne, par exemple) améliorerait sans doute le taux de compression de

la géométrie, car les codeurs adaptatifs construisent leur table de probabilité au fur et à mesure de la compression. Ils sont donc sous optimaux lorsque la table de probabilité n'évolue pas au cours du codage.

- La qualité de l'approximation par décomposition en ondelettes des coordonnées des maillages reste à évaluer quantitativement. L'emploi d'un outil tel que *Metro* [CIGN98] semble ici approprié, afin de relever les points faibles de la méthode. Actuellement, la fusion des faces est effectuée uniquement en prenant en compte la connectivité du maillage, par maximisation du nombre de faces fusionnées afin d'offrir le meilleur taux de compression sans pertes. Or, dans un soucis d'approximation efficace, les zones du maillage à forte courbure devraient être moins simplifiées que les zones peu accidentées. La prise en compte de la courbure locale des maillages au cours de la fusion sera donc nécessaire si on désire optimiser l'algorithme au niveau débit/distorsion pour un niveau de reconstruction donné.

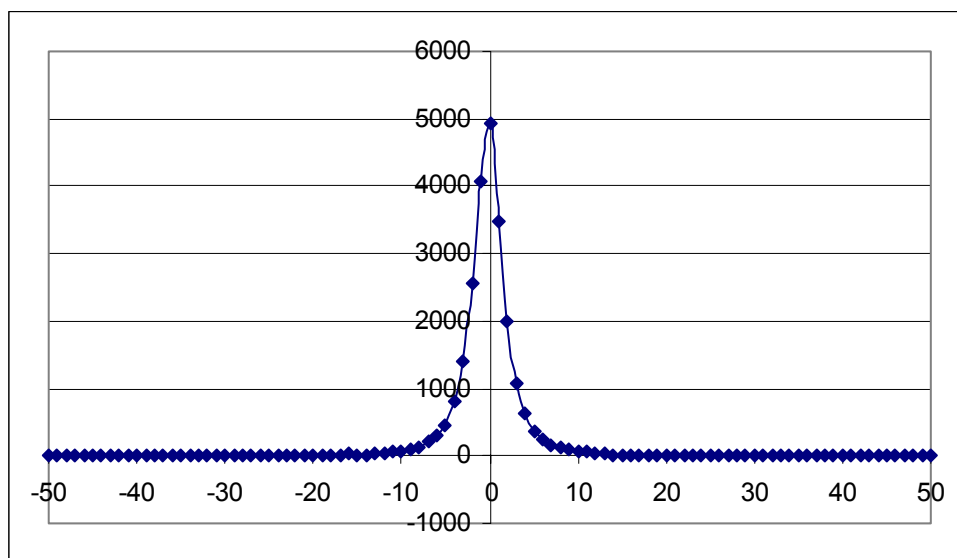


Figure 3.10 : Histogramme des coefficients d'ondelettes pour le maillage «blob» tronqué à l'intervalle [-50,50]

# **Chapitre III**

## **Maillages actif pour la segmentation et le suivi d'objets vidéo**

## ***1 INTRODUCTION***

La récente émergence de canaux de communication à haut débit tels que les réseaux câblés ou les satellites rend la transmission des informations audiovisuelles de plus en plus aisée. De plus, la puissance des ordinateurs évoluant exponentiellement (selon la loi de Moore), les informations transmises peuvent être de nature très différente et très complexe, afin d'offrir à l'utilisateur un contenu le plus riche possible, mélangeant images, voix et séquences vidéo. Elles peuvent être plus flexibles : plusieurs niveaux de qualité doivent être disponibles pour un même document, en fonction des ressources disponibles. De plus, une interactivité entre l'utilisateur et le document doit être envisagée.

### *1.1 Historique*

Le premier format standard de transmission de la vidéo numérique pour le grand public a été le format MPEG 1, qui propose une résolution de 352\*240 pixels, 24 images par seconde, équivalente au format analogique VHS, pour un débit variant entre 1 et 1,5 Mbits/s. Ce format est notamment utilisé par le Vidéo CD (VCD), offrant 74 minutes de vidéo sur un support optique de type CDROM, d'une capacité de 650 Mo. Mais le standard MPEG-1, arrivé très tôt (en 1990) sur le marché s'est peu répandu, mis à part dans les pays de l'Asie du sud-est.

L'élargissement des canaux de communication et l'augmentation de la puissance de calcul des processeurs de signaux a rendu possible la création d'un deuxième standard de compression : MPEG-2, offrant une meilleure qualité qu'une séquence MPEG-1, avec un débit se situant entre 4 et 15 Mbits/s. Ce standard c'est beaucoup répandu et est notamment utilisé pour la transmission de programmes télévisés par satellite et par le DVD (Digital Versatile Disc), un support optique d'une capacité allant jusqu'à 17 Go.

Un autre format (MPEG-3) pour la télévision haute définition (TVHD) a été imaginé, mais les performances du format MPEG-2 ont été jugées suffisantes pour la TVHD, et le standard MPEG-3 a donc été annulé [TEKA99].

### *1.2 Le standard MPEG-4*

Les formats MPEG-1 et MPEG-2 permettent la transmission de documents audiovisuel sous une forme brute (un flux vidéo et un flux audio), mais sont incapables de gérer simultanément plusieurs contenus. Le nouveau standard MPEG-4 (dont une version finale a été délivrée en 1998 [MPEG98]) a été créé pour ce type de gestion.

En effet, on peut imaginer une scène vidéo, où les différents objets présents n'utilisent pas tous le même type de représentation. La figure 1 montre un exemple de scène de type MPEG-4, comprenant plusieurs « objets », de nature visuelle ou sonore, dont la description se fait par flux binaires multiplexés dans le flux MPEG-4.

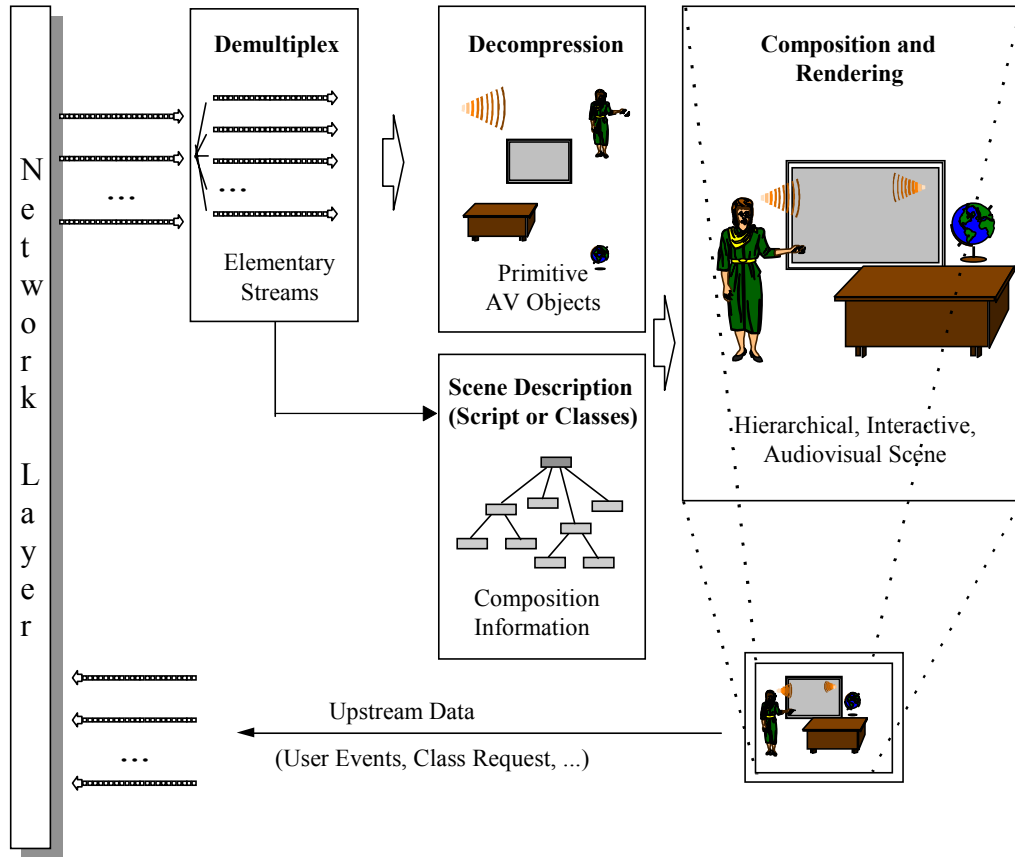


Figure 1.1 : décomposition d'une scène multimédia MPEG 4

Tiré du rapport MPEG 4 :N2323 (mars 98, Dublin)  
 plus d'information disponible à l'URL : <http://cselt.it/mpeg>

### 1.3 Le standard MPEG-7

La séparation des objets d'un même document multimédia offre une plus grande flexibilité quant à la recherche, au traitement et au classement des documents. MPEG-7 a été élaboré dans cette optique, proposant la description standardisée des objets (audio ou vidéo) présents dans un document MPEG-4, faisant, par exemple, le lien entre le document et une application nécessitant la recherche d'information dans une base de données multimédia. La version finale du format MPEG-7 est prévue pour l'année 2001.

#### 1.4 *Objectif du travail*

Dans le cadre du projet RNRT (Réseau National de Recherche en Télécommunications) OSIAM (Outils de Segmentation d'Images Animées pour MPEG-4/7), notre objectif a été la création d'un algorithme de segmentation d'objets vidéo.

En effet, le standard MPEG-4 fournit un moyen efficace de gérer les différents objets vidéo d'une même scène, mais ne donne aucun moyen pour la création de ces objets. Si on désire convertir une séquence vidéo déjà existante du format MPEG-2 vers le format MPEG-4 deux solutions sont à notre disposition :

- soit la vidéo est laissée telle quelle (MPEG-4 gère sans problèmes le format MPEG-2). Il s'agira alors d'une simple encapsulation, mais aucune fonctionnalité supplémentaire ne sera proposée à l'utilisateur.
- Si on désire plus de fonctionnalités, il faut pouvoir extraire les différents objets présents dans la séquence, pour convertir le flux MPEG-2 en plusieurs flux MPEG-4 multiplexés. La description d'un objet vidéo utilise principalement en 2 entités : l'objet vidéo, et un plan alpha (« alpha plane » en anglais), binaire, qui indique la forme de l'objet dans la séquence. La figure 1.2 montre un exemple de segmentation sur une image de la séquence « van » (figure 1.2.a). Le camion est extrait du fond de l'image (figure 1.2.b) et le plan alpha construit à partir de la segmentation (figure 1.2.c).

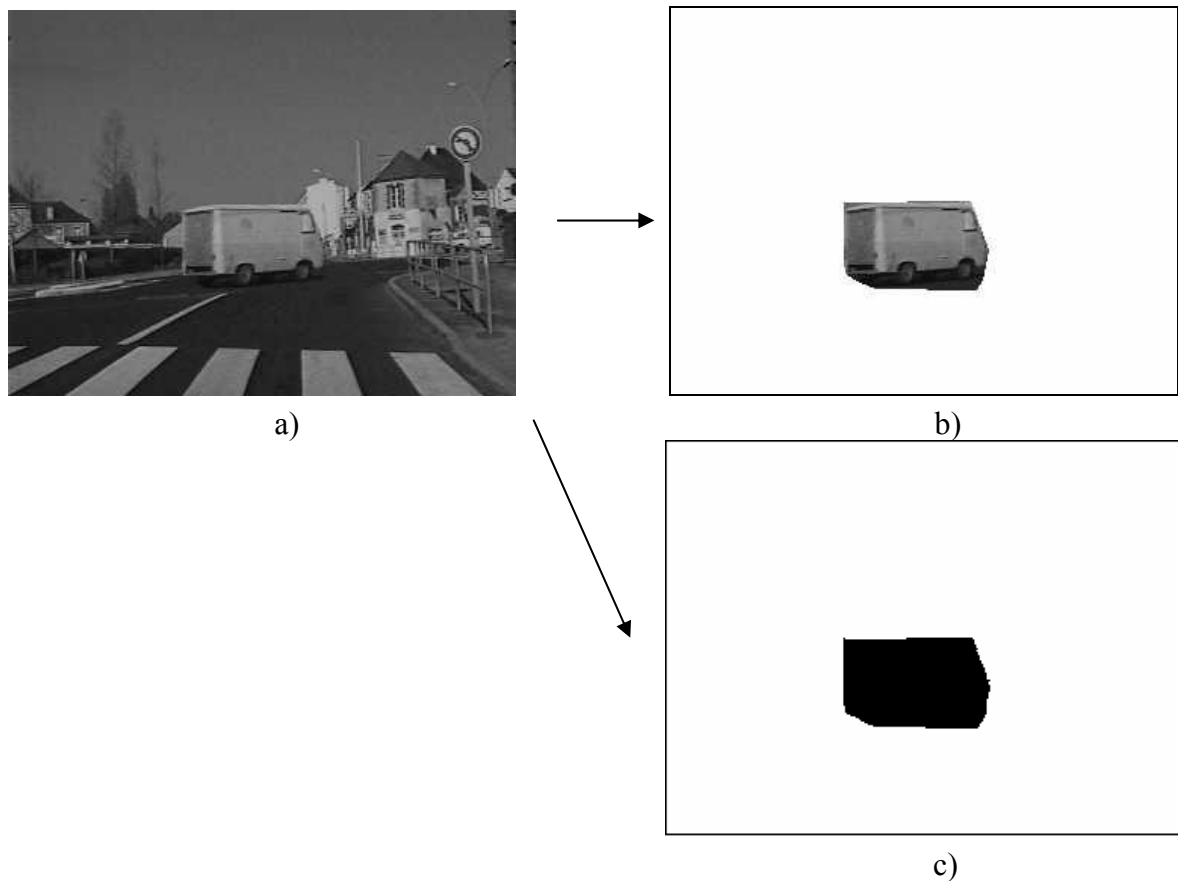


Figure 1.2 : exemple de segmentation spatio-temporelle :

- a) image originale
- b) segmentation du camion dans la scène
- c) « alpha plane » pour l'objet camion

La création d'outils de segmentation semble donc nécessaire si on désire exploiter complètement les fonctionnalités du format MPEG-4.

Mais la segmentation de séquences vidéo reste un problème ouvert et, à l'heure actuelle, aucun algorithme ne permet une telle segmentation dans le cas général. De plus, la notion d'objet vidéo est très subjective, et nécessite une approche *sémantique*. En effet, si on prend l'exemple de la séquence « van » (figure 1.2.a), on peut considérer que le camion est un objet vidéo, et que le fond immobile est un autre objet vidéo. Mais on pourrait aussi découper le fond immobile en plusieurs objets immobiles : la route, les bâtiments, le ciel, et partager le camion en plusieurs objets ayant des mouvements différents : la carrosserie (mouvement de translation) et les roues (mouvement cycloïde).

La figure 1.3 montre un cas réel de séquence MPEG-4, construite à partir de la séquence vidéo « stefan » de type MPEG. Le joueur de tennis a été préalablement séparé du cours de tennis. On obtient alors deux objets vidéo :

- Le cours de tennis, sous forme d'image fixe (figure 1.3.a).

- Le joueur de tennis, sous forme d'image animée (figure 1.3.b).

Ces deux objets vidéo peuvent ensuite être rassemblés en une seule séquence, par le biais du format MPEG-4 (figure 1.3.c). Il faut remarquer que dans ce cas, la « caméra » virtuelle qui sert à la représentation finale bouge au cours de la séquence (une partie seulement du fond est visible). Cette technique de synthèse d'image est appelée mosaïque (« mosaicking » en anglais, voir [TOKL00] et [NICO01])

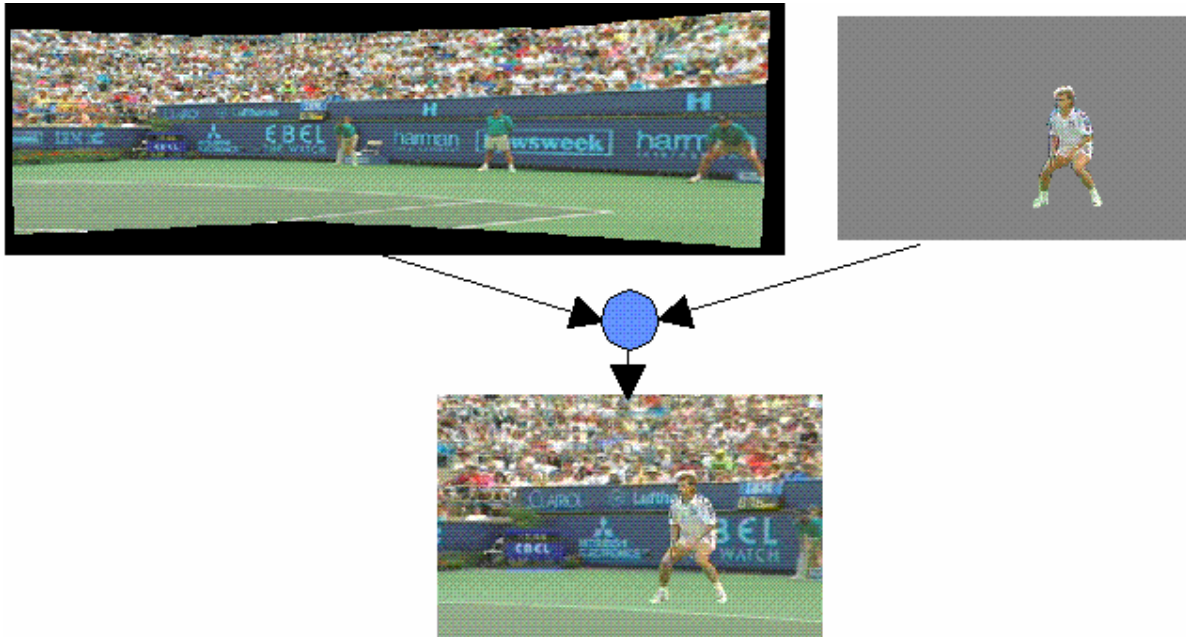


Figure 1.3 : synthèse d'une séquence MPEG-4 :

- Objet « fond »
- Objet « joueur de tennis »
- Scène MPEG-4

Notre objectif est de fournir un outil le plus général possible, en utilisant les maillages actifs comme cadre de recherche pour la segmentation de séquences vidéo. Cette segmentation génère les plans alpha donnant la description des différents objets présents dans une séquence.

## 2 ETAT DE L'ART

La segmentation vidéo est un problème très complexe, car pour segmenter correctement une séquence, plusieurs types d'information doivent être pris en compte :

- Les caractéristiques spatiales des objets (segmentation spatiale).
- La localisation et le déplacement des différents objets (segmentation par le mouvement)
- Les interactions entre les objets
- La cohérence temporelle entre les segmentations des images successives (suivi).

Les sections suivantes explicitent les différentes approches existantes pour segmenter les séquences vidéo. Le nombre élevé de travaux dans les 4 domaines cités précédemment rend l'exhaustivité de la description très ardue, et nous ne donnerons ici parfois que les approches généralement employées.

### 2.1 Segmentation spatiale

La segmentation spatiale sépare une image donnée en plusieurs partitions distinctes [COCQ95], que l'on construit à partir des caractéristiques internes à l'image comme:

- Les zones homogènes au sens de la luminance ou de la couleur
- Les contours présents dans l'image
- La texture

Les structures maillées ont rarement été employées pour effectuer une segmentation spatiale et le plus souvent été utilisées sur les images fixes dans un but de compression à bas débit. L'approximation d'une image par un maillage nécessite de définir une représentation des différentes régions que forment les cellules du maillage. Cette représentation peut avoir plusieurs formes :

- Une représentation uniforme : chaque cellule est remplie par une valeur constante, déterminée par la valeur moyenne des pixels qu'elle contient [BENO99]
- Une représentation interpolée : les valeurs des pixels de la cellule considérée sont interpolées à partir des valeurs des nœuds formant la cellule. L'interpolation peut être par exemple linéaire [LECH98], [MARQ00a], une interpolation de Shepard [SHEP68] ou une interpolation de Shepard modifiée [LEFL97]
- Une représentation à base de textures comprimées et plaquées sur les cellules [LEBU98], [BENO99]

## 2.2 Estimation de mouvement

L'estimation de mouvement dans une séquence d'image donne des informations sur la vitesse des différentes régions présentes dans une image. Une revue des approches développées pour l'estimation du mouvement est proposée dans [STIL99].

L'approche généralement employée pour estimer le mouvement présent entre deux images d'une séquence repose sur la conservation de l'intensité de l'image au cours du temps [HORN81]:

Soit un point  $p(x,y)$  à l'instant  $t$  dans la séquence ; son intensité est  $I(x,y,t)$ . Si on considère que son déplacement  $dp(dx,dy)$  au cours de l'intervalle de temps  $dt$  n'a pas modifié son intensité, on peut écrire :

$$I(x,y,t) = I(x+dx, y+dy, t+dt) \quad \text{III-1}$$

### 2.2.1 APPROCHE DIFFERENTIELLE : FLUX OPTIQUE

#### a) Principe

Le développement de Taylor de l'intensité de l'image  $I(x,y,t)$  nous donne également :

$$I(x+dx, y+dy, t+dt) = I(x,y,t) + \frac{\partial I}{\partial x} dx + \frac{\partial I}{\partial y} dy + \frac{\partial I}{\partial t} dt + o(\varepsilon) \quad \text{III-2}$$

où  $\varepsilon$  contient les termes d'ordre supérieur du développement de Taylor et tend vers 0 lorsque  $dt$  tend vers 0. En considérant que les grandeurs  $dx$ ,  $dy$  et  $dt$  sont suffisamment petites et que l'équation III-1 est respectée, il vient :

$$\begin{aligned} \frac{\partial I}{\partial x} dx + \frac{\partial I}{\partial y} dy + \frac{\partial I}{\partial t} dt &= 0 \\ \Leftrightarrow \frac{\partial I}{\partial x} \frac{dx}{dt} + \frac{\partial I}{\partial y} \frac{dy}{dt} + \frac{\partial I}{\partial t} &= 0 \\ \Leftrightarrow uI_x + vI_y + I_t &= 0 \end{aligned} \quad \text{III-3}$$

où  $I_x$ ,  $I_y$ ,  $I_t$  sont les dérivées partielles spatio-temporelles de l'intensité lumineuse au point  $p(x,y)$ , et  $u_p = (u,v)$  le vecteur vitesse associé à ce point.

Estimer le mouvement dans l'image par la méthode du flux optique [HORN81] reviendra à trouver le champ dense de vecteurs vitesses  $U$  satisfaisant l'équation III-3 en chaque point de l'image.

b) Problème d'ouverture

L'équation III-3 pose un problème de sous dimensionnement : en chaque point de l'image, deux inconnues sont à identifier alors que l'on ne dispose que d'une équation : la solution du problème n'est pas unique.

Ce problème appelé *problème d'ouverture* se traduit par le fait que l'on ne peut estimer correctement le vecteur vitesse d'un point  $p(x, y)$  que dans la direction du gradient local de l'image.

La résolution de ce problème s'effectue généralement par une approche globale d'estimation du flux optique, par la minimisation d'une énergie  $E$  :

$$E = \iint_{x,y} (uI_x + vI_y + I_t)^2 dx dy \quad \text{III-4}$$

La minimisation de l'énergie  $E$  par la prise en compte du voisinage de chaque point de l'image permet de limiter les effets du problème d'ouverture, mais induit un moyennage du champ de vecteurs estimé, qui devient alors continu. Or les scènes à analyser possèdent généralement plusieurs éléments ayant des mouvements distincts. Le champ de vecteurs mouvements réel présente donc des discontinuités aux frontières entre les objets. On peut en déduire que l'estimation du mouvement aux frontières entre les objets sera erroné.

c) Limitations de l'approche différentielle

L'équation III-3 est basée sur le développement de Taylor de l'intensité lumineuse  $I(x, y, t)$  au premier ordre, et n'est donc une approximation acceptable que pour les petits mouvements. De plus, le calcul des dérivées spatio-temporelles de  $I(x, y, t)$  réalisé à partir de données échantillonnées est une source supplémentaire d'erreurs. Une approche multirésolution rend cependant une résolution du problème moins sensible aux problèmes de discrétisation et améliore l'estimation de grand déplacements dans l'image.

## 2.2.2 APPROCHE PAR MISE EN CORRESPONDANCE

### Principe

L'estimation de mouvement par mise en correspondance est l'approche la plus généralement employée. Elle est basée sur l'utilisation directe de la contrainte de conservation de l'intensité (voir équation III-1) dans sa version discrétisée :

$$I(x, y, t) = I(x + dx, y + dy, t \pm 1) \quad \text{III-5}$$

L'estimation du champ de vecteurs mouvements dans l'image au temps  $t$  revient ici à essayer d'estimer l'image au temps  $t$  à partir de l'image au temps  $t-1$  (on parle alors d'estimation arrière (backward matching en anglais)) ou à partir de l'image au temps  $t+1$  (estimation avant (forward matching)) : l'image au temps  $t$  est découpée en un ensemble de  $n$  régions  $R_i$ , auxquelles on affecte un mouvement propre modélisé par une transformation  $\tau_i$  (voir figure 2.2.1). L'estimation de mouvement des différentes régions s'effectue en minimisant l'erreur de compensation de mouvement au temps  $t$  (appelé DFD pour Displaced Frame Difference):

$$DFD_t(x, y) = I(x, y, t) - I(x + dx, y + dy, t \pm 1) \quad \text{III-6}$$

Ce qui se traduit globalement par la minimisation de l'énergie  $E_{comp}$  :

$$E_{comp} = \sum_{i=1}^n \left( \sum_{p(x,y) \in R_i} e(I(p, t) - I(\tau_i(p), t \pm 1)) \right) \quad \text{III-7}$$

Il faut alors définir plusieurs paramètres pour le calcul de cette énergie :

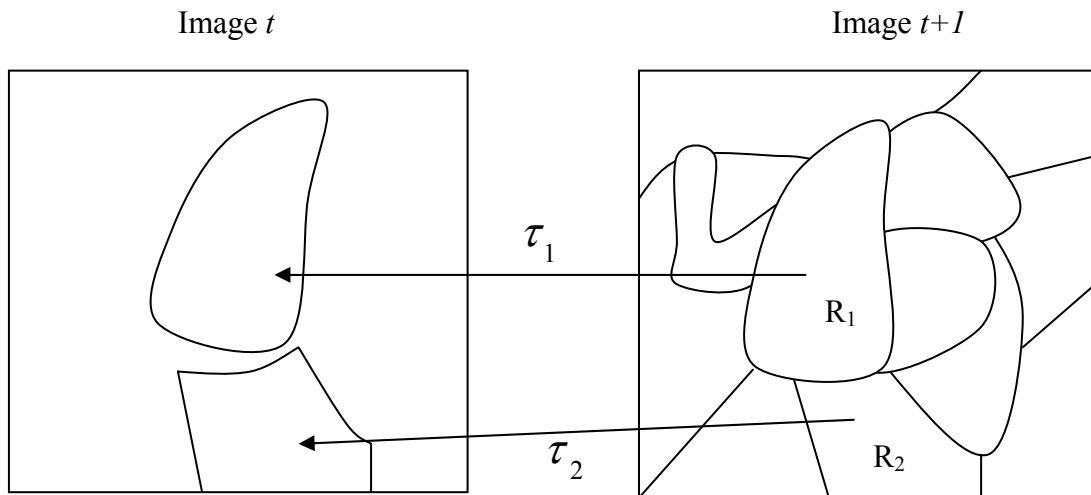
- Le terme d'erreur  $e(u)$ . Les deux formes les plus généralement employées sont l'erreur quadratique :  $e_{quad}(u) = u^2$  et l'erreur absolue  $e_{abs}(u) = |u|$ . L'emploi de l'erreur absolue se justifie souvent par sa complexité calculatoire moins importante.
- Le type de transformation  $\tau_i$  que l'on désire utiliser pour modéliser le mouvement affecté aux différentes régions  $R_i$ . Il peut être entre autres de type constant, affine ou polynomial. Le champ de mouvement alors estimé n'est pas forcément constitué de valeurs entières, et le calcul de  $I(\tau_i(p), t \pm 1)$  nécessite l'utilisation d'algorithmes d'interpolations. L'interpolation bilinéaire est la plus employée [WOLB92].

- Le partitionnement de l'image. Chaque région  $R_i$  se voit associé un type de mouvement unique. Le choix du découpage de l'image en régions est donc très important quant à la qualité de l'estimation du mouvement : il faudra veiller à ce que chaque région ne possède réellement un mouvement représentable par un jeu de paramètres uniques. Les discontinuités du champ réel de mouvement ne seront donc correctement estimées que si elles sont situées aux frontières entre les régions.
- Le choix de la direction de l'estimation (avant ou arrière, « forward » ou « backward » en anglais). Pour l'estimation de mouvement seule, les deux approches ont autant d'avantages que d'inconvénients que nous développerons par la suite. Pour la compression de séquences vidéo, l'estimation arrière est utilisée de préférence, car elle permet de prédire tous les pixels de l'image  $t+1$  à partir de l'image  $t$ . On parlera dans ce cas de compensation de mouvement (« motion compensation » en anglais)

La minimisation de l'énergie de compensation  $E_{comp}$  engendre nécessairement l'estimation du champ de mouvement dans l'image selon un partitionnement en régions  $R_i$ . Si ces régions sont fixes on dira que la méthode est rigide, dans le sens où le partitionnement est défini à l'avance, sans aucune prise en compte de la distribution du mouvement dans l'image. Si l'on estime le champ de mouvement et on construit un partitionnement adapté au contenu de l'image, ce qui dans ce cas revient à une segmentation de l'image par le mouvement, la méthode sera dite adaptative.

De nombreuses méthodes de partitionnement de l'image pour la mise en correspondance ont été développées. La plus simple est la mise en correspondance de blocs [JAIN81], à la plus évoluée effectuée une segmentation préalable de l'image en régions homogènes [SALE97], [PATE98]. Les sections suivantes détaillent quelques méthodes d'estimation de mouvement par mise en correspondance.

a)



b)

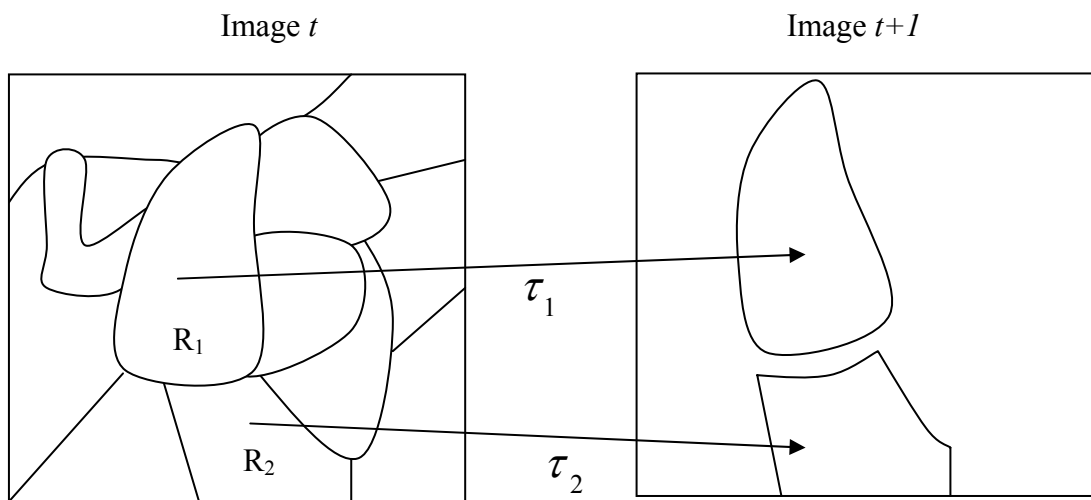


Figure 2.2.1 : mise en correspondance de deux régions  $R_1$  et  $R_2$  :

a) estimation arrière (entre les instants  $t$  et  $t-1$ )

b) estimation avant (entre les instants  $t$  et  $t+1$ )

### 2.2.3 MISE EN CORRESPONDANCE DE BLOCS

L'algorithme de mise en correspondance de blocs [JAIN81] utilise un découpage régulier de l'image en blocs carrés (de dimensions 16x16 pixels en général) et un modèle de mouvement constant pour tous les pixels d'un bloc donné (2 paramètres par bloc):

$$p(x, y) \xrightarrow{\tau_i} p'(x', y') \begin{cases} x' = x + u_i \\ y' = y + v_i \end{cases} \quad \text{III-8}$$

Avec ce modèle, le vecteur mouvement des différents blocs est estimé indépendamment d'un bloc à l'autre (figure 2.2.2).

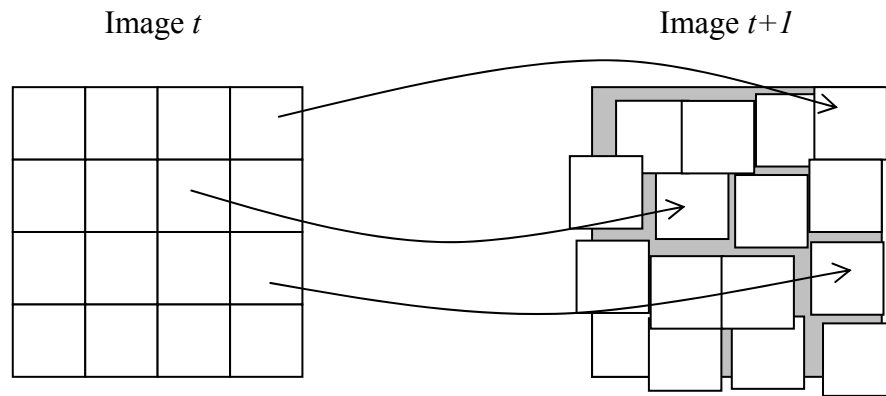


Figure 2.2.2 : mise en correspondance de blocs (estimation avant)

#### *Méthodes d'estimation*

Ces méthodes nécessitent l'estimation d'un vecteur à deux dimensions pour chaque bloc. Plusieurs approches sont envisageables pour la détermination du vecteur mouvement d'un bloc. La plus simple est la recherche exhaustive du vecteur donnant la meilleure compensation parmi un ensemble discret de vecteurs, par exemple l'espace des paires d'entiers  $[-16, 16] \times [-16, 16]$ , nous dirons que cette méthode est optimale. Des algorithmes plus rapides ont cependant été développés, afin de ne pas effectuer une recherche exhaustive, (par exemple l'algorithme « Diamond Search » [ZHU00], et le « Successive Elimination Algorithm », SEA [WANG99]). Ces algorithmes utilisent des tests comparatifs sur un échantillonnage de vecteurs candidats, entraînant un raffinement itératif du vecteur déplacement estimé.

La mise en correspondance de blocs présente de nombreux avantages : son partitionnement est régulier, sa complexité est relativement faible (2 paramètres à estimer par cellule) ; son rapport qualité de compensation/coût de codage l'a rendu très populaire pour la compression de séquences vidéo, c'est pourquoi il est utilisé dans les standards vidéo MPEG-1 et MPEG-2 entre autres.

## 2.2.4 MAILLAGES DYNAMIQUES

### a) Grilles quadrangulaires

La mise en correspondance par grilles (CGI pour « Control Grid Interpolation » en anglais) a été introduite par Sullivan [SULL91] : le mouvement de l'image est estimé à partir du déplacement des points de contrôle (ou nœuds) formant une grille uniforme. Cette grille partitionne l'image en quadrangles.

La grille est construite sur l'image au temps  $t$  (figure 2.2.3.a). Le déplacement des nœuds est effectué sur l'image au temps  $t+1$  afin de diminuer l'erreur de compensation de mouvement (figure 2.2.3.b).

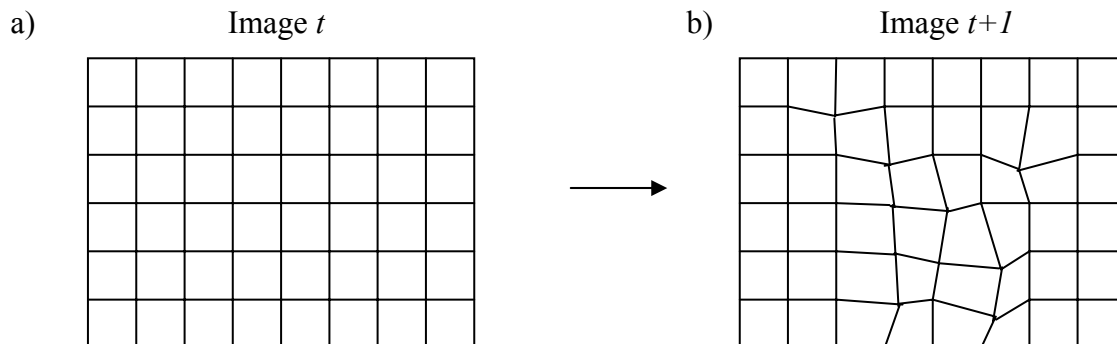


Figure 2.2.3 : exemple de grille active (quadrangles)

Le mouvement de chaque pixel est déterminé par interpolation du mouvement des nœuds des points de contrôle autour du pixel considéré. Chaque région possède alors son modèle de mouvement, déterminé par le mouvement des points de contrôle. Plusieurs modèles d'interpolation (reflétant le modèle du mouvement) peuvent être envisagés :

- Le modèle bilinéaire :

$$p(x, y) \xrightarrow{r_i} p'(x', y') \begin{cases} x' = a_1x + a_2y + a_3xy + a_4 \\ y' = a_5x + a_6y + a_7xy + a_8 \end{cases} \quad \text{III-9}$$

- Le modèle perspective :

$$p(x, y) \xrightarrow{\tau_i} p'(x', y') / \begin{cases} x' = \frac{a_1 x + a_2 y + a_3}{a_7 x + a_8 y + 1} \\ y' = \frac{a_4 x + a_5 y + a_6}{a_7 x + a_8 y + 1} \end{cases} \quad \text{III-10}$$

Les grilles actives présentent l'avantage de représenter une gamme de mouvements plus riche que l'algorithme de mise en correspondance de blocs, qui ne gère correctement que les translations.

Néanmoins, le mouvement estimé est continu, ce qui pose problème lorsque des discontinuités de mouvement apparaissent dans une séquence (ce qui est le cas dans pratiquement toutes les séquences naturelles). Pour résoudre ce problème, Nosratinia [NOSR01] propose un modèle de mouvement adaptatif en fonction de la scène, optimisant la compensation de mouvement par un noyau d'interpolation du mouvement des nœuds variable avec la séquence. Cependant, cette méthode d'optimisation est effectuée sur la globalité de l'image et ne tient donc pas compte du fait que les discontinuités apparaissent localement dans l'image. Dans un souci de compression, une autre méthode hybridant les modèles continus et discontinus en fonction du contenu est proposée dans [ALTU98].

### b) Maillages triangulaires

La méthode de mise en correspondance par maillages triangulaire [BRUS90], [NAKA91] est similaire à l'approche CGI, à la différence près que les régions formées sont triangulaires au lieu d'être rectangulaires (figure 2.2.4).

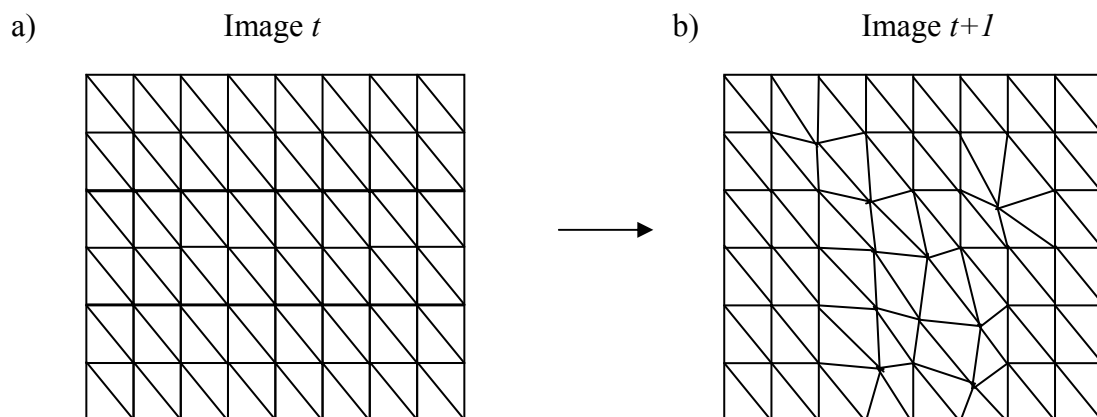


Figure 2.2.4 : mise en correspondance par maillage triangulaire

Le modèle d'interpolation à partir du mouvement des nœuds devient alors un modèle affine à 6 paramètres :

$$p(x, y) \xrightarrow{\tau_i} p'(x', y') / \begin{cases} x' = a_1 x + a_2 y + a_3 \\ y' = a_4 x + a_5 y + a_6 \end{cases} \quad \text{III-11}$$

c) Maillages adaptés au contenu ou « maillages actifs »

Les deux méthodes citées précédemment (maillages triangulaires et quadrangulaires) possèdent l'avantage de représenter un mouvement riche, mais leur structure est rigide 'sur l'image au temps  $t$ ). Ainsi, l'idée d'utiliser un *maillage actif* [WANG94b] est apparue, afin d'adapter le maillage à l'image sur laquelle il est plaqué, ceci pour deux raisons :

- Le mouvement au temps  $t$  d'une séquence donnée n'est pas uniformément réparti, certaines zones sont plus riches que d'autres. Il est donc naturel de vouloir construire un maillage dont la précision (c'est-à-dire la taille des mailles) varie avec le mouvement d'une zone donnée. La figure 2.2.5 illustre cet argument : l'image représentée figure 2.2.5.a possède deux objets différents : le cube et le fond. L'utilisation d'un maillage régulier (figure 2.2.5.b) impose des calculs supplémentaires inutiles dans la zone du fond. L'utilisation d'un maillage irrégulier (figure 2.2.5.c) donnera des résultats comparables, pour une gestion plus aisée, car le maillage ainsi construit possède moins de cellules que le maillage régulier.
- Pour une plus grande précision dans la compensation du mouvement et donc dans la segmentation des différents objets présents dans la scène, la géométrie du maillage ne doit plus être fixe i.e. il faut adapter la localisation des nœuds dans l'image pour faire correspondre les arêtes du maillage avec les frontières entre les objets présents dans la séquence, comme le montre la figure 2.2.6 : l'image au temps  $t$  est la même que celle de la figure 2.2.5.a, à la différence près que le rectangle grisé a été déplacé d'un petit nombre de pixels. Le maillage à géométrie rigide (figure 2.2.6.b) n'est ici plus adapté à l'image, car ses arêtes ne coïncident plus avec les arêtes du rectangle. Par contre, le maillage à géométrie adaptative (figure 2.2.6.c) épouse les contours du rectangle.

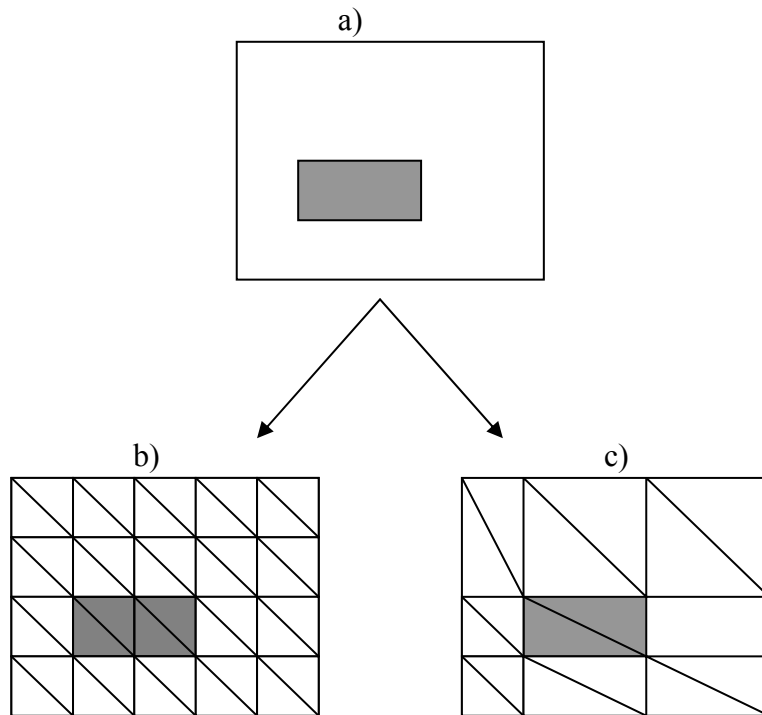


Figure 2.2.5 : connectivité régulière/connectivité irrégulière

- a) image au temps  $t$
- b) maillage à connectivité régulière
- c) maillage à connectivité irrégulière

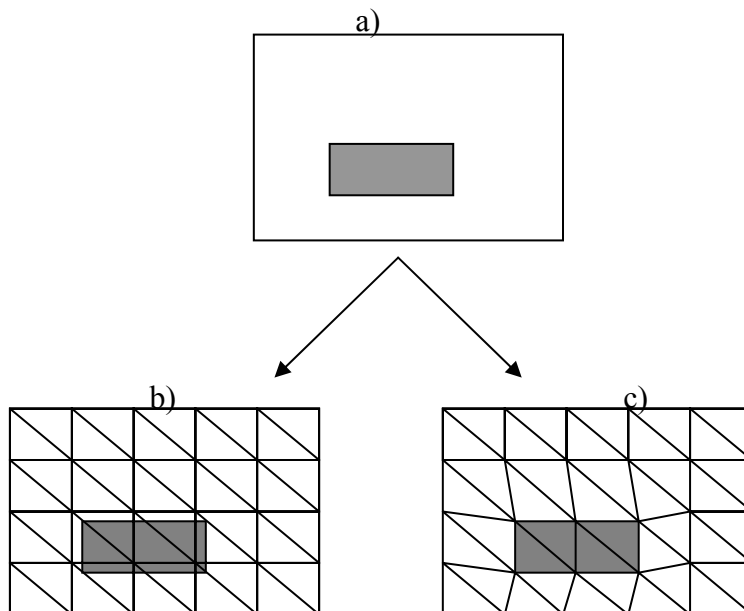


Figure 2.2.6 : géométrie rigide / géométrie adaptative

- a) image au temps  $t$
- b) maillage à géométrie rigide
- c) maillage à géométrie adaptative

Deux approches pour créer un maillage adapté au contenu de l'image ont été développées :

- L'approche basée mailles : Le maillage est formé par découpages (ou subdivisions) successifs des mailles d'un maillage initial régulier et grossier. *Huang et al.* [HUAN94] proposent l'algorithme hiérarchique d'interpolation par grille qui construit une grille par décomposition quadarbre (« quadtree » en anglais) (figure 2.2.7). Une méthode similaire est proposée dans [LAUR98], avec un maillage hiérarchique triangulaire (figure 2.2.8) cette dernière méthode insère une étape d'adaptation du maillage entre les étapes de découpage. Cette adaptation rend le maillage conforme en découpant en 2 ou en 3 les cellules voisines des cellules venant d'être subdivisées. La notion de découpage implique la définition d'un critère de subdivision, qui doit indiquer pour une cellule donnée si son découpage est judicieux ou non. *Huang et al.* proposent d'utiliser la variance temporelle comme critère de découpage. *Wang et al.* [WANG96a], [WANG96b] proposent une approche par maillage quadarbre déformable: entre chaque étape de raffinement, le maillage est déformé pour faire coïncider ses arêtes avec les contours de l'image (figure 2.2.8). La déformation est effectuée par minimisation d'une énergie constitué de trois termes : un terme d'énergie interne (modélisé par un système de masses - ressorts), un terme de mouvement (la DFD) et un terme spatial, tendant à déplacer les arêtes du maillage sur les zones à fort gradient de l'image. Le critère de découpage proposé est basé sur l'erreur de compensation, qui semble plus approprié que la variance temporelle. Ces méthodes hiérarchiques peuvent engendrer une estimation du mouvement par une approche grossier à fin, offrant une robustesse accrue.
- L'approche basée nœuds [ALTU96]: Le maillage est construit par insertions successive de nœuds dans l'image, sur les zones à fort gradient spatio-temporel de l'image. Chaque nœud inséré rend impossible l'insertion d'un autre nœud dans un disque centré sur le nœud ajouté. Le rayon du disque dépend de l'activité temporelle de la zone considérée (figure 2.2.9) : la somme cumulée de la DFD sur la surface de chaque disque est la même. Un maillage est ensuite construit par triangulation contrainte de Delaunay des nœuds ainsi créés.

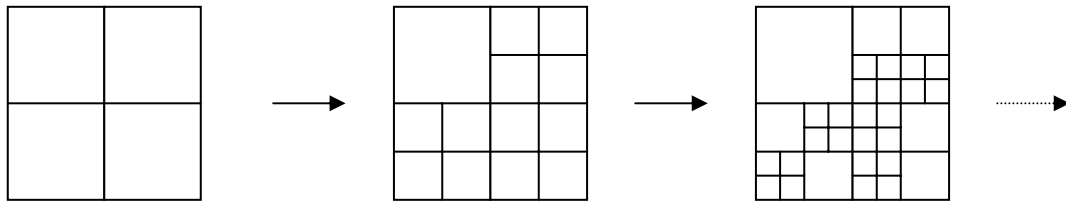


Figure 2.2.7 : approche « quadtree »

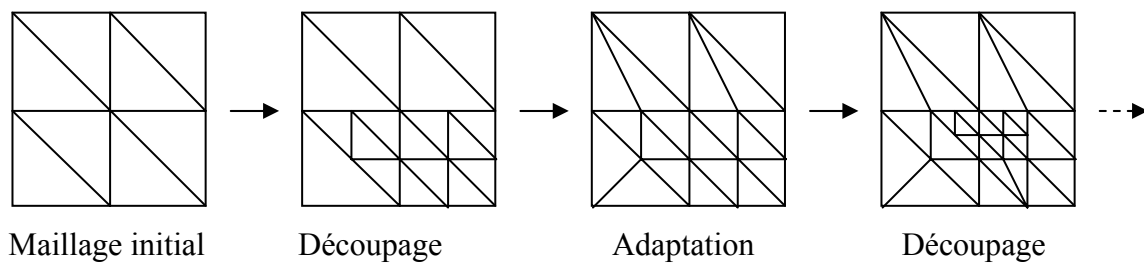


Figure 2.2.8 : approche « quadtree » par maillage triangulaire

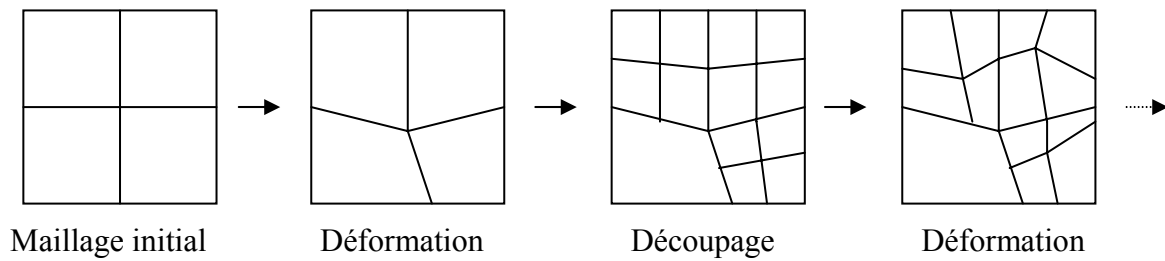


Figure 2.2.9 : approche « quadtree » + déformation

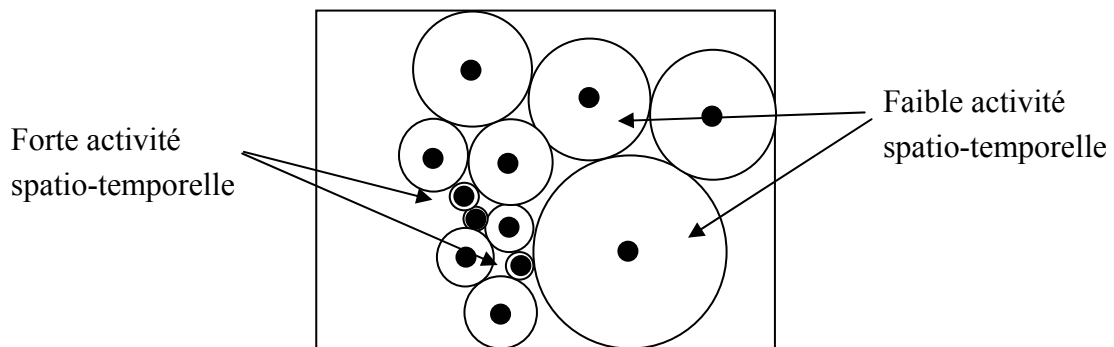


Figure 2.2.10 : approche par insertion de nœuds

d) Maillage « objet »

Plusieurs travaux utilisent la notion de maillage « objet » pour effectuer une poursuite ou un codage des objets vidéo. Un maillage objet se distingue d'un maillage classique par le fait qu'il ne recouvre pas entièrement l'image, mais uniquement l'objet vidéo auquel il est associé (voir figure 2.2.11).

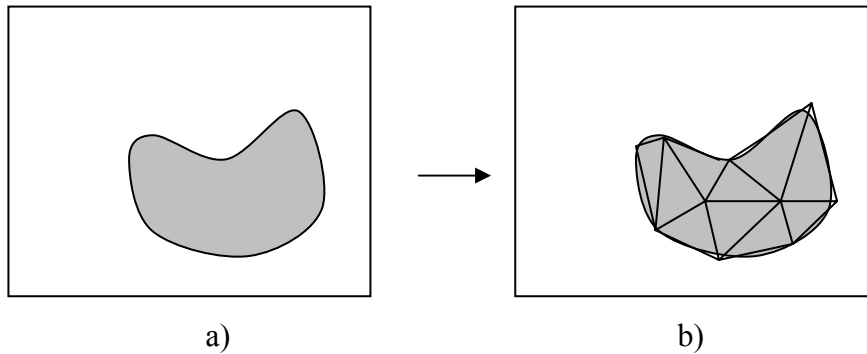


Figure 2.2.11 : exemple de maillage objet

- a) Image contenant deux objets vidéo : le fond et la forme en mouvement
- b) Maillage construit sur l'objet en mouvement

Une scène complète contiendra donc autant de maillages que d'objets vidéo. Une telle représentation permet une approximation des objets vidéo par un contour polygonal dans un but de compression [VANBE99], [LEBU98], ou pour effectuer le suivi de l'objet vidéo considéré [TOKL98][LECH99].

Cette approche est apparue car l'estimation de mouvement par maillage n'est réellement efficace que pour les mouvements continus. Or dans une séquence vidéo, des discontinuités apparaissent entre les différents objets en mouvement. L'utilisation de maillages objets permet donc une estimation robuste du mouvement à l'intérieur de chaque objet.

Mais une telle approche possède l'inconvénient de ne pas estimer réellement les discontinuités de mouvement, ce qui peut poser des problèmes aux frontières entre les objets (lorsque la forme de l'objet change, par exemple, dans le cas d'auto-occultations). De plus, du fait des différents déplacements des objets dans une scène, des régions apparaissent au cours du temps (régions découvertes) et des régions disparaissent (régions occultées), comme l'illustre la figure 2.2.12. Ainsi des trous et des collisions peuvent apparaître entre les différents maillages objets d'une même scène. Toklu [TOKL98] propose de résoudre le problème en détectant à chaque instant les régions découvertes et occultées, et adapte chaque maillage objets en fonction de ces régions.

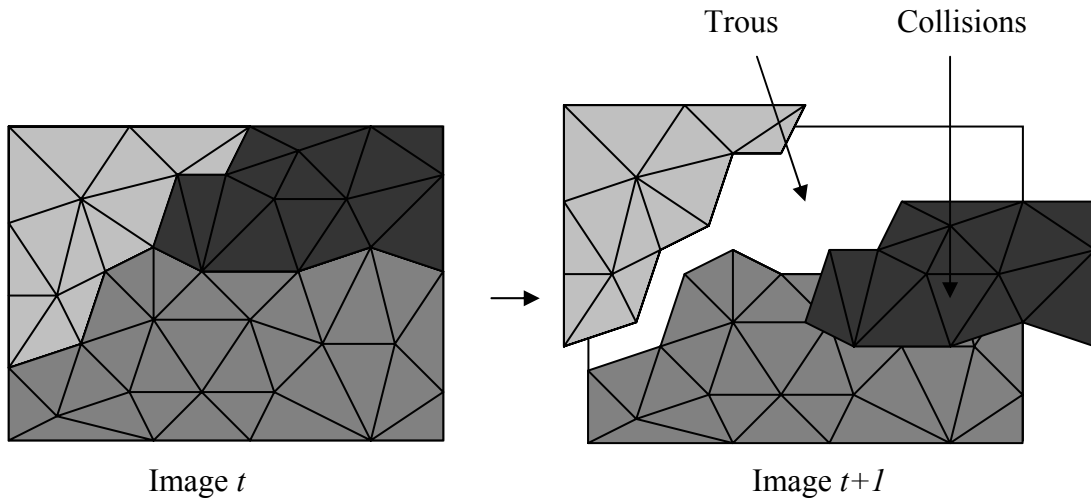


Figure 2.2.12 : Trous et collisions engendrés par les mouvements différents des objets (représentés par des maillages individuels) présents dans la scène.

e) Approche par « lignes de rupture »

Marquant [MARQ00b] invente la notion de « ligne de rupture », afin de faciliter l'estimation de mouvement dans les zones discontinues. En partant du principe que les discontinuités de mouvement apparaissent le plus souvent aux frontières entre les objets en mouvement dans l'image, l'estimation de mouvement s'effectue indépendamment sur chaque objet, dont le masque de segmentation est fourni. Un maillage hiérarchique est créé sur le masque de segmentation, et le mouvement de l'objet est estimé par déformation du maillage dans l'image au temps  $t+1$  (ou  $t-1$ ). Le maillage englobe complètement l'objet, mais seuls les pixels à l'intérieur du masque seront pris en compte pour l'erreur de compensation de mouvement. Les frontières du masque de segmentation représentent alors les lignes de rupture du mouvement. Le mouvement estimé est donc continu à l'intérieur des objets et continu aux frontières entre les objets. Ce schéma d'estimation est illustré par la figure 2.2.11, où l'image possède deux objets distincts : le fond et la forme grisée.

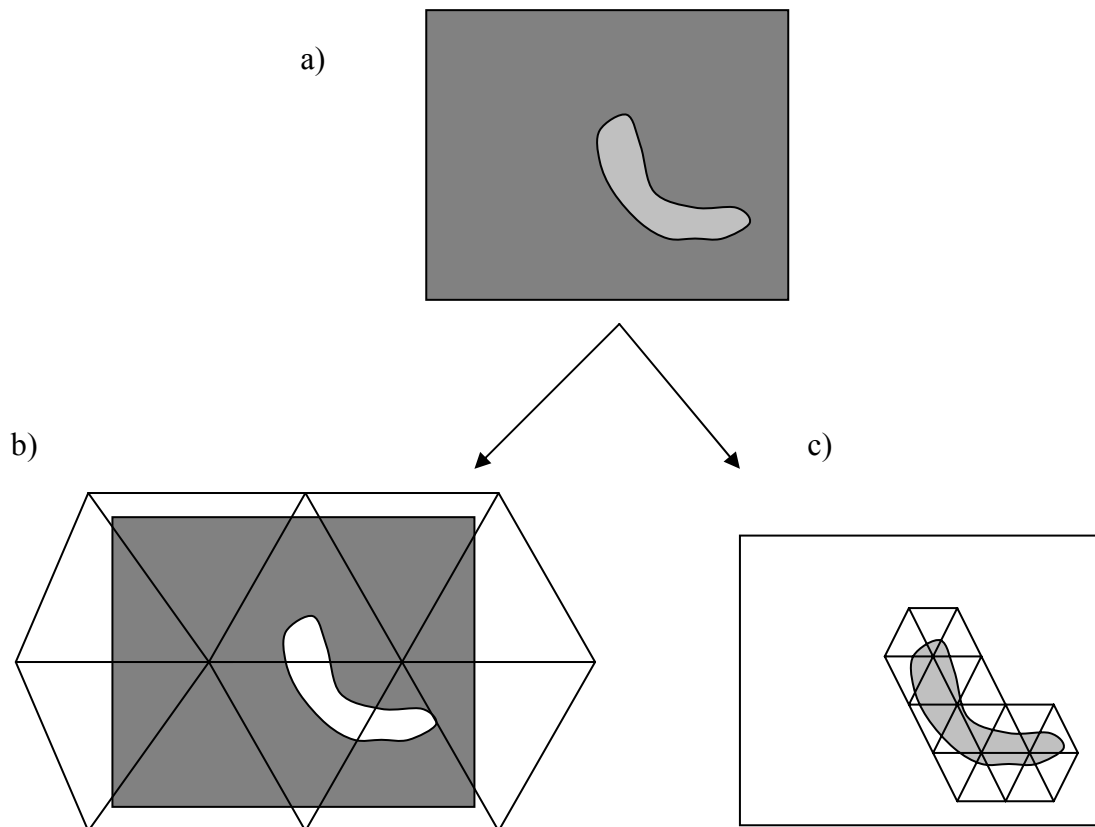


Figure 2.2.11 : lignes de rupture : l'image est constituée de deux objets : le fond et la forme grise (a). Un maillage est créé pour chaque objet (b et c). Les frontières des deux masques représentent les lignes de rupture où les discontinuités de mouvement sont autorisées.

f) Méthodes numériques d'estimation du mouvement par maillage

L'estimation des déplacements des nœuds peut s'effectuer selon plusieurs schémas :

- par relaxation déterministe : les déplacements des nœuds sont optimisés itérativement, de manière à faire diminuer l'erreur de compensation [NAKA91], [HUAN94], [TOKL98], [VANB99], [LEFL97]. Dans le cas des maillages triangulaires, le terme de mise en correspondance hexagonale (« hexagonal matching » en anglais) est employé : pour chaque nœud, on estime le déplacement minimisant l'erreur de compensation dans la région entourant le nœud (cette zone est hexagonale dans le cas d'un maillage triangulaire régulier, voir figure 2.2.12)
- par approche différentielle (similaire à la méthode du flux optique) : l'estimation de mouvement est effectuée par la recherche des déplacements des nœuds annulant les dérivées partielles de l'erreur de compensation de mouvement [BRUS90], [ALTU97b], [LECH99], [WANG94], [MARQ00a]. Cette approche est plus rapide que la méthode de relaxation déterministe, mais présente le défaut d'être sous optimale, car les calculs sont basés sur la linéarisation du critère d'erreur de compensation de mouvement, qui n'est pas toujours valable si la contrainte de conservation d'intensité (équation III-3) n'est pas respectée.
- par approche markovienne : Chrétien-Planat [CHRE99] propose une estimation du mouvement basée sur un maillage quadtree déformable. Le mouvement estimé à l'intérieur de chaque cellule est une translation. Un terme de régularisation force le champ estimé à être continu, à l'exception de certaines zones de l'image où la contrainte de régularité est relâchée, ceci permettant une estimation efficace des discontinuités de mouvement. L'estimation des différents paramètres est effectuée par recuit simulé [KIRK83].

Pour améliorer la qualité de l'estimation de mouvement, des stratégies multirésolution ou multiéchelle ont été employées. De nombreux travaux proposent une représentation hiérarchique du maillage d'estimation : l'estimation du mouvement s'effectue tout d'abord sur le maillage le plus grossier. Les vecteurs estimés sont ensuite propagés sur les niveaux de hiérarchie inférieurs où l'estimation de mouvement est raffinée. La structure hiérarchique proposée dans [TOKL98] est une structure régulière (figure 2.2.13), alors que les approches décrites dans [LECH99], [VANB99] et [CELA00] sont basées sur une pyramide de Delaunay, créée par simplifications successives du maillage (un exemple de construction d'une telle hiérarchie est donnée figure 2.2.14). [WANG96b], [MARQ00c] et [CHRE99] proposent de plus la création d'une pyramide d'images pour effectuer une estimation multirésolution, ceci pour accélérer les calculs et améliorer l'estimation des grands déplacements.

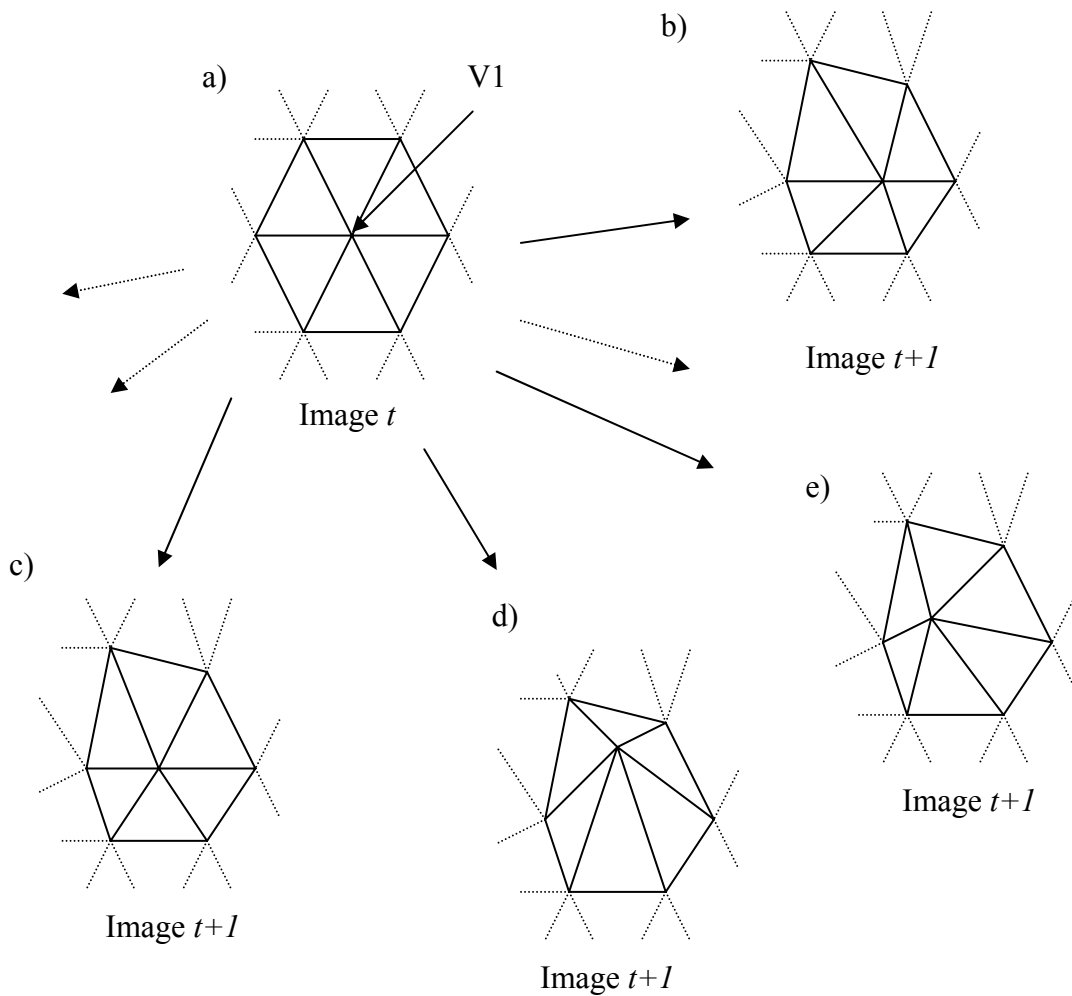


Figure 2.2.12 : algorithme de mise en correspondance hexagonale : optimisation du déplacement d'un nœud donné

Le nœud dont le déplacement est à optimiser est le nœud V1 (figure a) sur l'image au temps  $t$ . On cherche le déplacement optimal du nœud V1 en ne changeant pas les déplacements des nœuds l'entourant (quelques cas sont montrés figures b, c, d et e). Le déplacement optimal sera celui donnant l'erreur de compensation la plus petite possible sur la zone hexagonale entourant V1. Le déplacement optimal est affecté au nœud V1, et l'optimisation s'effectue alors sur un autre nœud. L'algorithme s'arrête lorsque l'erreur de compensation ne peut plus être diminuée, lorsqu'elle devient inférieure à un seuil donné, ou lorsque le nombre de passes d'optimisations sur les nœuds du maillage est égal à une valeur maximale donnée.

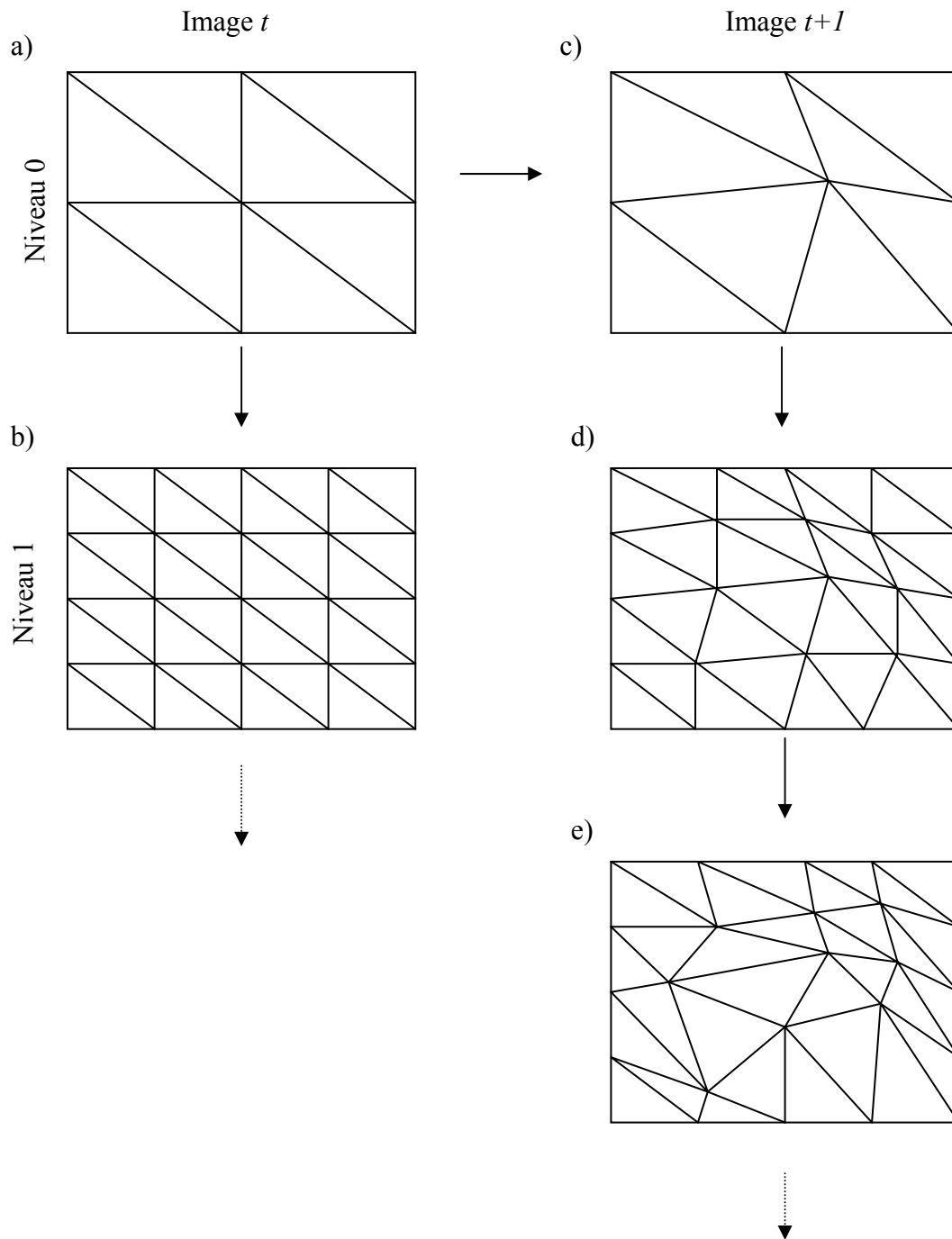


Figure 2.2.13 : estimation de mouvement hiérarchique :

On dispose d'un maillage à plusieurs niveaux de hiérarchie (a) et (b). L'estimation de mouvement entre l'image  $t$  et  $t+1$  est effectuée avec le maillage à la plus basse résolution. L'estimation est ensuite propagée sur le niveau supérieur (c). Les vecteurs mouvement des nouveaux nœuds sont initialisés par interpolation des vecteurs des nœuds de hiérarchie supérieure (d). Les vecteurs déplacement sont ensuite optimisés par une nouvelle estimation de mouvement. L'opération peut ensuite se poursuivre sur les éventuels niveaux de hiérarchie inférieure.

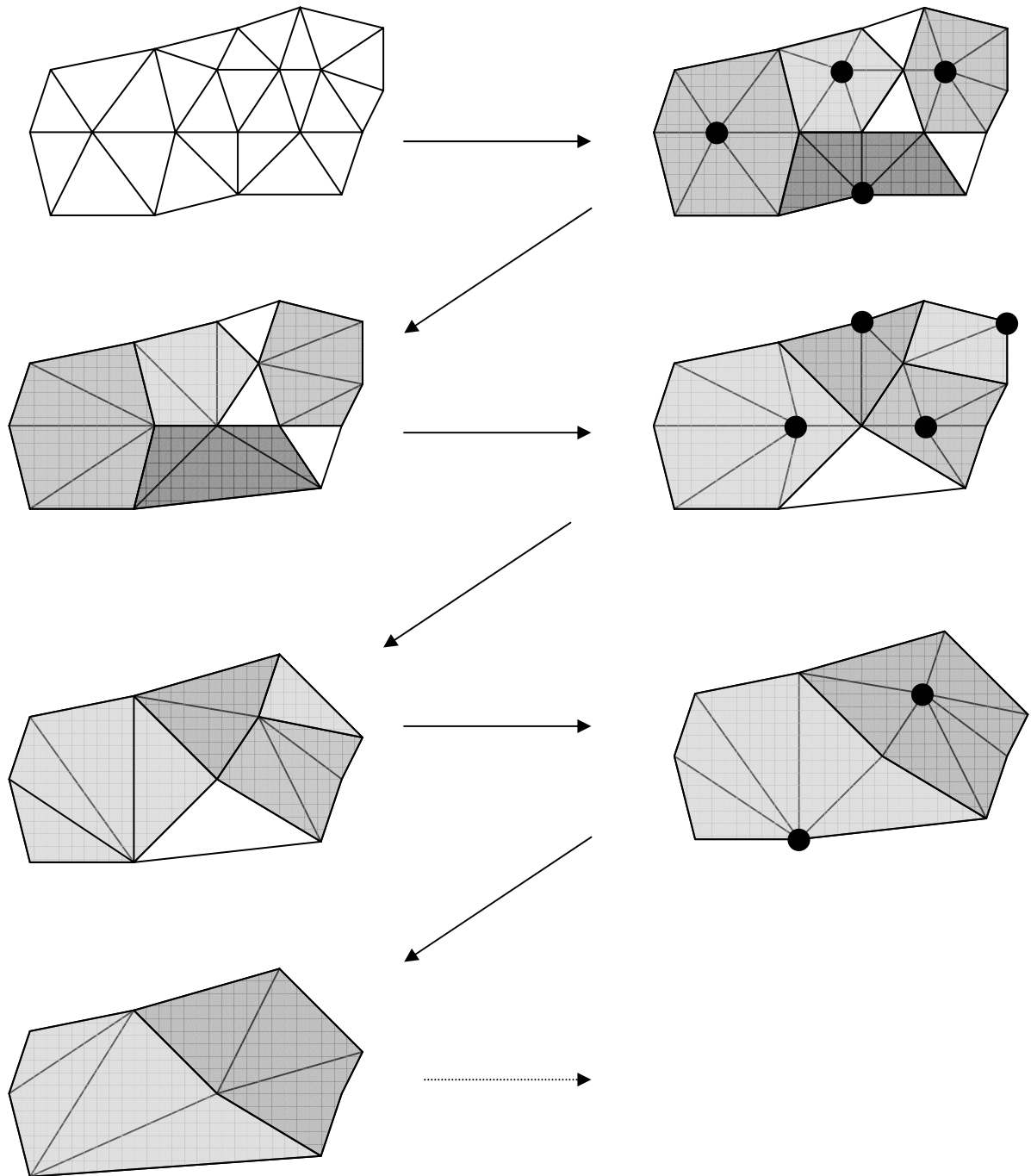


Figure 2.2.14 : création d'une hiérarchie de maillages par suppression de nœuds :  
On sélectionne un ensemble de nœuds dans le maillage, qui seront supprimés durant la simplification. Les lieux de suppression des nœuds sont ensuite remaillés par triangulation contrainte de Delaunay. L'estimation de mouvement peut ensuite s'effectuer hiérarchiquement comme le montre la figure 2.2.12.

g) Problèmes liés à l'estimation de mouvement

L'estimation de mouvement est basée sur la contrainte de conservation de l'intensité (équation III-3), et la qualité d'une estimation dépend principalement du type d'image à analyser. En effet, le mouvement est très difficile à estimer dans les zones spatialement homogènes, car dans ces zones, l'erreur de compensation reste toujours faible, que le mouvement soit bien ou mal estimé. La figure 2.2.15 illustre cette difficulté: un objet texturé de forme complexe effectue une translation sur un fond blanc uniforme (figures 2.2.15.a et 2.2.15.b). En effectuant une estimation de mouvement par mise en correspondance de grands blocs, on arrive à avoir une erreur de compensation nulle sur toute l'image, en affectant des vecteurs déplacements à tous les blocs ayant une partie en commun avec l'objet en mouvement, et en affectant un mouvement nul aux autres (figure 2.2.15.c). Pourtant, les vecteurs mouvements estimés sont erronés dans certaines parties des blocs contenant l'objet (régions grisées figure 2.2.15.d). Ainsi, si les bords des régions servant à l'estimation de mouvement ne coïncident pas avec les frontières de mouvement, le mouvement estimé sera erroné, et une segmentation selon le mouvement ne sera pas précise. Ceci est une autre illustration du problème d'ouverture, qui est une difficulté pour l'estimation de mouvement, mais un avantage pour la compression vidéo, où l'erreur de compensation minimale est recherchée.

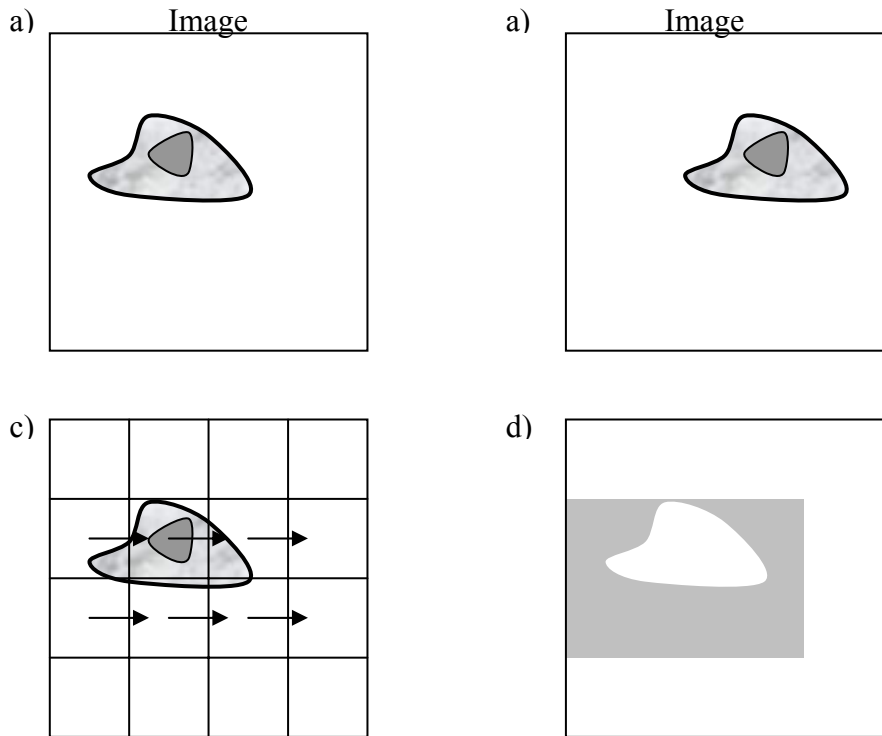


Figure 2.2.15 : problème d'ouverture

La Figure 2.2.16 montre un exemple réel d'estimation par mise en correspondance de blocs, sur la séquence J7, où la caméra suit un camion. On observe que, dans les zones texturées, le mouvement est correctement estimé mais, dans les zones homogènes, comme le ciel ou la route, les vecteurs mouvement suivent une distribution aléatoire, due au bruit présent dans l'image.

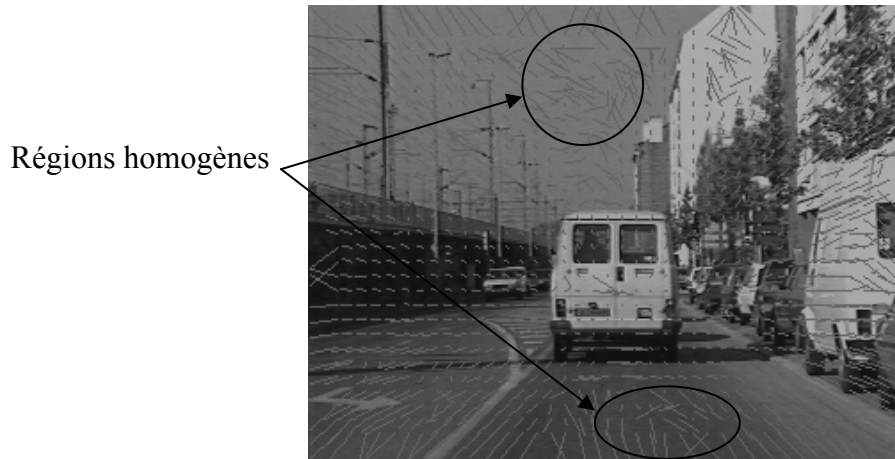


Figure 2.2.16 : exemple de mise en correspondance de blocs sur la séquence «J7 »

Les discontinuités de mouvement posent aussi des problèmes aux algorithmes de mise en correspondance, du fait des zones découvertes et des zones occultées dans l'image, selon que l'on effectue une estimation avant ou arrière, comme le montre la figure 2.2.17. Dans une estimation avant, les zones occultées sont mal estimées. En effet, une zone occultée est présente sur l'image au temps  $t$ , mais disparaît au temps  $t+1$ . L'erreur de compensation sera donc importante dans cette zone (il est impossible de reconstruire la zone occultée sur l'image  $t$  à partir de l'image  $t+1$ , étant donné qu'elle n'y figure plus. D'une manière analogue, les zones découvertes posent problème à l'estimation arrière.

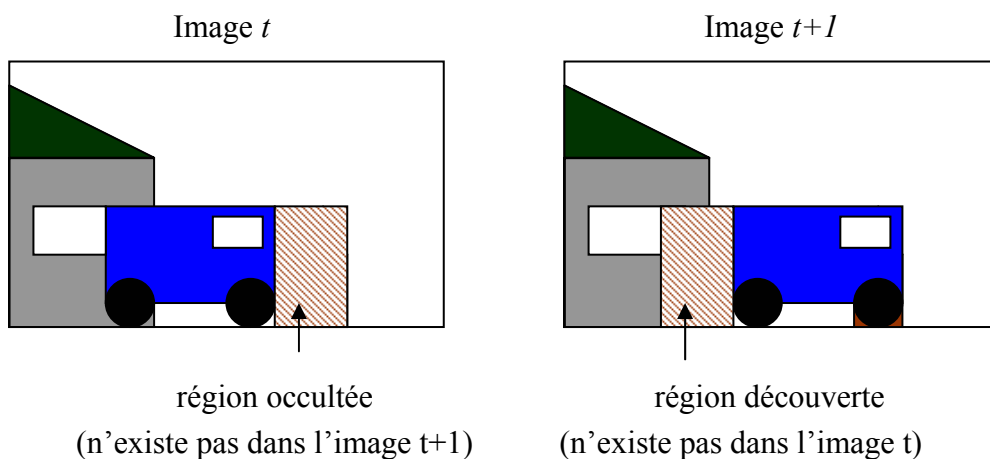


Figure 2.2.17 : régions occultées/régions découvertes

## 2.2.5 ANALYSE ET SEGMENTATION DE SEQUENCES VIDEO

### a) Segmentation par le mouvement

De nombreuses approches ont été présentées sur la segmentation de séquences vidéo par analyse du mouvement.

Wang et al. [WANG94a] utilisent une représentation de l'image par couches (« layered image representation » en anglais) pour segmenter l'image: Un champ dense de mouvement est estimé entre les images aux instants  $t$  et  $t-1$ . L'image est ensuite partitionnée en régions homogènes au sens du mouvement. Le modèle de mouvement utilisé est un modèle affine. Les régions sont itérativement découpées, fusionnées et raffinées, afin d'améliorer la qualité de la classification [BORS97], [MOST99]. Zhong et al. [ZHON01] apportent une amélioration à la méthode en utilisant plus de deux images pour établir la classification en introduisant la notion de consistance temporelle du mouvement.

Cette approche permet une classification automatique des régions dans l'image, mais ne permet pas la segmentation automatique des objets articulés, qui sont alors découpés en plusieurs régions de mouvements différents.

L'analyse du mouvement peut aussi s'effectuer sur un empilement d'images, constituant ainsi un volume, comme décrit dans [PORI01]. Des volumes homogènes au sens de la couleur ou du niveau de gris sont extraits et fusionnés selon leur trajectoire estimée.

### b) Segmentation selon un modèle prédéfini

La segmentation dans le cas général est une tâche extrêmement difficile. Elle est rendue plus aisée lorsque l'on connaît la nature des images à segmenter, et lorsque l'on connaît à priori les objets à segmenter. Cette connaissance permet d'introduire une modélisation des objets dans l'algorithme, ce qui offrira, pour une application donnée une segmentation supérieure à celle obtenue avec un algorithme générique.

Un algorithme de segmentation de la cavité du ventricule gauche du cœur humain dans les séquences IRM est proposé dans [VINC00], et son extension à la segmentation du cœur dans les images IRM 3D est décrit dans [PHAM01]. Ces algorithmes utilisent une modélisation du cœur humain par éléments finis.

c) Segmentation par détection de changement

La détection de changement utilise les interactions entre le fond et les objets vidéo présents dans une séquence, par une soustraction d'images. Cette technique revient au calcul du gradient temporel de l'image. La figure 2.2.18 illustre cette approche : en effectuant une différence (FD pour « frame difference ») entre les images 100 (figure 2.2.18.a) et 101 (2.2.18.b) de la séquence van, on obtient une image donnant une bonne indication des objets en mouvement (figure 2.2.18.d). Si l'on dispose d'une image représentant le fond sans les objets (figure 2.2.18.c), la différence entre le fond et une image de la séquence donne une indication encore plus précise (figure 2.2.18.e) : dans ce cas, les piétons sur la droite de l'image ayant un mouvement faible sont aussi détectés. La segmentation des régions en mouvement s'effectue alors dans l'image de différence, en cherchant les régions où la différence d'images est non nulle. Ceci peut être effectué par différentes méthodes :

- par seuillage [LATZ01]
- en approchant les contours par des B-Splines [PREC01] ou des contours actifs géodésiques [MANS99], [JEHA00].
- En créant un maillage actif, dont les nœuds sont placés dans les zones où la différence d'images est importante [ALTU97a].

Une amélioration de la détection de changement est proposée dans [KUHN01] : le gradient temporel est calculé à partir de l'empilement de toutes les images de la séquence, pour rendre l'estimation du gradient temporel plus robuste car basée sur plus de deux images.

La différence d'images consécutives est bon indicateur de mouvement, mais est restreinte aux séquences d'images où le mouvement de la caméra est nul. Jehan et al. [JEHA01] proposent de compenser le mouvement de la caméra par un modèle global de mouvement affine, pour s'affranchir de cette restriction.

La différence entre le fond fixe et l'image est, elle aussi contrainte, aux scènes à fond immobile. Elle présente de mauvais résultats sur des séquences de longue durée, lorsque l'aspect du fond change, à cause des changements de luminosité, par exemple. Afin de limiter ces effets, un algorithme de mise à jour de l'image du fond au cours du temps est proposé dans [CAPL01].

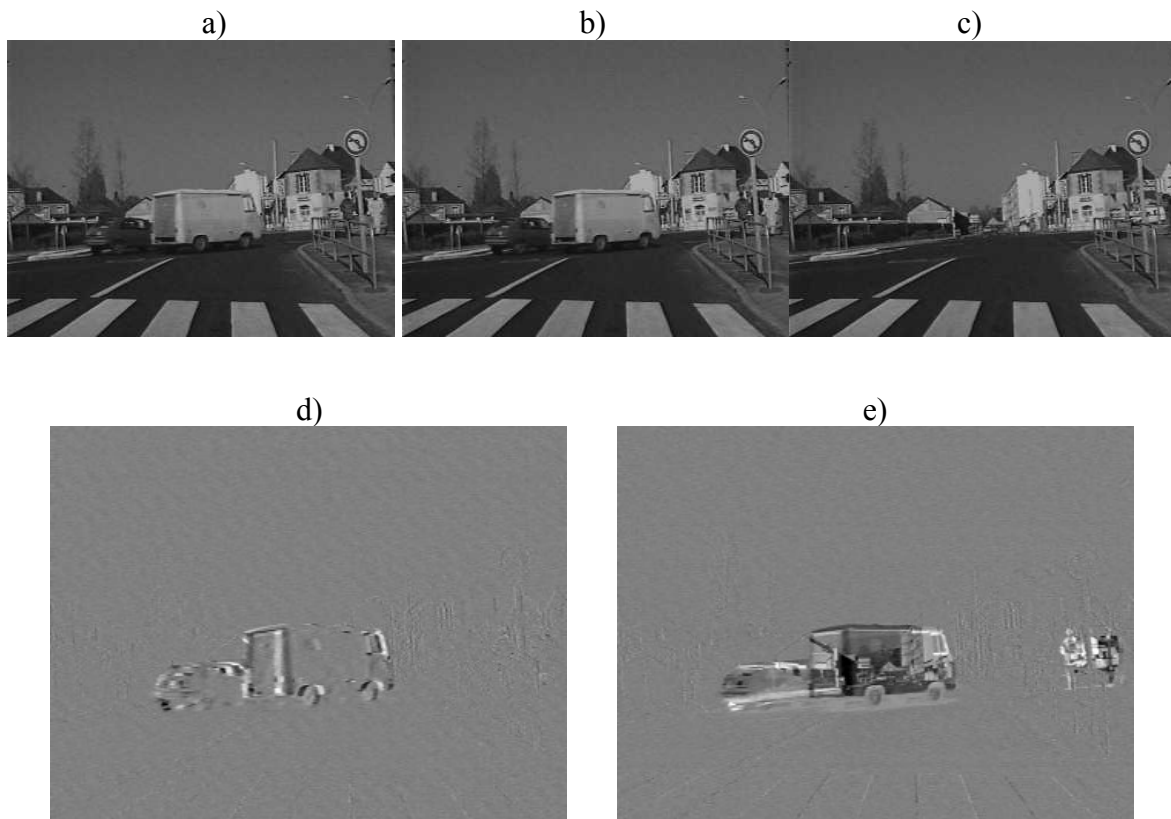


Figure 2.2.18 : Détection de changement

- a) Image du fond fixe
- b) Image 100
- c) Image 101
- d) Différence Image 100-Image 101
- e) Différence Fond-Image 100

d) Suivi d'objets vidéo

On parle de suivi d'objets lorsqu'une segmentation de référence est fournie à l'algorithme, qui n'a alors plus qu'à effectuer un suivi de ces objets le long de la séquence. Plusieurs méthodes de suivi sont possibles :

Dans [GU98], [PATE00], une carte de segmentation existe sur la première image. L'image  $t$  est segmentée en régions homogènes au sens de la luminance (ou de la couleur). Une estimation de mouvement est effectuée sur chaque région entre l'instant  $t$  et l'instant  $t-1$ . chaque région est projetée sur la segmentation à  $t-1$  en estimant le mouvement arrière. Cette projection va permettre d'associer chaque région à un objet vidéo, comme l'illustre la figure 2.2.19. Une gestion des régions découvertes doit être effectuée, car l'estimation de mouvement arrière de ces régions n'est pas possible (voir section 2.2.4.g).

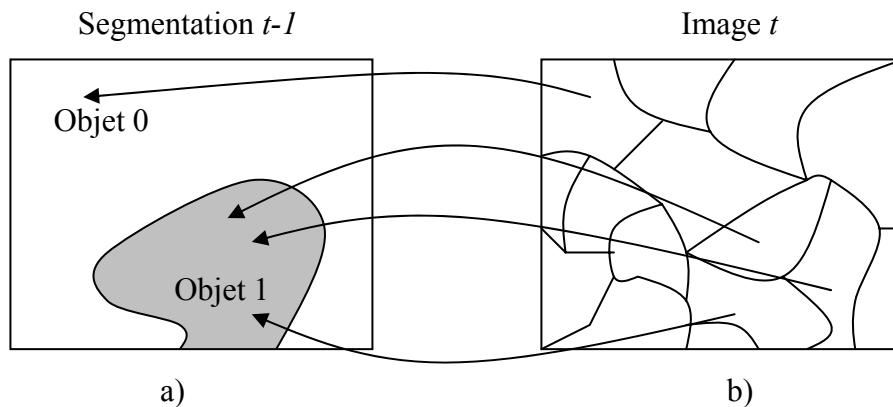


Figure 2.2.18 : suivi par projection arrière : l'image au temps  $t$  est segmentée spatialement en régions homogènes. Les régions sont ensuite projetées sur la segmentation vidéo au temps  $t-1$  et associées à l'objet qu'elles recouvrent le plus.

Le suivi des objets vidéo par maillages objets est proposé dans [LECH99] et [VANB99]. Un mouvement continu est estimé à l'intérieur du maillage, et les nœuds sont déplacés d'une image à l'autre selon le mouvement estimé.

Un algorithme de suivi par contour actif est proposé dans [FU00] : le contour est placé manuellement sur la première image. Le contour est ensuite projeté sur l'image suivante, à l'aide d'un estimateur de mouvement utilisant une estimation de type arrière et avant, pour s'affranchir des problèmes liés aux régions occultées et découvertes. Il subit enfin une déformation basée gradient, afin de le placer sur les frontières des objets qui sont aussi les frontières spatiales de l'image.

Dans [YOSH01], l'algorithme de la ligne de partage des eaux 2D (« watershed algorithm » en anglais) est étendu au volume 3D formé par l'empilement des images de la séquence à segmenter. L'utilisateur fournit un masque de segmentation dans une des images, et les différentes régions obtenues par l'algorithme de partage des eaux sont fusionnées selon la carte de segmentation.

Dans [LOUT01], un algorithme de suivi de régions d'intérêt fixée par l'utilisateur est proposée. La robustesse de l'algorithme vis-à-vis des occultations partielles ou totales de la région d'intérêt est assurée par une estimation sur laquelle on applique un filtrage de Kalman.

### 3 NOTRE PROPOSITION

Le but de ce chapitre est de fournir une méthode de segmentation et/ou de suivi d'objets vidéo, en utilisant un *maillage actif*. Les sections suivantes détaillent les choix que nous avons effectués pour construire l'algorithme proposé.

#### 3.1 Stratégie de construction

Deux solutions s'offrent à nous pour la construction du maillage : l'approche basée mailles ou l'approche basée nœuds (voir section III-2.2.4-c). Dans un souci de qualité de segmentation, l'idée de pouvoir optimiser successivement les nœuds du maillage pour l'adapter au contenu de l'image semble plus intéressante que l'option de placer les nœuds dans les zones à haute activité temporelle et de créer ensuite le maillage par triangulation de Delaunay. En effet, dans ce dernier cas, le positionnement des nœuds est effectué sans connaître la topologie du maillage (sa connectivité), et les arêtes ainsi créées ne seraient peut-être pas situées sur les frontières des objets.

Ainsi, l'optimisation des positions des nœuds semble être une étape difficilement évitable. De ce fait, nous utilisons une stratégie hiérarchique descendante, où la connectivité du maillage est fixée à l'avance, et où la position des nœuds est optimisée : on opère une déformation intra image.

Dans un souci de précision, on doit pouvoir raffiner le maillage dans les régions ayant besoin de plus de précision. Les cellules du maillage qui seront considérées comme trop grossières pour la zone à représenter seront donc subdivisées.

Il faut ensuite choisir la nature des cellules du maillage : triangulaire ou quadrangulaire. Les maillages hiérarchiques existent dans les deux versions, mais les maillages hiérarchiques à base de quadrangles sont plus difficilement déformables. En effet, des contraintes de colinéarité sont imposées entre les nœuds créés et les nœuds déjà existant avant de la subdivision, afin de garder la forme quadrangulaire des cellules. Ceci a pour conséquence qu'un nœud peut être déplacé au moment de sa création, mais une fois que des nouveaux nœuds sont créés dans un niveau de hiérarchie inférieur, le nœud ne peut plus être déplacé, comme l'illustre la figure 3.1.1. Wang [WANG96b] introduit des « nœuds fictifs » à l'intérieur des cellules de hiérarchie supérieure pour s'affranchir de ce problème. Les maillages hiérarchiques triangulaires ne possèdent pas cet inconvénient, si entre chaque étape de subdivision on effectue une adaptation du maillage, ce qui est beaucoup plus simple à implémenter que des nœuds fictifs (figure 2.2.8). Nous avons donc choisi d'employer un maillage hiérarchique triangulaire.

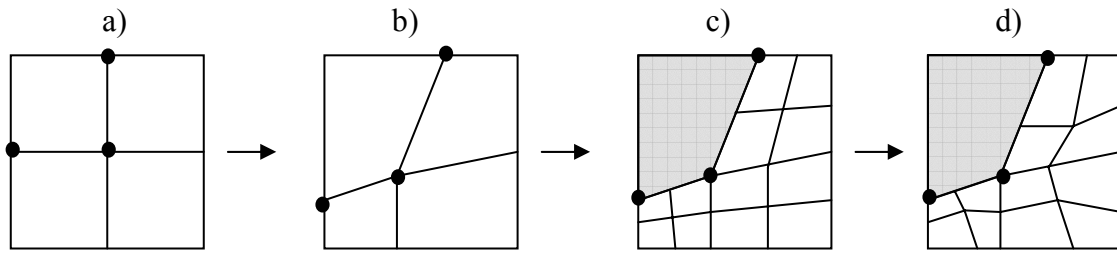


Figure 3.1.1 : maillage hiérarchique à base de quadrangles :

Les nœuds du maillage de résolution la plus basse (a) sont déplacés. Une fois la déformation effectuée, des cellules sont subdivisées et des nouveaux nœuds sont créés (c). Les nœuds marqués en noir sont alors figés, car la zone grisée ne serait plus quadrangulaire s'ils sont déplacés.

Le maillage sera donc créé à partir d'un maillage régulier, sur lequel on effectue une succession de déformations et de subdivisions sélectives, pour améliorer sa précision.

La construction d'un maillage différent sur chaque image de la séquence semble inutile, car nous opérons dans un schéma d'optimisation :

- nous utilisons le maillage créé à l'image  $t-1$  comme maillage initial sur l'image  $t$ .
- Les nouvelles coordonnées du maillage sont prédites par déplacement des nœuds au temps  $t-1$  en fonction du mouvement.

Cette technique présente l'avantage d'accélérer les calculs.

Pour que la forme des cellules ne dégénère pas au fur et à mesure des déformations successives, une stratégie de remaillage des zones du maillage trop déformées doit être utilisée, car un maillage dégénéré ferait diminuer l'efficacité de la méthode. Ainsi, après chaque passe d'optimisation (déformation), l'aspect général des cellules du maillage est observé, et les cellules considérées comme dégénérées seront remaillées.

La figure 3.1.2 montre le schéma global de fonctionnement que nous avons choisi pour notre algorithme. Les sections suivantes décrivent les blocs fonctionnels présents dans le diagramme de la figure 3.1.2: La déformation, la subdivision, le remaillage et la prédiction du maillage pour l'image suivante.

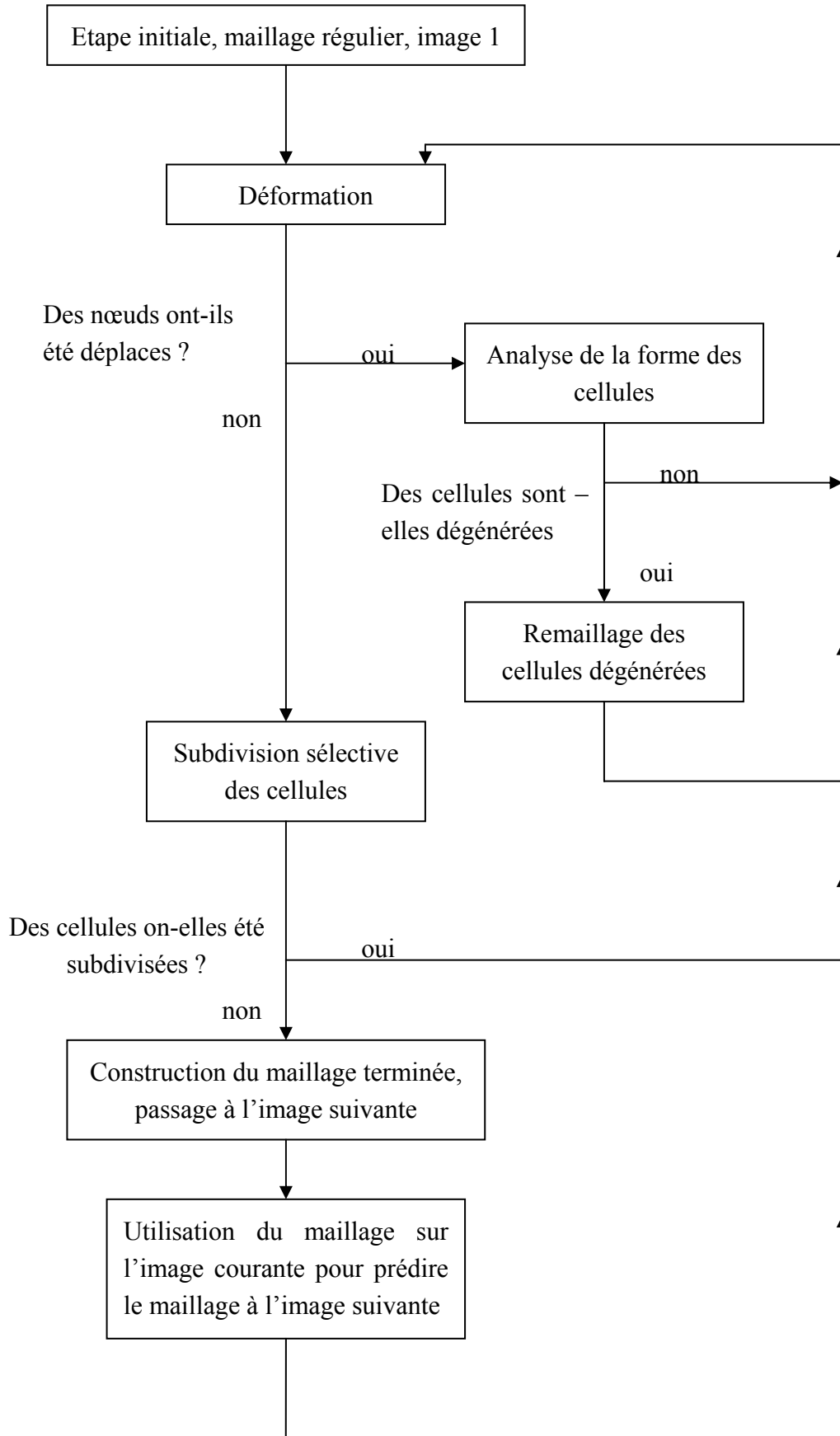


Figure 3.1.2 : Vue d'ensemble du fonctionnement l'algorithme proposé

## 3.2 Déformation du maillage

### 3.2.1 PRINCIPE

Cette section décrit la méthode employée pour déformer le maillage afin qu'il épouse le contenu de l'image c'est-à-dire pour que les arêtes du maillage coïncident avec les frontières des objets vidéo.

La notion de déformation implique la définition d'un critère d'évaluation de cette déformation. Nous utilisons un critère d'énergie, le but recherché étant de trouver la déformation du maillage qui minimise le critère. Ainsi, la définition du critère d'énergie est la tâche fondamentale dans l'algorithme proposé, puisque c'est de celui-ci que va dépendre l'efficacité de l'algorithme (figure 3.2.1).

*Wang et al.* [WANG96b] proposent un critère décomposable en trois termes :

- Un terme d'erreur de compensation de mouvement
- Un terme d'énergie visant à faire coïncider les arêtes du maillage avec les forts gradients de l'image.
- Un terme d'énergie interne, modélisé par un système de ressorts, reflétant la régularité du maillage, et empêchant une déformation exagérée, qui réduit le risque d'apparition de cellules dégénérées.

Nous avons utilisé les deux premiers termes d'énergie, car ils semblent tout à fait appropriés pour la création d'un maillage actif. Néanmoins, le troisième terme (l'énergie interne) ne nous est pas nécessaire, sachant qu'une partie de notre algorithme aura pour fonction de remailler les zones dégénérées.

Nous avons aussi souhaité régulariser l'aspect visuel de la segmentation. Deux possibilités s'offrent à nous :

- La régularisation spatiale des contours des objets (en tenant compte de la courbure des contours, par exemple)
- La régularisation temporelle de la définition des objets (la carte de segmentation)

La régularisation spatiale implique une contrainte sur les contours des objets. La qualité de la segmentation dépendra donc de la forme de l'objet considéré. Une contrainte trop forte lissera la segmentation d'objets de forme complexe, et une contrainte trop faible apporte une segmentation « accidentée ».

La régularisation temporelle implique une contrainte sur l'évolution de la forme des objets. Une contrainte trop forte peut pénaliser la segmentation d'objets non rigides ou dans les cas d'auto - occultation. Elle peut aussi engendrer un phénomène de *dérive*, qui diminue progressivement la qualité de la segmentation durant le déroulement de la séquence. Elle permet toutefois d'introduire la notion d'objet sémantique au cœur de l'algorithme : comme nous l'avons dit en introduction à ce chapitre, la segmentation sémantique d'objets vidéo purement automatique est à l'heure actuelle impossible à réaliser. Une segmentation manuelle ou semi-automatique sur la première image de la séquence (ou une autre, selon besoin) permet une définition sémantique fiable car assistée par l'utilisateur. Une régularisation temporelle permet de propager cette segmentation sur l'intégralité de la séquence.

C'est pour cela que nous avons choisi d'employer une régularisation temporelle sur la segmentation, ce qui se traduit dans notre cas par la mise en correspondance de la segmentation à l'instant  $t$  avec la segmentation à l'instant  $t-1$ .

Nous proposons donc de déformer le maillage selon le critère d'énergie  $E(t)$  défini par:

$$E(t) = E_m(t) + \alpha E_s(t) + \beta E_b(t) \quad \text{III-12}$$

où  $E_m(t)$  est le terme relatif au mouvement (inter images),  $E_s(t)$  le terme spatial (intra image) et  $E_b(t)$  est le terme de régularisation temporelle.  $\alpha$  et  $\beta$  sont des paramètres de pondération, servant à équilibrer la contribution respective des différents termes d'énergie, et sont fixés par l'utilisateur.

Les sections suivantes décrivent en détail les trois critères d'énergie employés dans notre algorithme : le terme relatif au mouvement, le terme de régularisation temporelle et le terme spatial.

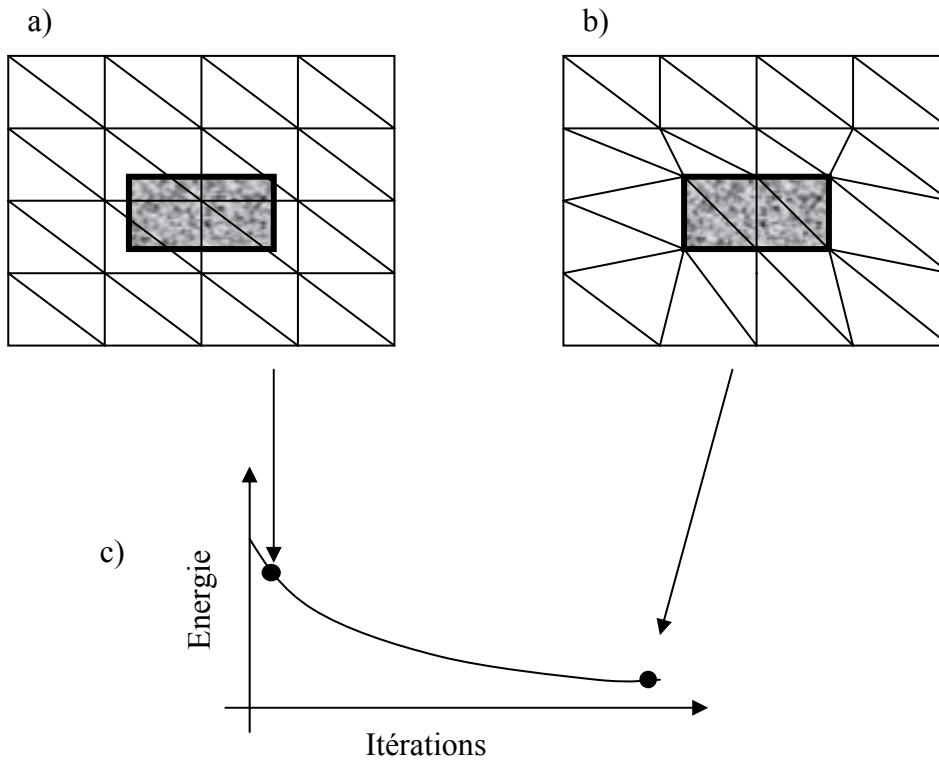


Figure 3.2.1 : Déformation par minimisation d'un terme d'énergie : le maillage initial (a) est itérativement déformé jusqu'à une configuration (b) minimisant le critère choisi (c), reflétant la qualité de la segmentation.

### 3.2.2 CRITERE DE MOUVEMENT

La définition d'un critère de mouvement permet de déformer le maillage afin que les cellules soient homogènes au sens du mouvement. Une estimation du mouvement des objets doit donc être effectuée. Comme nous l'avons expliqué précédemment, cette estimation est rendue difficile à cause du problème d'ouverture et des discontinuités de mouvement qui engendrent des zones découvertes et des zones occultées. Nous allons montrer que l'algorithme d'estimation de mouvement proposé prend en compte les discontinuités de mouvement, et présente une capacité à estimer correctement le mouvement dans les zones occultées ou découvertes, par une estimation bidirectionnelle du mouvement.

Ainsi, nous proposons d'effectuer l'estimation du mouvement dans l'image  $t$  par une mise en correspondance dans les deux directions (figure 3.2.2):

- en avant (entre les images  $t$  et  $t+1$ ). Le mouvement des régions découvertes à l'instant  $t$  est correctement estimé, mais le mouvement des zones occultées entre les instant  $t$  et  $t+1$  ne l'est pas (figure 3.2.2.a).
- en arrière (entre les images  $t$  et  $t-1$ ) : le mouvement des régions occultées entre les instants  $t$  et  $t+1$  est correctement estimé à l'instant  $t$ , mais pour les zones découvertes à l'instant  $t$  (n'existant pas à l'instant  $t-1$ ) le mouvement ne peut pas être estimé (figure 3.2.2.b).

Chaque cellule du maillage se verra donc attribuer deux mouvements (avant et arrière), dont un seul sera correct dans le cas où la cellule contient une zone découverte ou occultée. Nous proposons alors pour chaque cellule de choisir la direction de l'estimation de mouvement de chaque cellule prise en compte pour le calcul du terme d'énergie en fonction de l'erreur de compensation (figure 3.2.2.c) :

- Si le mouvement estimé en avant donne une erreur de compensation plus faible qu'en arrière, la cellule considérée contient probablement une zone découverte, on considère alors uniquement la compensation avant.
- Si la compensation arrière est meilleure que la compensation avant, la cellule est le lieu d'une occultation à l'image  $t+1$ , on choisit la méthode arrière.
- Dans le cas où la cellule ne possède aucune région occultée ou découverte, les deux directions d'estimation doivent fournir une erreur de compensation équivalente. Le choix de la direction de l'estimation est donc peu important. On choisit alors la direction donnant l'erreur de compensation la plus faible.

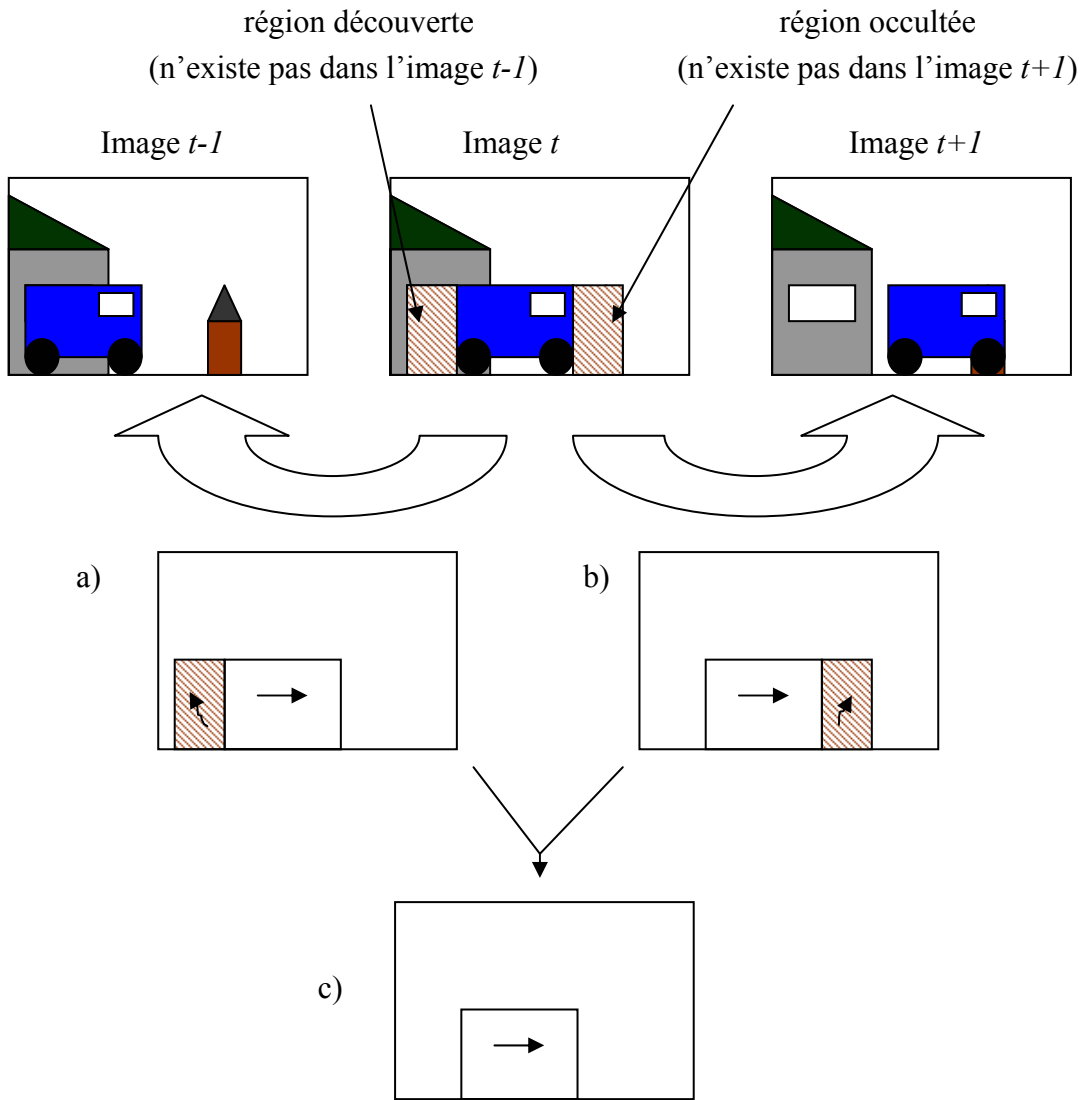


Figure 3.2.2 : estimation de mouvement combinée avant - arrière :

- a) estimation arrière (zones découvertes non prises en compte)
- b) estimation avant (zones occultées non prises en compte)
- c) combinaison des deux estimations : l'estimation de mouvement est meilleure

Nous définissons le terme d'énergie relatif au mouvement  $E_m(t)$  par :

$$E_m(t) = \frac{\sum_i \left( \sum_{p \in c_i} [I(\tau_i(p), t + d_i) - I(p, t)]^2 \right)}{\sum_i a_i} \quad \text{III-13}$$

où :

- $I(p, t)$  est l'intensité du pixel  $p$  de coordonnées  $(x, y)$  dans l'image au temps  $t$ .
- $a_i$  est l'aire de la cellule triangulaire  $c_i$ .
- $\tau_i$  est une transformation reflétant le mouvement interne à chaque cellule et peut être une transformation avant ( $\tau_i^+$ , entre les images  $t$  et  $t+1$ ) ou arrière ( $\tau_i^-$ , entre les images  $t$  et  $t-1$ ).
- $d_i$  est la direction (avant ou arrière) choisie pour le mouvement interne de la cellule considérée.

Ce terme d'énergie est l'erreur quadratique de compensation de mouvement normalisée (la DFD). Plusieurs modèles de mouvement ont été implantés dans notre algorithme. En effet, l'estimation de mouvement est le bloc fonctionnel le plus coûteux en termes de calculs. Un modèle de mouvement simple rendra l'estimation de mouvement plus rapide, mais fournira une estimation médiocre dans les scènes où le mouvement est complexe. A l'inverse, un modèle de mouvement plus évolué permettra l'analyse de scènes complexes, mais pénalisera l'algorithme au niveau de la rapidité d'exécution.

Dans la description qui suit, nous ne considérerons qu'une seule direction d'estimation, mais dans notre implémentation, l'estimation s'effectue en avant et en arrière. Les modèles de mouvement que nous avons utilisés sont :

- Un modèle de translation (2 paramètres par cellule). Un vecteur déplacement unique est estimé sur chaque cellule, indépendamment d'une cellule à l'autre. Des discontinuités peuvent donc être estimées sur toutes les arêtes du maillage (figure 3.2.3.a).
- Un modèle affine (6 paramètres par cellule) : pour chaque cellule, le mouvement est défini par les trois vecteurs déplacements estimés sur les trois nœuds formant la cellule considérée. Ici encore, les discontinuités sont autorisées sur chaque arête du maillage. Chaque nœud du maillage se verra donc attribuer autant de vecteurs déplacements qu'il n'a de cellules autour de lui (figure 3.2.3.b)

- Un modèle affine contraint (6 paramètres par cellule), où l'on interdit les discontinuités de mouvement à l'intérieur des objets segmentés. Un nœud à l'intérieur d'un objet vidéo se verra donc attribuer un seul vecteur déplacement, et un nœud situé sur une frontière entre deux objets possèdera deux vecteurs mouvements, comme l'illustre la figure 3.2.3.c.
- Un modèle de mouvement utilisant un maillage pour chaque objet, utilisant la notion de ligne de rupture [MARQU00b]. L'approche par ligne de rupture implique l'utilisation d'un masque de segmentation et d'un maillage pour chaque objet. Dans notre implémentation, nous ne disposons pas du masque de segmentation sur toutes les images (car notre but est la création de ces masques). Néanmoins, le maillage que nous cherchons à créer peut faire office de masque. La figure 3.2.4 illustre cette modification : si la scène contient trois objets (Le fond et les deux formes en mouvement, figure 3.2.4.a), nous aurons simultanément quatre maillages : le maillage que l'on cherche à déformer pour effectuer la segmentation (figure 3.2.4.b) et un maillage pour chaque objet vidéo, représentant son mouvement propre (figure 3.2.4.c, 3.2.4.d et 3.2.4.e). Les maillages « mouvement » que nous avons utilisés sont de type régulier.

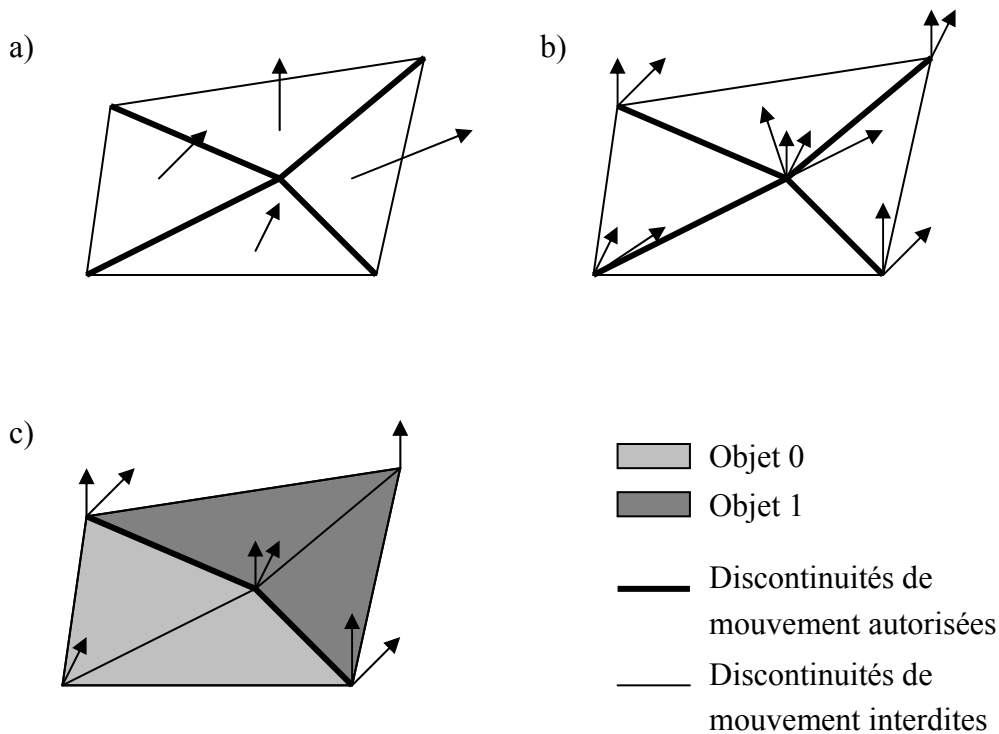


Figure 3.2.3 : modèles de mouvement : translation (a), affine (b) et affine contraint (c)

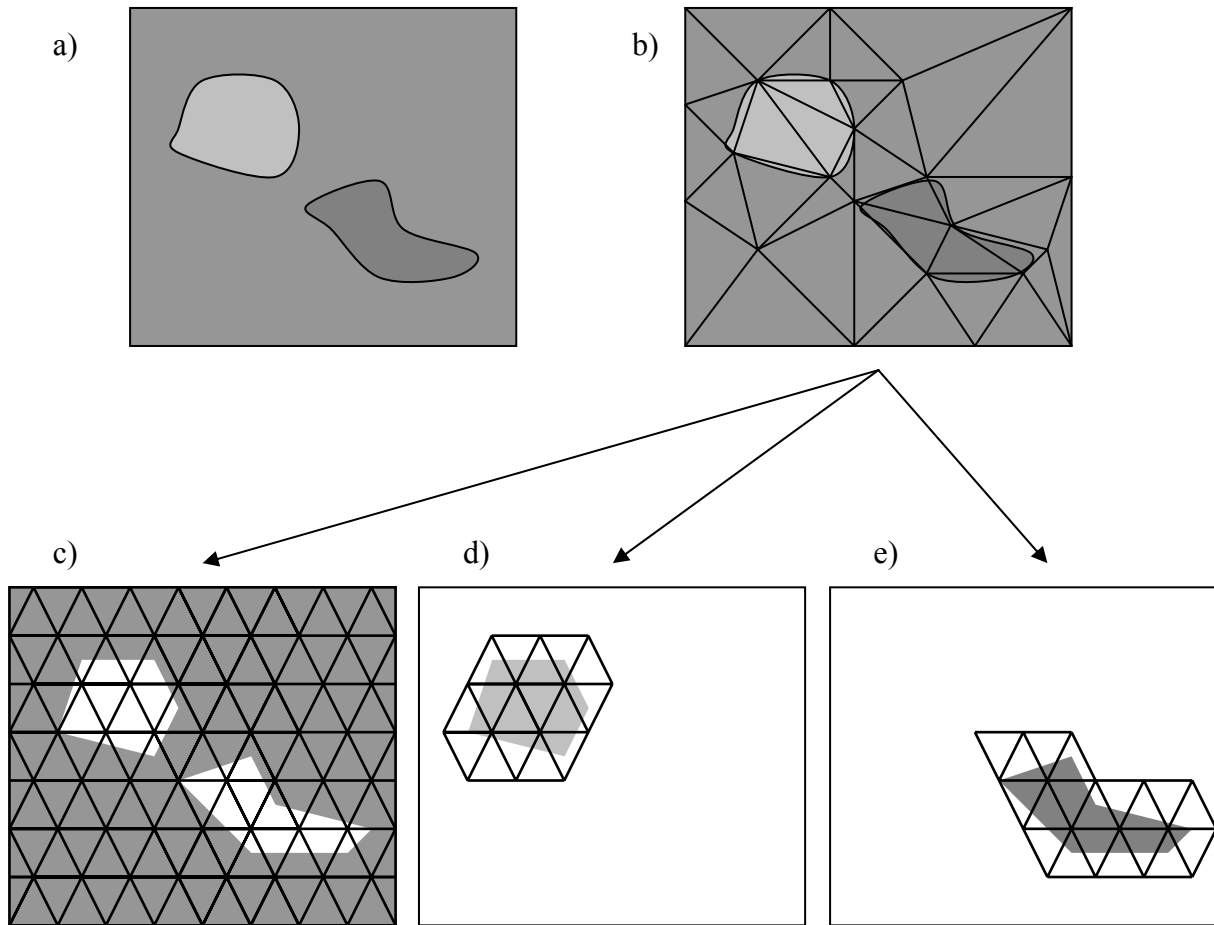


Figure 3.2.4 : adaptation de l'approche par lignes de rupture pour notre algorithme : le maillage actif (b) sert de masque pour l'estimation de mouvement avec lignes de ruptures (c, d et e)

Nous avons considéré, dans notre implémentation, la gestion des séquences en couleurs : le terme d'énergie de mouvement est alors la somme des erreurs de compensation de mouvement sur les trois composantes RVB de l'image.

Les deux derniers modèles de mouvement utilisés (le modèle affine contraint et le modèle utilisant les lignes de rupture) impliquent la localisation des objets vidéo c'est à dire que l'algorithme doit connaître pour chaque cellule du maillage à quel objet vidéo elle est associée. Cette association est effectuée au moment du calcul du terme d'énergie de régularisation temporelle de la segmentation, décrit ultérieurement.

Le terme d'énergie de mouvement implique la déformation du maillage pour que les cellules possèdent un mouvement homogène (au sens du modèle choisi), mais n'est pas affranchi du problème d'ouverture qui rend l'estimation du mouvement très difficile dans les zones homogènes spatialement, comme l'illustre la figure 3.2.5 : le maillage est construit et déformé sur l'image 100 de la séquence « van », où un camion et une voiture sont en mouvement. La

déformation est effectuée par prise en compte unique du critère de mouvement. Dans les régions texturées, les arêtes du maillage sont bien placées sur les discontinuités de mouvement, mais dans la région la plus homogène (la route), la frontière est mal placée, car dans cette région, l'erreur de compensation reste faible. Ainsi, une partie de la route est associée avec la partie basse avant du camion. Le critère spatial décrit dans la section suivante permet une meilleure localisation des contours des objets dans les zones homogènes.

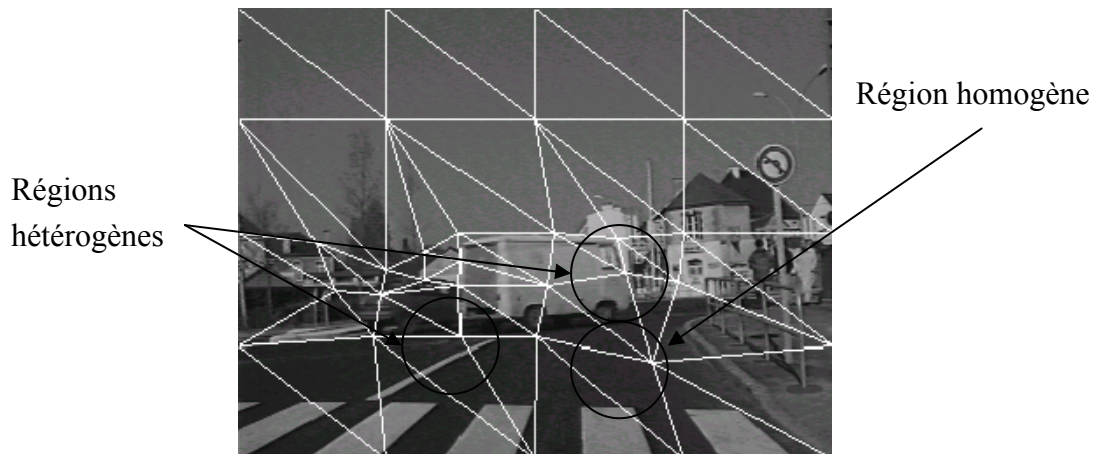


Figure 3.2.5 : maillage déformé uniquement par prise en compte d'un critère de mouvement

### 3.2.3 CRITERE SPATIAL

Nous avons introduit un critère spatial dans le terme d'énergie global, pour s'affranchir du problème d'estimation de mouvement dans les zones homogènes au sens de la luminance. La prise en compte simultanée d'un critère spatial et d'un terme de mouvement doit être effectuée de manière à ce que ces deux critères ne soient pas contradictoires. Deux approches sont envisageables :

- Définir un critère visant à placer les arêtes du maillage sur les discontinuités spatiales (où le gradient spatial de l'image est le plus grand) [WANG96b], [BENO99]. C'est une approche basée contours.
- Définir un critère visant à maximiser l'homogénéité des cellules au sens de la luminance ( un critère utilisant la variance de chaque cellule, par exemple, [LEFL97], [LECH98]). C'est une approche basée régions.

Les deux approches sont illustrées par la figure 3.2.6 : Un maillage quadrarbre est construit sur l'image « Lena » (figure 3.2.6.a) , par l'approche basée contours (en déplaçant les arêtes du maillage sur les forts gradients de l'image, figure 3.2.6.b) et par l'approche basée régions (en minimisant la variance de chaque cellule, figure 3.2.6.c). L'intérieur de chaque cellule du maillage est remplacé par la luminance moyenne des pixels de la cellule considérée. L'approche basée régions donne un rapport signal sur bruit équivalent à l'approche basée contours (pour un nombre de cellules beaucoup plus petit : 1444 contre 3659). Dans un soucis de codage d'images fixes, une approche basée régions semble donc à première vue plus efficace qu'une approche basée contours.



Figure 3.2.6 : déformation en fonction du contenu spatial de l'image :

- a) image « Lena » originale
- b) Maillage construit avec un critère basé contours, 3659 cellules, PPSNR=23.6 dB.
- c) Maillage construit avec un critère basé régions, 1444 cellules, PPSNR=23.8 dB.

Cette conclusion ne s'applique pas à la compression de séquences vidéo ou la segmentation spatio-temporelle, pour deux raisons :

- L'estimation du mouvement dans une séquence est intrinsèquement pénalisée par le problème d'ouverture, qui apparaît lorsque le support d'estimation de mouvement est une région spatialement homogène. Un critère visant à rendre les cellules homogènes au sens de la luminance va donc diminuer l'efficacité de l'estimation de mouvement.
- Un terme d'énergie basé régions implique la définition d'un critère d'hétérogénéité (que l'on cherchera à minimiser en déformant le maillage), qui aura une valeur importante dans les régions hétérogènes et une valeur faible dans les régions homogènes. Si on associe ce critère spatial à un critère de mouvement (ce qui est notre cas), il est préférable que ces deux termes d'énergie soient complémentaires. Or, dans les régions fortement hétérogènes, le critère spatial ainsi défini prendrait une valeur élevée et sa minimisation ne mènerait pas forcément au déplacement des arêtes du maillage sur les frontières de mouvement, si on considère par exemple deux régions voisines ayant des propriétés spatiales (textures) identiques, mais des mouvements différents. Ce terme d'énergie pourrait donc diminuer l'efficacité de la déformation dans les zones hétérogènes, où le terme basé mouvement est, lui, fiable.

Nous avons donc décidé d'utiliser un terme d'énergie spatiale basé contours  $E_s(t)$  défini par :

$$E_s(t) = \frac{\sum_j \left( r(e_j) \sum_{p \in e_j} dist(p) \right)}{\sum_j l_j} \quad \text{III-14}$$

où :

- $e_j$  est une arête du maillage
- $l_j$  est la longueur de l'arête  $e_j$
- $dist(p)$  est la distance du pixel  $p(x,y)$  au contour détecté le plus proche [REISS97](figures 3.2.7.b et 3.2.7.c).
- $r(e_j) = 1$  si l'arête  $e_j$  est sur une frontière entre deux objets vidéo. Sinon,  $r(e_j) = 0$ .

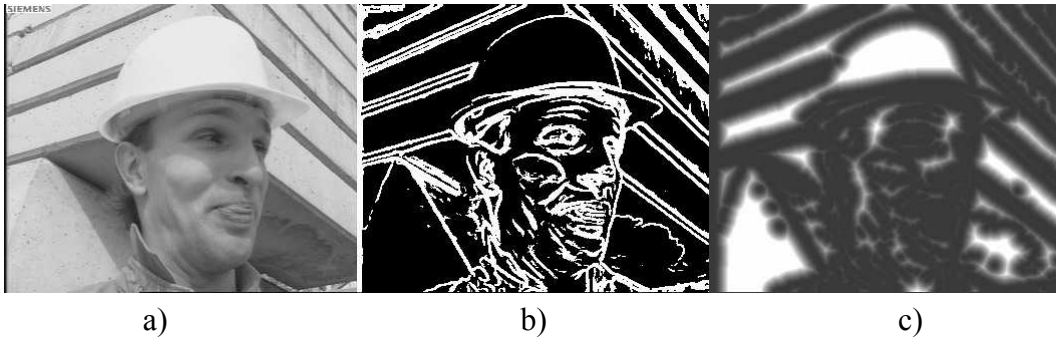


Figure 3.2.7 : construction d'une carte de distance :

- a) image originale
- b) contour détecté sur l'image
- c) carte de distance au contour le plus proche

Notre algorithme de détection de contours est élaboré ainsi (figure 3.2.7.b):

- On applique un filtrage médian sur l'image originale (la fenêtre utilisée est de dimension 5x5 pixels)
- On calcule le gradient morphologique de l'image filtrée (le motif utilisé est un motif 6-connexe)
- Les contours sont finalement détectés par seuillage de l'image du gradient morphologique (valeur du seuil utilisé : 10)

L'énergie spatiale ainsi définie est similaire à une énergie de contour actif, par l'insertion du terme  $r(e_j)$  (voir figure 3.2.8). En effet, la prise en compte des arêtes à l'intérieur des objets est inutile, car seule la frontière entre les objets est importante.

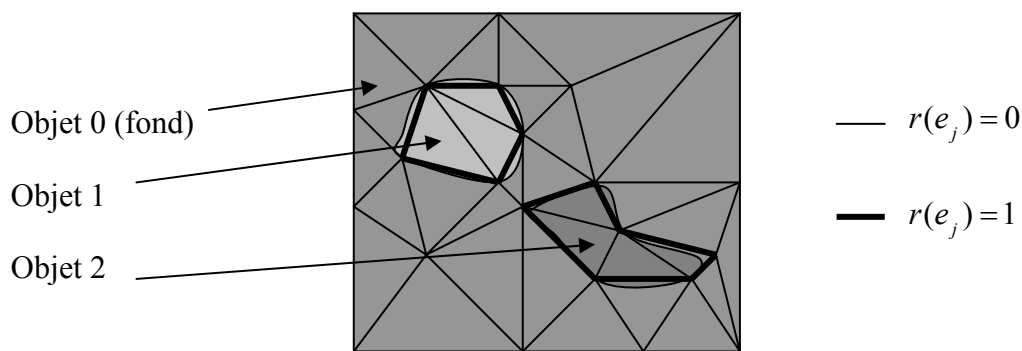


Figure 3.2.8 : détermination de  $r(e_j)$  pour chaque arête  $e_j$

Nous avons donc défini un critère spatial complémentaire au critère de mouvement, puisque si une arête est située dans une région fortement texturée, beaucoup de contours seront détectés et la distance au contour le plus proche sera donc faible quelque soit la position de l'arête. La

déformation du maillage se fera donc principalement par la minimisation du critère de mouvement. Inversement, si l'arête est située dans une zone homogène traversée par un contour, le terme d'énergie relatif au mouvement sera peu significatif par rapport au terme spatial, qui permet à notre algorithme de se placer correctement sur le contour spatial de l'image.

Le terme spatial ainsi défini nécessite la localisation des objets vidéo pour le calcul du terme  $r(e_j)$  pour chaque arête du maillage. Cette localisation est effectuée lors du calcul du terme de régularisation temporelle de la segmentation, détaillé dans la section suivante.

### 3.2.4 CRITERE DE REGULARISATION TEMPORELLE

L'introduction d'un terme de régularisation temporelle de la segmentation a deux utilités dans notre algorithme :

- Assurer une stabilité visuelle de la segmentation spatio-temporelle (c'est son utilité première)
- Permettre une segmentation sémantique des objets vidéo, c'est-à-dire fusionner les régions possédant des caractéristiques spatiales (luminance) ou temporelles (mouvement) différentes, mais appartenant au même objet vidéo, selon une définition propre au système de vision humaine. Ainsi, on pourra définir correctement les objets vidéo, leurs intérieurs et leurs frontières, tâche indispensable pour le calcul de l'énergie de mouvement et de l'énergie spatiale.

Nous déterminons le terme de régularisation temporelle de la segmentation  $E_b(t)$  par :

$$E_b(t) = \frac{\sum_i m_i}{\sum_i a_i} \quad \text{III-15}$$

où :

- $a_i$  est l'aire de la cellule  $c_i$  considérée
- $m_i$  est le nombre de pixels mal classés dans la cellule  $c_i$

Le calcul de  $m_i$  est explicité par la figure 3.2.9 : chaque cellule du maillage construit sur l'image au temps  $t$  est projetée en arrière sur la segmentation effectuée sur l'image précédente (au temps  $t-1$ ). La projection arrière utilisée est déterminée lors de l'estimation de mouvement par méthode arrière (entre les images  $t$  et  $t-1$ ). La cellule considérée est alors associée à l'objet

vidéo qu'elle recouvre le plus [GU98][PATE00] (sur la figure 3.2.9, la cellule  $c_1$  est associée à l'objet vidéo 1 et la cellule  $c_2$  est associée à l'objet vidéo 0). Le nombre de pixels mal classés  $m_i$  est égal au nombre de pixels projetés en dehors de l'objet vidéo associé à la cellule considéré (sur la figure 3.2.9,  $m_1$  est le nombre de pixels dans la zone en gris clair et  $m_2$  est le nombre de pixels dans la zone en gris foncé).

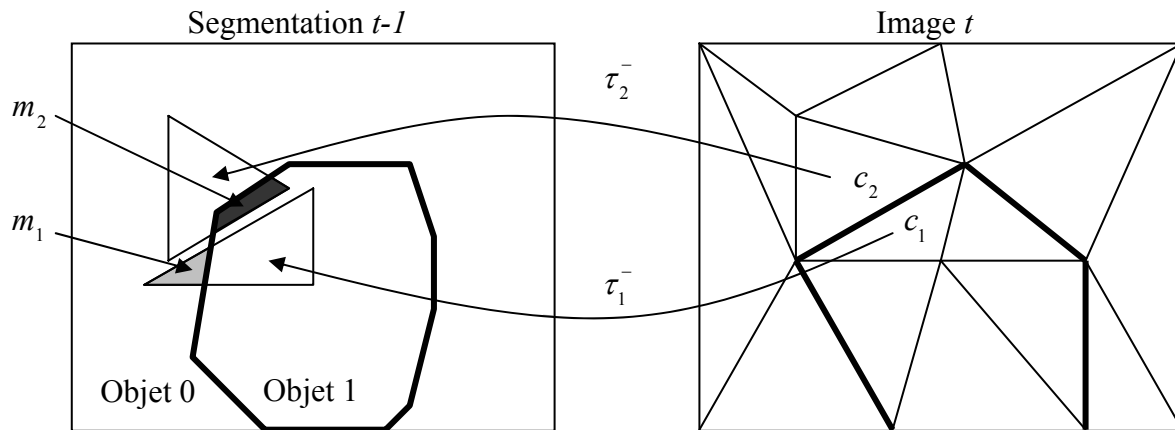


Figure 3.2.9 : calcul du nombre de pixels mal classés  $m_i$

Le terme de régularisation ainsi défini permet de garantir que la segmentation au temps  $t$  est en accord avec la segmentation au temps  $t-1$ . Il sera donc nécessaire de fournir à l'algorithme une carte de segmentation sur la première image de la séquence à traiter pour pouvoir propager cette segmentation sur les images suivantes.

Un défaut va cependant diminuer l'efficacité d'une telle régularisation : elle est basée sur la projection arrière des cellules au temps  $t$  sur la segmentation au temps  $t-1$ . L'estimation de mouvement de type arrière ne donne pas de bons résultats dans les zones découvertes (voir section 2.2.4.g). Le terme de régularisation risque donc de pénaliser l'efficacité de la déformation du maillage dans les zones découvertes de l'image entre les instants  $t-1$  et  $t$ .

### 3.2.5 MINIMISATION DU CRITERE D'ENERGIE

Finalement, le critère d'énergie global que nous avons employé dans notre algorithme est:

$$E(t) = \alpha \frac{\sum_i \left( \sum_{p \in c_i} [I(\tau_i(p), t + d_i) - I(p, t)]^2 \right)}{\sum_i a_i} + \beta \frac{\sum_j \left( r(e_j) \sum_{p \in e_j} \text{dist}(p) \right)}{\sum_j l_j} + \frac{\sum_i m_i}{\sum_i a_i} \quad \text{III-16}$$

Nous avons employé un algorithme de relaxation déterministe pour minimiser cette énergie : l'algorithme va déplacer itérativement chaque nœud du maillage (dans une fenêtre de dimensions  $(2d+1) \times (2d+1)$  centrée sur le nœud considéré) jusqu'à ce que le critère d'énergie ne puisse plus être diminué.

Pour accélérer les calculs, la déformation du maillage n'intervient que sur les nœuds placés sur les frontières entre les objets. En effet, le déplacement des nœuds intérieurs aux objets vidéo peut faire diminuer le critère d'énergie mais ne changera en rien le résultat de la segmentation. Nous avons limité le nombre d'itérations de déformation du maillage dans un souci de réduction du temps de calcul et afin d'assurer la fin des itérations lorsque l'énergie reste constante.

Le critère d'énergie contenant des termes relatif au mouvement, un algorithme d'estimation de mouvement a dû être implémenté. Après chaque déformation, l'estimation du mouvement des cellules ayant été déformées doit être recommencée, car le support d'estimation a lui aussi changé. L'algorithme doit de plus s'assurer que les cellules ne deviennent pas trop dégénérées (c'est à dire trop petites ou trop étirées) et, le cas échéant, remailler les cellules dégénérées.

L'estimation de mouvement est elle aussi réalisée par relaxation déterministe : les vecteurs mouvements des différents nœuds sont itérativement optimisés, selon une stratégie de recherche déterministe.

L'algorithme de déformation se déroule ainsi :

- Estimer le mouvement sur toutes les cellules  $c_i$  du maillage.
- Marquer tous les nœuds  $n_j(x, y)$  placés sur les frontières entre les objets.
- $Compteur=1$
- $Maillage\_déformé = 1$
- Tant que ( $Maillage\_déformé = 1$  et  $Compteur < seuil\_itérations$ ) faire
  - {
    - $Maillage\_déformé = 0$
    - Pour chaque nœud  $n_j(x, y)$  marqué faire :
      - {
        - Chercher la nouvelle position  $(x', y')$  du nœud minimisant  $E_m(t)$ , avec  $|x'-x| < d$  et  $|y'-y| < d$ .
        - Si ( $x' = x$  et  $y' = y$ ) :
          - {
            - la position  $(x, y)$  était déjà optimale : le nœud  $n_j(x, y)$  n'est plus marqué.
        - Sinon :
          - {
            - Placer le nœud  $n_j(x, y)$  sur ses nouvelles coordonnées  $(x', y')$
            - Marquer le nœud  $n_j(x, y)$  ainsi que tous ses nœuds voisins situés sur des frontières entre les objets.
            - $Maillage\_déformé=1$
- Remailler les cellules  $c_i$  dégénérées.
- Estimer le mouvement sur les cellules  $c_i$  possédant un nœud  $n_j(x, y)$  marqué.
- $Compteur = Compteur + 1$

Dans nos expérimentations, nous avons fixé  $seuil\_itérations$  20 et  $d$  à 4.

### 3.2.6 NORMALISATION ET PARAMETRES DE PONDERATION

Le critère global d'énergie à minimiser est une somme de trois critères pondérés par les coefficients  $\alpha$  et  $\beta$ .

Les trois critères d'énergie  $E_m(t)$ ,  $E_s(t)$  et  $E_b(t)$  sont normalisés (par la surface des cellules pour  $E_m(t)$  et  $E_b(t)$ , et par la longueur des arêtes pour  $E_s(t)$ , voir équation III-16) pour garantir leur invariance avec l'échelle. En effet,  $E_m(t)$  et  $E_b(t)$  sont des critères basés régions, et s'il n'étaient pas normalisés, leur grandeur évoluerait quadratiquement avec la dimension des objets, alors que le critère  $E_s(t)$  non normalisé évoluerait de manière linéaire. Ceci nécessiterait l'ajustement des paramètres  $\alpha$  et  $\beta$  pour que la contribution respective des trois énergies reste constante vis-à-vis de la taille des objets.

### 3.3 Stratégie de découpage

Dans la plupart des travaux effectués sur les maillages à subdivision non uniforme, le critère de découpage est un critère basé sur l'erreur de compensation de mouvement. Or, dans un but de suivi d'objets, la qualité de compensation de mouvement ne signifie pas forcément que les objets ont été correctement suivis, à cause du problème d'ouverture (voir section 2.2.4.g). Nous proposons donc la formulation d'un nouveau critère de découpage, basé sur la mise en correspondance de la segmentation au temps  $t$  avec la segmentation au temps  $t-1$ . Le critère d'énergie relatif à la régularisation temporelle de la segmentation implique déjà le calcul d'un tel critère, par le biais du terme  $m_i$ , qui détermine pour chaque cellule  $c_i$  le nombre de pixels mal classés par rapport à la segmentation au temps  $t-1$  (voir section 3.2.4). Nous proposons donc d'effectuer un test de découpage de chaque cellule vis-à-vis du nombre de pixels mal classés  $m_i$  : si  $m_i$  est inférieur à un seuil  $S_{\text{découp}}$  donné, la cellule reste intacte, sinon, la cellule est découpée pour améliorer la précision du maillage dans cette région.

### 3.4 Méthodes de remaillage

Le maillage subissant de nombreuses déformations durant la segmentation, un contrôle de l'aspect de ses cellules doit être effectué, pour remailler les zones dégénérées. Nous avons employé deux méthodes de remaillage pour améliorer l'aspect du maillage : la contraction d'arête (figure 3.4.a) et la permutation d'arête (figure 3.4.b)

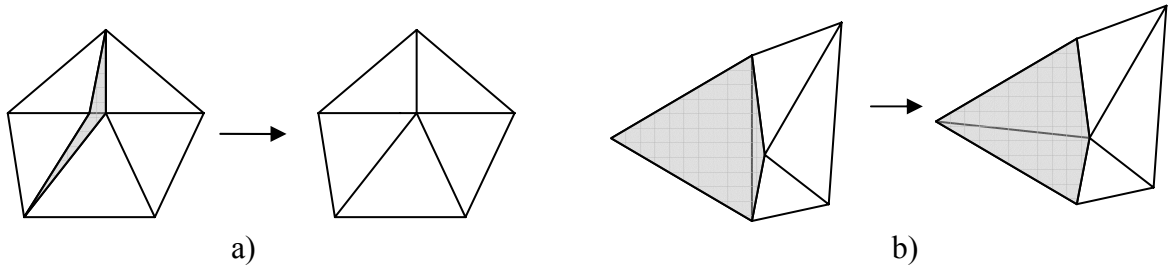


Figure 3.4.1 : cas de remaillage :  
 a) Contraction d'arête  
 b) Permutation d'arête

Pour ces deux types de remaillage, il a fallu définir un critère de déclenchement, indiquant si la zone considérée doit être remaillée.

#### 3.4.1 CRITERE DE CONTRACTION D'ARETE

Pour le calcul du critère de contraction d'arête, on considère deux cellules voisines, dont on calcule la longueur des différentes arêtes (figure 3.4.2).

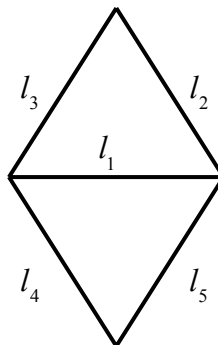


Figure 3.4.2 : longueurs calculées pour le critère de contraction d'arête

Nous définissons le critère de contraction  $C_{cont}$  :

$$C_{cont} = \frac{l_1}{l_1 + l_2 + l_3} \times \frac{l_1}{l_1 + l_4 + l_5} \quad \text{III-17}$$

Si le critère de contraction de deux cellules voisines est inférieur à un seuil  $S_{cont}$  donné et la longueur  $l_1$  de l'arête commune est inférieure à 7 pixels, leur arête commune est contractée, et les deux cellules sont supprimées. Dans nos expérimentations,  $S_{cont}$  est fixé à 0.018.

### 3.4.2 CRITERE DE PERMUTATION D'ARETE

Le calcul du critère de contraction d'arête fait lui aussi intervenir deux cellules voisines, mais les termes géométriques employés sont différents : on calcule la longueur des arêtes ainsi que la longueur des hauteurs incidentes à l'arête commune (figure 3.4.3.a).

#### *préliminaires*

Une permutation d'arête peut s'effectuer uniquement si le résultat de la permutation ne donne pas lui aussi une cellule dégénérée (comme l'illustre la figure 3.4.4), car sinon, la permutation serait de nouveau effectuée, plaçant l'algorithme dans une boucle infinie. Il faut donc aussi considérer la géométrie des cellules sur lesquelles on simule une permutation d'arête (figure 3.4.3.b). De plus, la permutation peut être effectuée uniquement si le quadrilatère formé par les deux cellules voisines est convexe (figure 3.4.5)

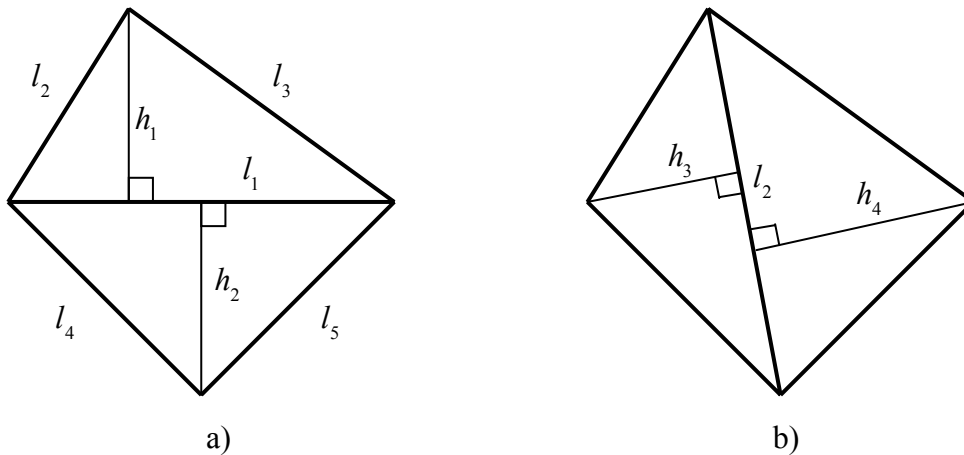


Figure 3.4.3 : longueurs calculées pour le critère de permutation d'arête

Nous définissons alors quatre rapports de longueur  $r_1$ ,  $r_2$ ,  $r_3$  et  $r_4$  :

$$r_1 = \frac{h_1}{l_1} \quad \text{III-18}$$

$$r_2 = \frac{h_2}{l_1} \quad \text{III-19}$$

$$r_3 = \frac{h_3}{l_2} \quad \text{III-20}$$

$$r_4 = \frac{h_4}{l_2} \quad \text{III-21}$$

Finalement, une permutation d'arête entre deux cellules voisines sera effectuée si :

$$(r_1 < S_{perm} \text{ ou } r_2 < S_{perm}) \text{ et } (\min(r_1, r_2) < \min(r_3, r_4))$$

Dans nos expérimentations, nous avons fixé  $S_{perm}$  à 0,2.

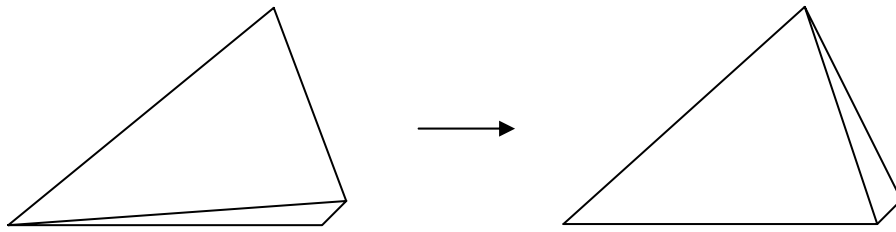


Figure 3.4.4 : cas où une permutation d'arête n'améliore pas la qualité du maillage

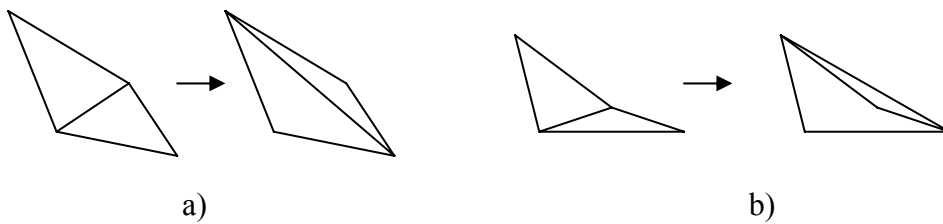


Figure 3.4.5 : permutation d'arête : convexité du polygone englobant les cellules voisines :

- a) polygone convexe : permutation autorisée
- b) polygone non convexe : permutation interdite

### 3.5 Prédiction du maillage d'une image à l'autre

Lorsque la construction du maillage est terminée sur l'image  $t$ , ce maillage peut servir de maillage initial pour l'image  $t+1$ , après une étape de prédiction. Dans cette étape, chaque nœud du maillage est déplacé selon un vecteur mouvement, qui est défini comme étant la moyenne des mouvements estimés (par l'estimation de type avant) des cellules voisines à ce sommet. Ceci permet d'une part d'effectuer une prédiction du maillage au temps  $t+1$  par le maillage au temps  $t$ , mais permet aussi de déplacer les nœuds intérieurs aux objets vidéo, ce qui améliore visuellement l'évolution du maillage au cours du temps. En effet, pendant l'étape de déformation du maillage, seuls les nœuds situés aux frontières entre les objets sont déplacés. Si l'on n'effectuait pas l'étape de prédiction, les nœuds intérieurs aux objets vidéo seraient immobiles, même si l'objet vidéo dont ils font partie est en mouvement.

## 4 RESULTATS

Cette section présente les résultats obtenus avec notre algorithme. Nous étudions tout d'abord la minimisation du critère global d'énergie. Une étude quantitative de la qualité du suivi des objets est ensuite effectuée sur une séquence synthétique. Enfin, des résultats obtenus sur des séquences réelles sont montrés. Dans les sections suivantes, nous appellerons :

- Modèle 1 : le modèle de mouvement basé sur des translations.
- Modèle 2 : le modèle de mouvement affine.
- Modèle 3 : le modèle de mouvement affine contraint.
- Modèle 4 : le modèle de mouvement utilisant les lignes de ruptures.

Le mouvement estimé dans chaque image est représenté sous forme de traits dont la couleur indique la direction de l'estimation: si les traits sont bleus, le mouvement de la cellule considérée a été estimé par la méthode avant, s'ils sont rouges, le mouvement a été estimé de façon arrière.

### 4.1 Minimisation d'énergie

La figure 4.1.1 illustre la minimisation itérative du terme d'énergie effectuée lors de la construction du maillage sur l'image 6 de la séquence « foreman ». Le masque de segmentation initial sur l'image 5 est présenté figure 4.1.3.a. L'évolution du maillage aux itérations 0, 1, 2, 4, 6, 8 et 20 est montré sur les images de la figure 4.3.b à 4.3.h. Le modèle de mouvement utilisé est le modèle 4. Les paramètres sont :  $\alpha = 1$ ,  $\beta = 200$ ,  $S_{decoup} = 150$ .

La figure 4.1.2 illustre la minimisation sur l'image suivante (image 7). Le maillage construit sur l'image 6 sert de maillage initial pour l'image 7, après prédiction. Cette initialisation permet une minimisation plus rapide : l'algorithme s'est arrêté au bout de 7 itérations, alors que sur l'image 6, il a atteint le nombre maximum d'itérations autorisées (20).

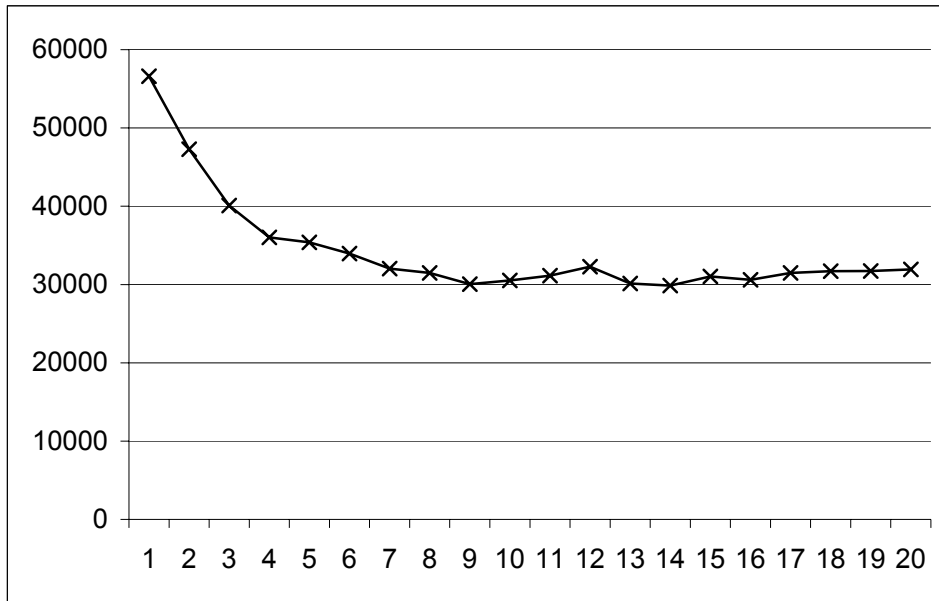


Figure 4.1.1 : évolution du critère d'énergie en fonction des itérations pour l'image 6 de la séquence « Foreman »

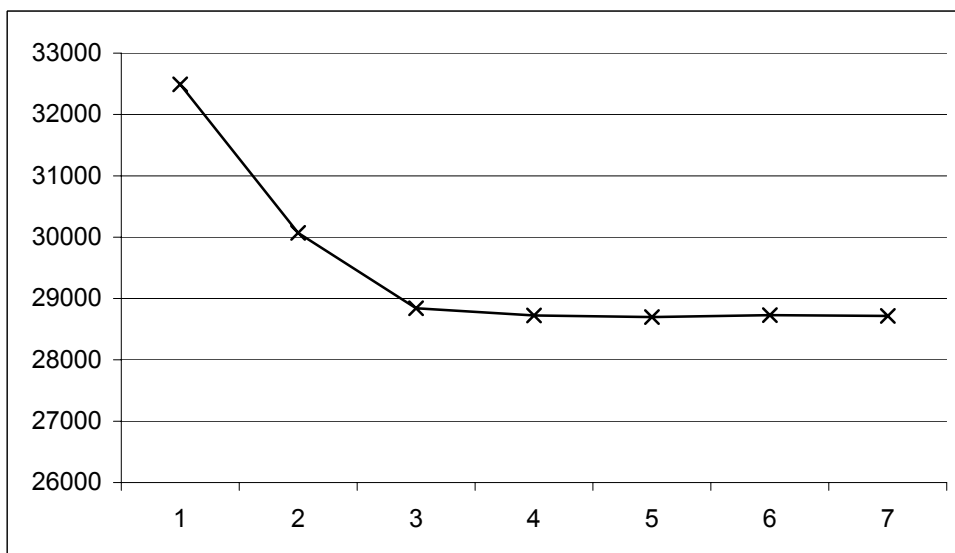


Figure 4.1.2 : évolution du critère d'énergie en fonction des itérations pour l'image 7 de la séquence « foreman »

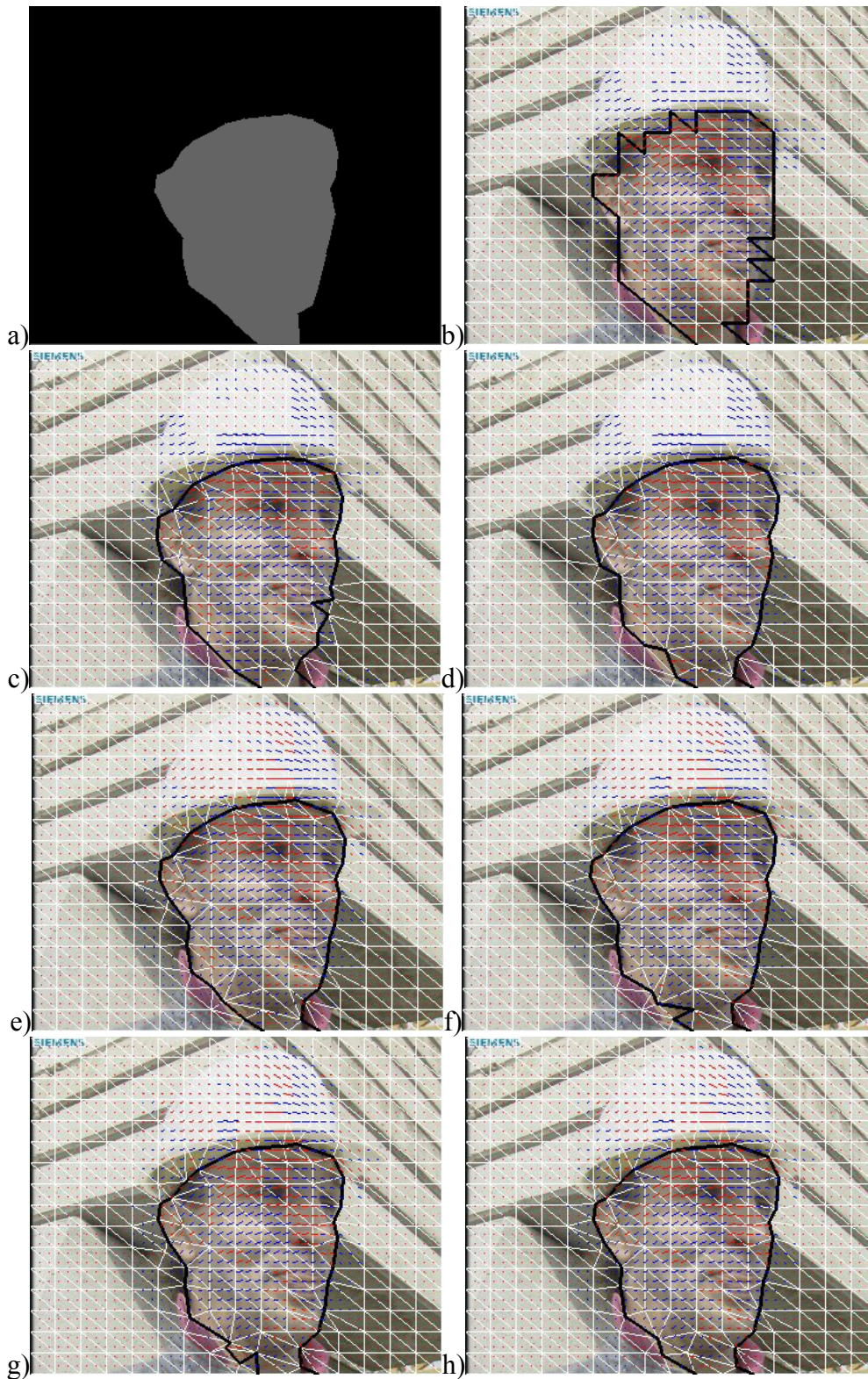


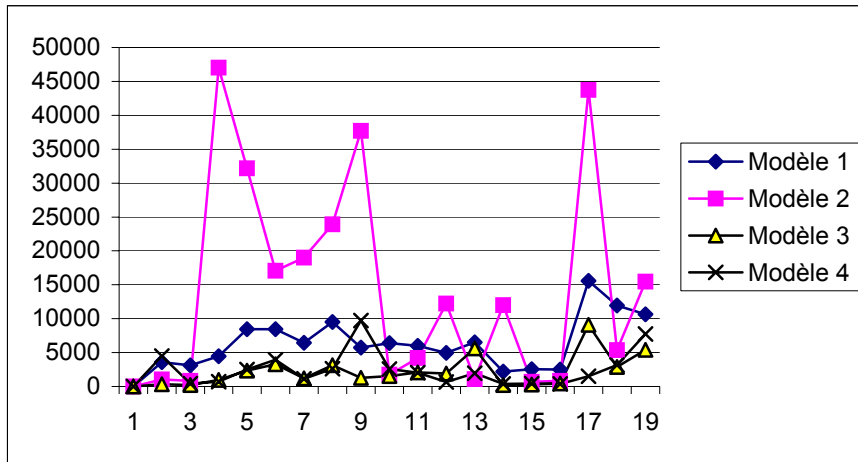
Figure 4.1.3 : déformation du maillage lors de la minimisation du critère d'énergie : masque de segmentation sur l'image 5(a), forme du maillage sur l'image 6 aux itérations 0, 1, 2, 4, 6, 8 et 20 (b-h)

## 4.2 Séquence synthétique

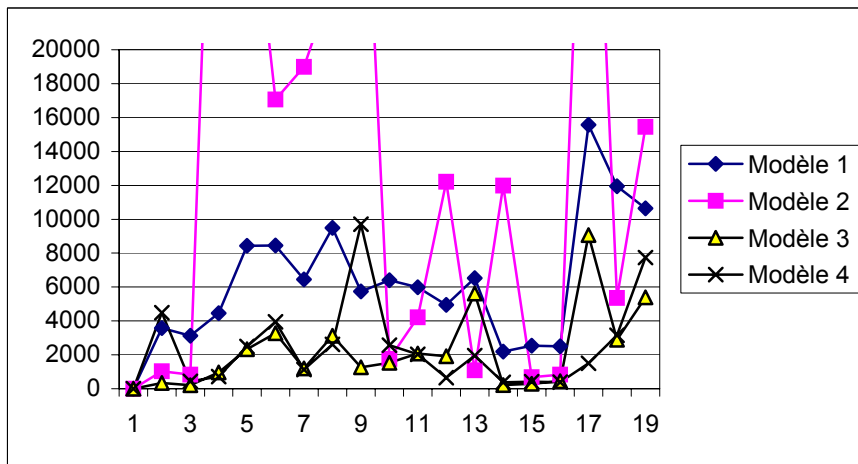
Nous avons créé deux séquences synthétiques pour évaluer quantitativement la qualité de notre algorithme en fonction du modèle de mouvement utilisé. Nous connaissons la carte de segmentation et le champ de mouvement de chaque image. Ces données de référence sont comparées avec les résultats de notre algorithme : estimation du mouvement et segmentation sur chaque image. Dans les modèles de mouvement, les paramètres employés sont  $\alpha = 1$ ,  $\beta = 200$  et  $S_{decoup} = 150$ .

Les deux séquences de 20 images représentent un quadrilatère texturé se déplaçant et subissant des déformations :

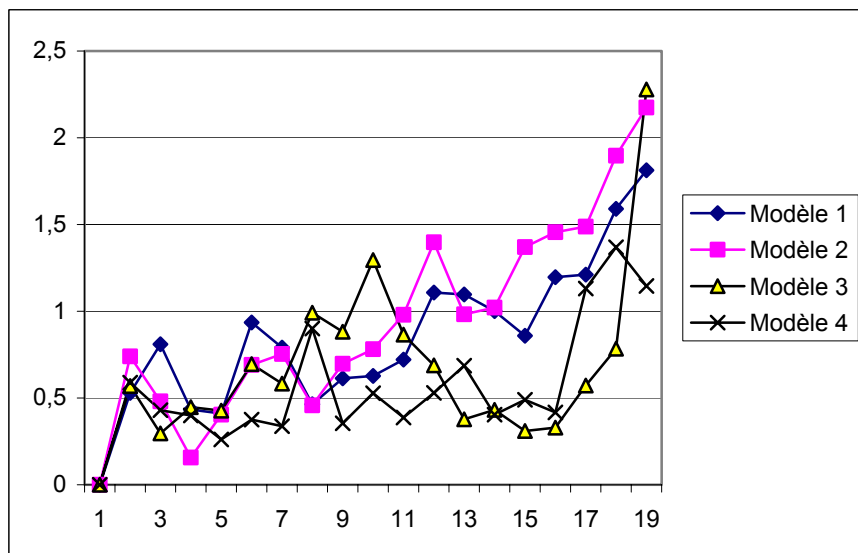
- Dans la première séquence, le fond est en translation constante. Les graphiques des figures 4.2.1.a et 4.2.1.b montrent l'erreur quadratique moyenne sur l'estimation du mouvement, selon les quatre modèles de mouvement employés. Le modèle de mouvement affine présente les résultats les moins réguliers, parfois moins bons que pour le modèle de translation qui est plus robuste. Les modèles 3 et 4 apportent les meilleurs résultats, par l'estimation d'un champ de mouvement régulier à l'intérieur des objets. La figure 4.2.1.c montre l'erreur de segmentation (en pourcentage de pixels mal classés dans l'image) de notre algorithme en fonction des 4 modèles de mouvement. On remarque que la qualité de la segmentation avec les modèles 1 et 2 se dégrade au fur et à mesure des images, faisant apparaître une dérive de notre algorithme. Les modèles 3 et 4 fournissent une segmentation de meilleure qualité, moins sensible à la dérive. Le masque de segmentation employé pour la première image est montré figure 4.2.2.a. Les figures 4.2.2.b à 4.2.2.h illustrent l'évolution du maillage sur les images 2,5,8,11,14,17,19 de la séquence (le modèle de mouvement employé est le modèle 3).
- Dans la deuxième séquence, le fond est en rotation constante, et la nature plus complexe du mouvement à estimer accentue les différences observées sur les résultats avec la première séquence. Les figures 4.2.3.a et 4.2.3.b montrent l'erreur quadratique moyenne d'estimation de mouvement. Le modèle 0 (translation) est toujours robuste, mais le modèle 1 (affine) offre parfois de meilleurs résultats. Le modèle 4 offre les meilleures performances, suivi par le modèle 3, comme l'illustre la figure 4.2.3.c. Le masque de segmentation initial utilisé sur la première image est présenté figure 4.2.4.a. Les figures 4.2.4.b à 4.2.4.h montrent les maillages construits sur les images 2,5,8,11,14,17,19 de la séquence, à l'aide du modèle 4. La figure 4.2.5 montre l'erreur d'estimation de mouvement par le modèle 4, sur les images 2, 3, 5, 8, 11, 14, 17 et 19 (l'histogramme des images est recadré).



a) Erreur quadratique moyenne d'estimation de mouvement x1000



b) Erreur quadratique moyenne d'estimation de mouvement x1000 (agrandissement)



c) Erreur de segmentation (en pourcentage de pixels mal classés dans l'image)

Figure : 4.2.1 : Résultats sur une séquence synthétique (fond en translation)

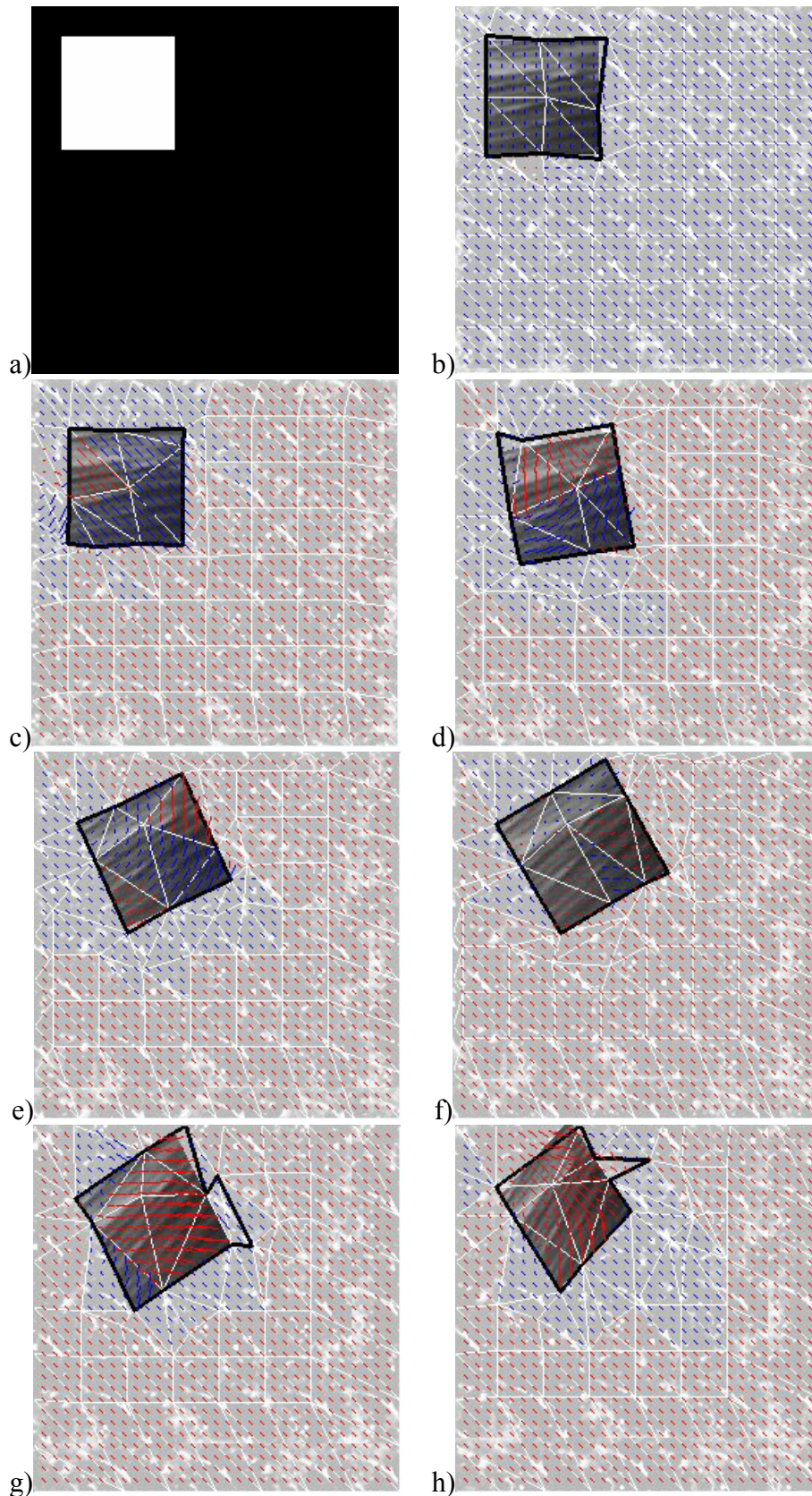
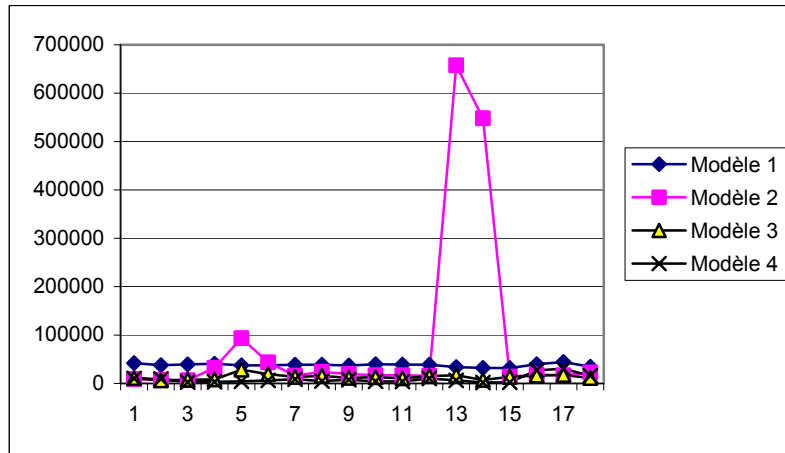
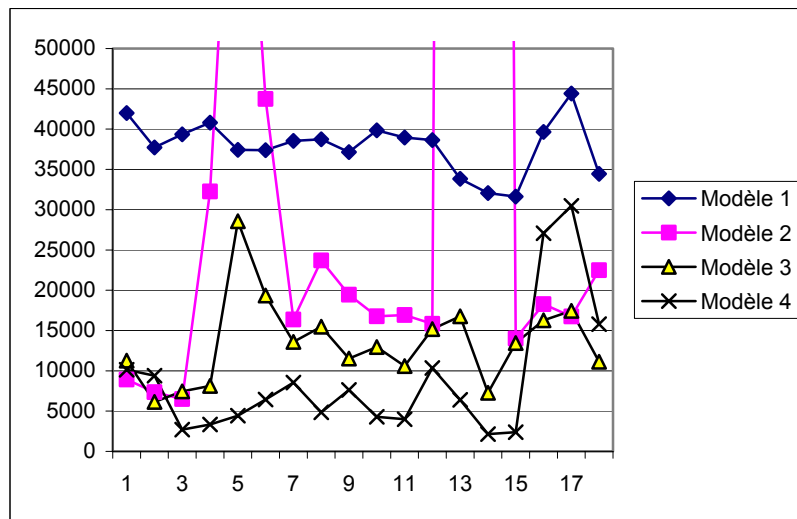


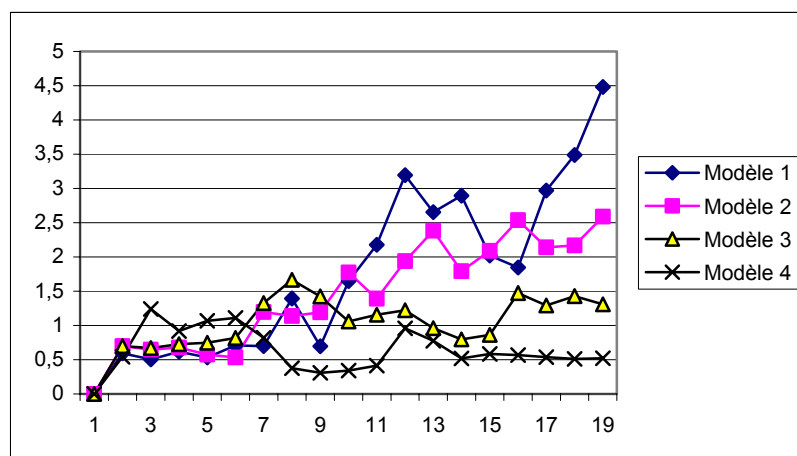
Figure 4.2.2 : évolution du maillage au cours de la séquence (fond en translation, modèle de mouvement n° 4) : masque de segmentation initiale sur l'image 1 (a) et évolution du maillage sur les images 2,5,8,11,14,17,19 (b-h)



a) erreur quadratique moyenne d'estimation de mouvement X1000

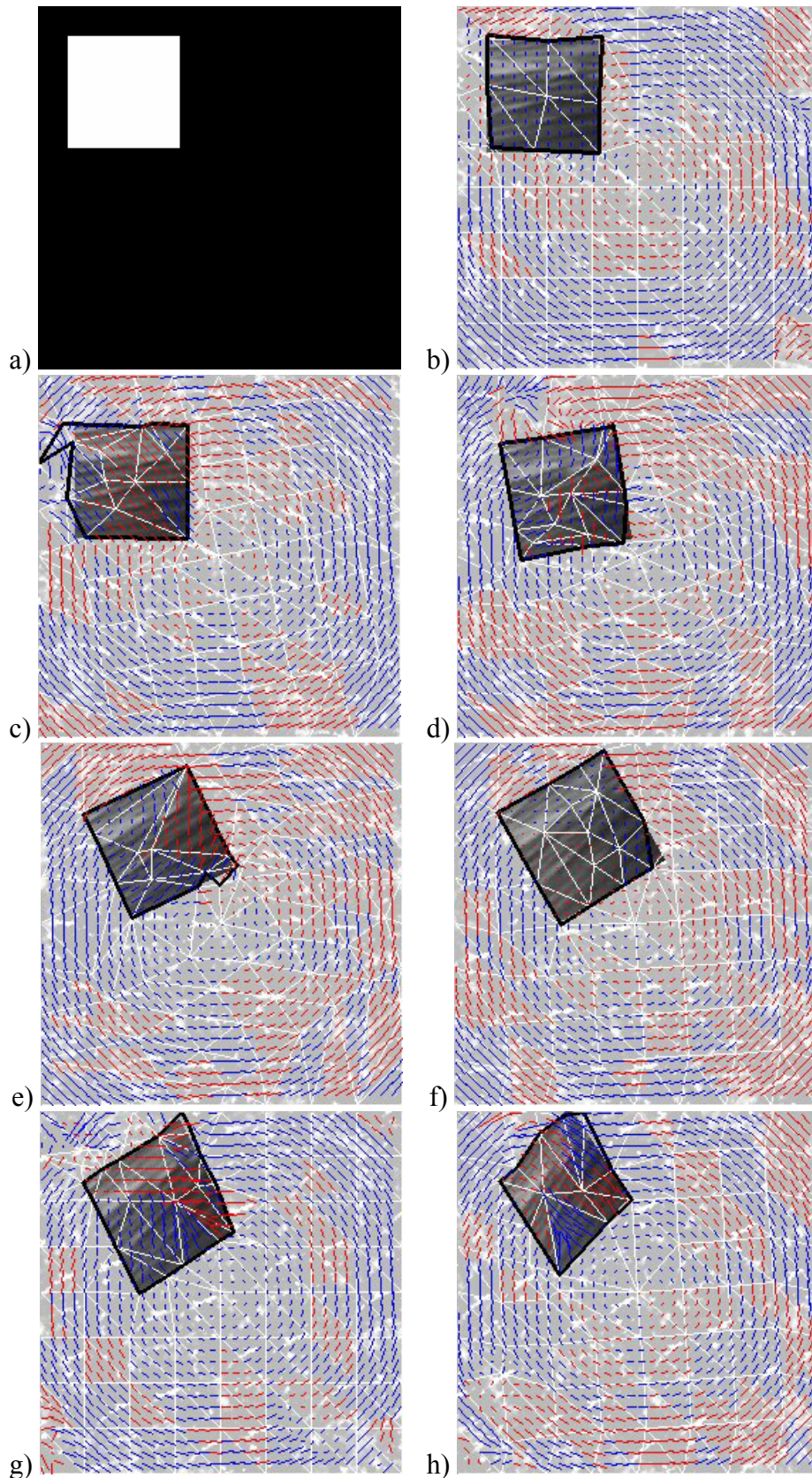


b) erreur quadratique moyenne d'estimation de mouvement X1000 (agrandissement)



c) erreur de segmentation (en pourcentage de pixels mal classés dans l'image)

Figure : 4.2.3 : Résultats sur une séquence synthétique (fond en rotation)



4.2.4 : évolution du maillage au cours de la séquence (fond en rotation, modèle de mouvement n° 4) : masque de segmentation initiale sur l'image 1 (a) et évolution du maillage sur les images 2,5,8,11,14,17,19 (b-h)

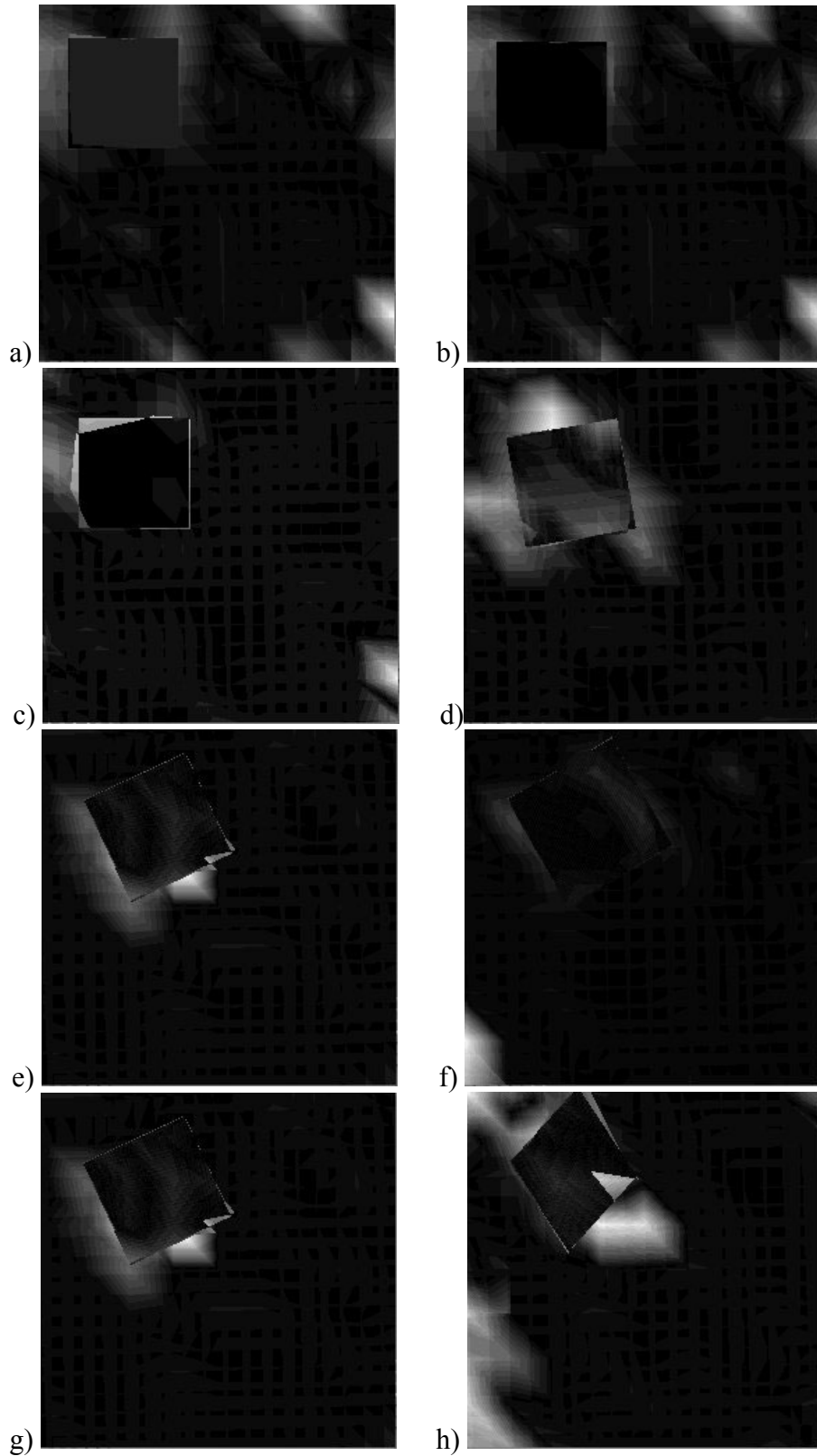


Figure 4.2.5 : erreur d'estimation de mouvement par le modèle 4 sur les images 2, 3, 5, 8, 11, 14, 17 et 19 de la séquence (fond en rotation).

### 4.3 Séquences réelles

Nous avons testé notre algorithme sur plusieurs séquences réelles :

- La figure 4.3.1 montre un exemple de segmentation semi-automatique : le maillage est construit sur l'image 100 de la séquence van (figure 4.3.1.a), en utilisant uniquement le critère d'énergie relatif au mouvement, avec le modèle de translation (modèle 1). L'algorithme est donc dans ce cas purement automatique. Or, à cause du problème d'ouverture, les frontières de mouvement sont mal déterminées (figure 4.3.1.b). En utilisant une interface graphique, nous pouvons modifier manuellement la position des nœuds du maillage et définir les objets vidéo dont on désire effectuer le suivi (figure 4.3.1.c et 4.3.1.g). Ensuite, l'algorithme poursuit automatiquement les objets dans la séquence (images 4.3.1.d à 4.3.1.f et 4.3.1.h à 4.3.1.j)

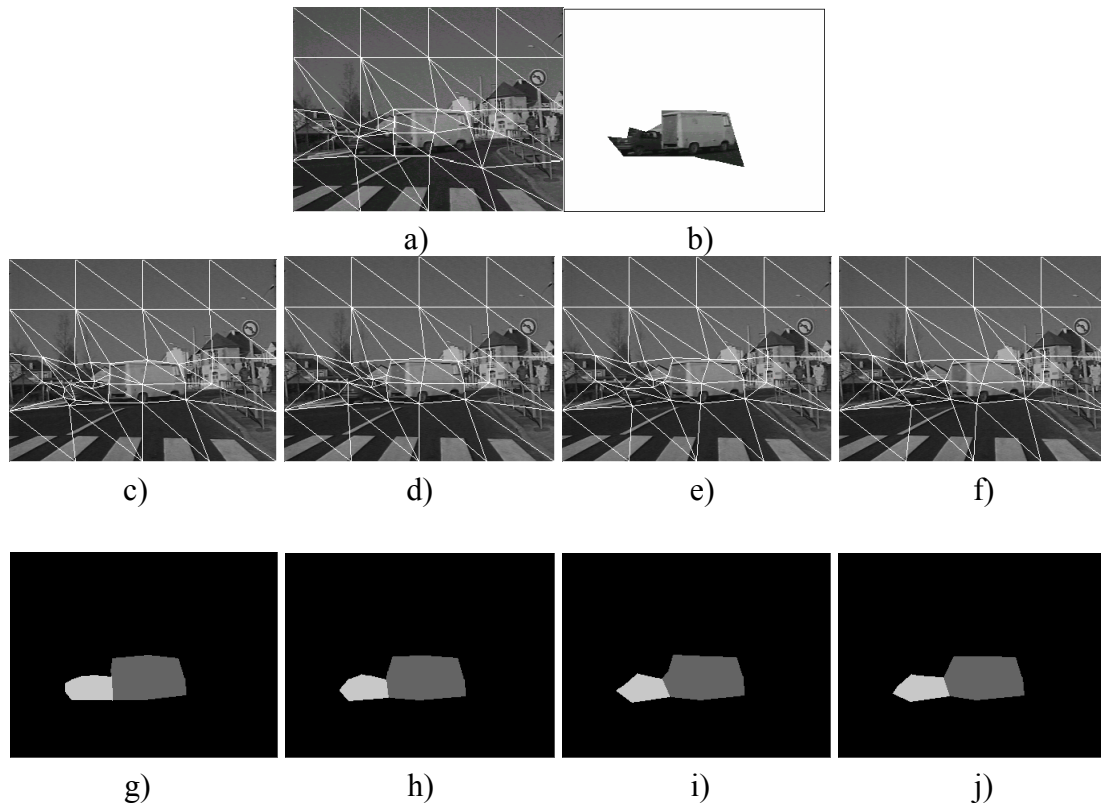


Figure 4.3.1 : segmentation semi-automatique et suivi sur la séquence « van » :

- maillage créé automatiquement (a) et segmentation initiale (b)
- maillage déformé manuellement (c) et segmentation résultante (g)
- suivi automatique sur les images suivantes (d-f) et segmentation automatique (h-j)

- La figure 4.3.2 montre les résultats obtenus avec la séquence van, mais ici, un masque de segmentation sur l'image 100 est fourni. Le modèle de mouvement employé est le modèle 1, les paramètres utilisés sont :  $\alpha = 6$ ,  $\beta = 100$  et  $S_{decoup} = 100$ . Le masque de segmentation sur l'image 100 est montré figure 4.3.2.a. Le suivi des objets sur les images 101, 104, 107, 110, 113, 116, 119 est montré figure 4.3.2.b à 4.3.2.h. On peut remarquer que dans la séquence, la voiture se déplaçant vers la gauche de l'image n'est pas détectée, car elle n'existait pas dans la segmentation initiale.
- La figure 4.3.3 montre le suivi des parois d'un cœur humain dans une séquence d'Images par Résonance Magnétique (IRM). Ici, le maillage est déformé manuellement sur l'image 1 (figure 4.3.3.a). Les parois du cœur sont ensuite suivies sur les 6 images suivantes de la séquence (figures 4.3.3.b à 4.3.3.g). Le modèle de mouvement utilisé est le modèle 4, les paramètres utilisés sont :  $\alpha = 1$ ,  $\beta = 100$ , et  $S_{decoup} = 200$ .
- La figure 4.3.4 illustre le suivi des deux bateaux présents dans la séquence « Coastguard ». Un masque de segmentation est fourni pour l'image 100 (figure 4.3.4.a). Les figures 4.3.4.b à 4.3.4.h montrent le résultat du suivi des objets sur les images 101, 104, 107, 110, 113, 116 et 119 de la séquence. Le modèle de mouvement utilisé est le modèle 4, les paramètres utilisés sont :  $\alpha = 1$ ,  $\beta = 100$  et  $S_{decoup} = 150$ .
- Sur la figure 4.3.5, on peut voir le suivi du visage du personnage de la séquence « foreman ». Le masque de segmentation est fourni pour l'image 5 (figure 4.3.5.a). Le suivi des images 6, 9, 12, 15, 18, 20 et 22 est illustré par les figures 4.3.5.b à 4.3.5.h. Le modèle de mouvement utilisé est le modèle 4, les paramètres utilisés sont :  $\alpha = 1$ ,  $\beta = 200$  et  $S_{decoup} = 200$ .

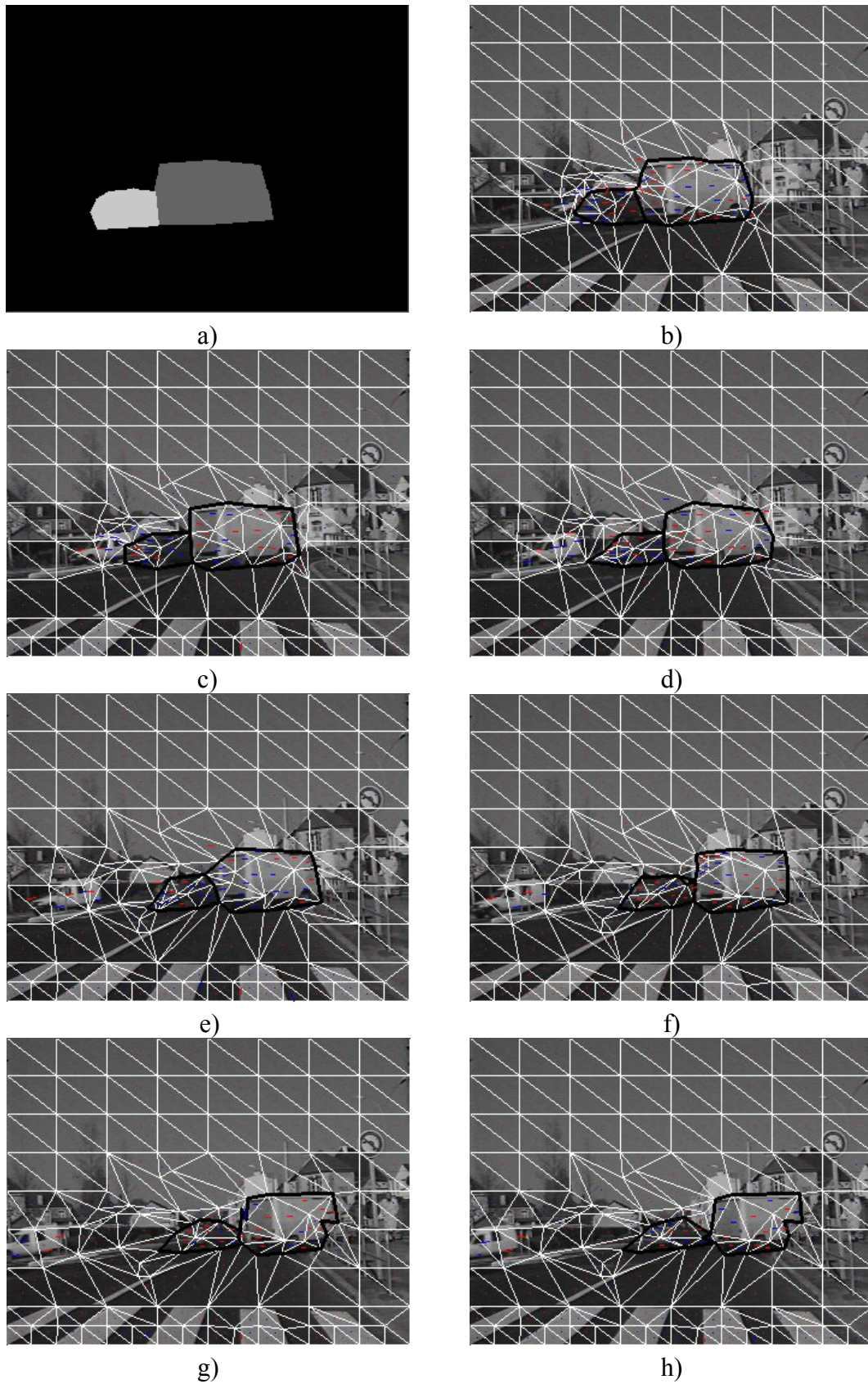


Figure 4.3.2 : suivi d'objets sur la séquence « van » : masque de segmentation initiale sur l'image 101 (a) et évolution du maillage sur les images 101, 104, 107, 110, 113, 116 et 119 (b-h)

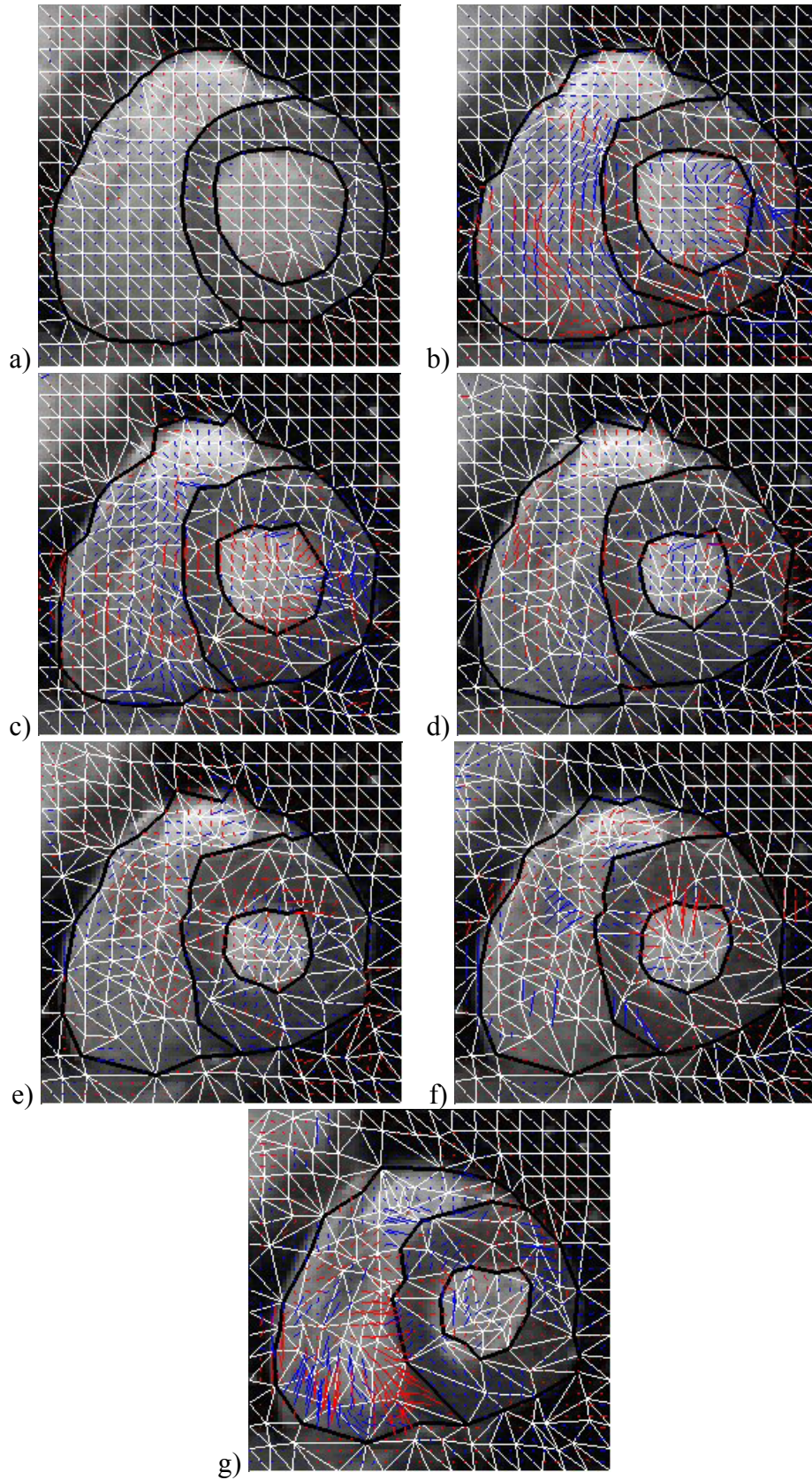


Figure 4.3.3 : Séquence IRM : maillage initial déformé manuellement sur l'image 1 (a) et suivi des parois sur les images 2 à 7 (b-g)

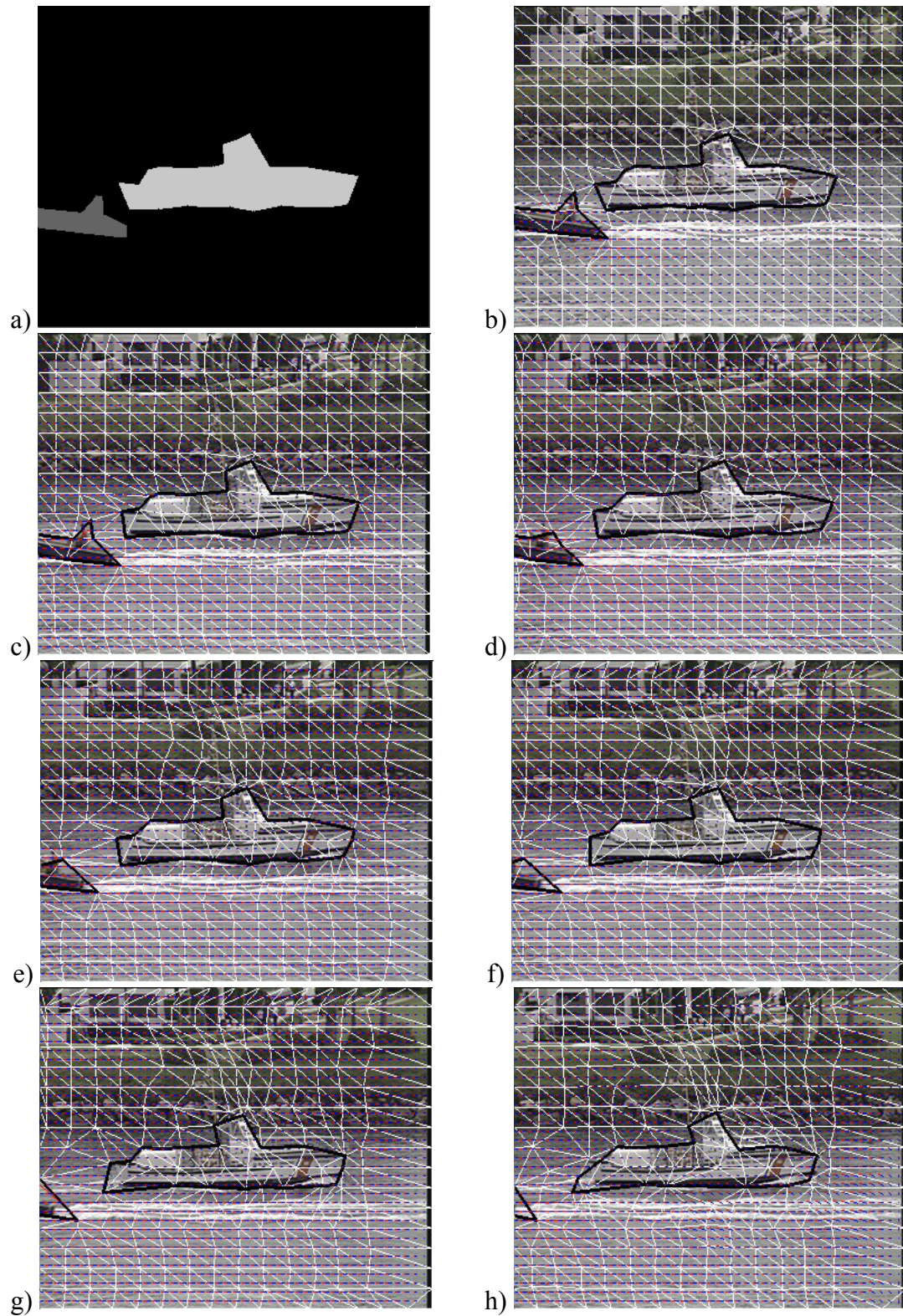


Figure 4.3.4 : Séquence « coastguard » : segmentation initiale sur l'image 100 (a), évolution du maillage sur les images 101, 104, 107, 110, 113, 116 et 119 (b-h)

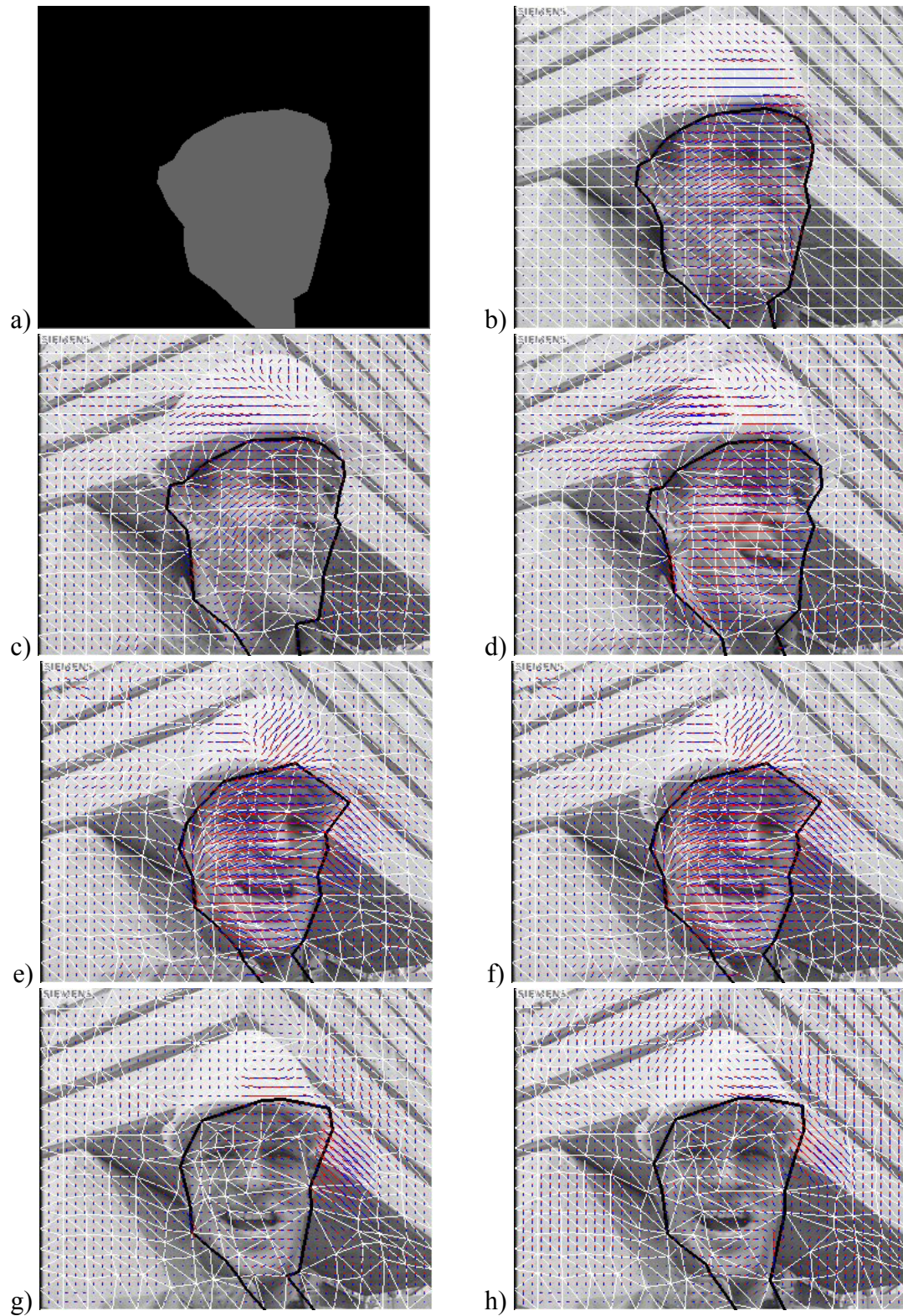


Figure 4.3.5 : séquence « Foreman » : masque de segmentation initiale (a) et évolution du maillage sur les images 6, 9, 12, 15, 18, 20 et 22 (b-h)

## 5 CONCLUSION ET PERSPECTIVES

Nous avons proposé, dans ce chapitre, un algorithme de segmentation semi-automatique et de suivi d'objets vidéo par l'utilisation d'un maillage actif. L'estimation de mouvement des cellules du maillage est effectuée selon un schéma d'estimation bidirectionnel (avant et arrière) prenant en compte les discontinuités de mouvement. Un critère spatial est introduit, améliorant la localisation des contours des objets dans les régions homogènes. Une contrainte de régularisation temporelle stabilise le suivi des objets au cours du temps.

Quelques pistes pour la poursuite de nos travaux sont proposés ci-dessous :

- Estimation du mouvement :

L'ajout d'une stratégie multiéchelle et/ou multirésolution offrirait une plus grande robustesse de l'algorithme pour les mouvements complexes ou de grande amplitude.

- Définition du critère spatial :

L'algorithme de détection de contours servant au calcul du critère spatial est basé sur des opérations de bas niveau (filtrage médian+seuillage sur le gradient morphologique de l'image). L'utilisation d'un algorithme de segmentation spatiale de plus haut niveau (utilisant la ligne de partage des eaux, par exemple) améliorerait la détection des contours de l'image, et par conséquent, la qualité du suivi des objets.

- Calcul du terme de régularisation temporelle :

Ce calcul est basé sur une projection arrière des cellules à l'instant  $t$  sur la segmentation au temps  $t+1$ . Or la projection arrière donne de mauvais résultats dans les régions découvertes de l'image, ce qui pénalise donc le critère de régularisation temporelle. Une perspective possible d'amélioration du calcul du terme de régularisation temporelle serait par exemple de ne pas prendre en compte les pixels dans les régions découvertes pour ce calcul. Ceci vaut aussi pour le calcul du critère de découpage qui est basé, lui aussi, sur une projection arrière.

- Choix des paramètres de pondération entre les différents critères d'énergie :

Une analyse préalable de la séquence à traiter pourrait fournir des informations statistiques permettant de fixer automatiquement les paramètres de pondération  $\alpha$  et  $\beta$ .

- Algorithme entièrement automatique :

L'utilisation d'un algorithme utilisant la détection de changement semble être une bonne méthode pour détecter correctement les objets en mouvement dans la séquence (voir section 2.2.5.C). Une telle implémentation nécessiterait cependant une redéfinition du critère global d'énergie actuel.

La grande force des maillages actifs est leur capacité à représenter les mouvements, il est fort probable que les futurs standards de compression de séquences vidéo utiliseront des algorithmes de compensation de mouvement utilisant les maillages. Cependant, dans un but de segmentation ou de suivi, une grande précision spatiale est demandée sur les contours des objets vidéo. Or dans le cas des maillages actifs, le contour des objets vidéo est polygonal. Si l'on désire une grande précision, le maillage devra être formé d'un grand nombre de cellules, afin que le polygone d'approximation soit assez précis. Utiliser un contour actif (polygonal ou à base de B-splines, par exemple) dont la précision serait au moins équivalente à celle d'un maillage actif semble donc plus intéressant. En étudiant le critère d'énergie, au cœur de notre algorithme, on peut remarquer :

- que notre algorithme fournit les meilleurs résultats quand le modèle de mouvement utilisé est basé sur les lignes de rupture. Dans ce cas, seule la frontière entre les objets (polygone) est prise en compte, le mouvement n'étant plus estimé par le maillage actif, mais par des maillages auxiliaires, de nature régulière.
- que le critère spatial est un critère basé contour.

L'adaptation de notre algorithme à un contour actif semble donc une tâche aisée, sachant que seul le terme de régularisation temporelle est à reformuler.

# **Chapitre IV**

## **Bilan et perspectives**

## **1 BILAN ET PERSPECTIVES**

Nous avons présenté dans ce mémoire une méthode d'analyse multirésolution des maillages irréguliers et un algorithme de segmentation et de suivi d'objets vidéo. Ces deux approches peuvent être combinées, afin de fournir un algorithme de représentation multirésolution des objets vidéo : le maillage créé sur chaque image peut être simplifié. Les objets vidéo seraient donc représentables à plusieurs niveaux de précision.

Le nouveau standard audio-visuel MPEG-4 introduit la notion de représentation des scènes multimédia sous plusieurs niveaux de qualité, en fonction du canal de transmission des informations, ou de la puissance de calcul du terminal de l'utilisateur. On parle alors de codage scalable des scènes multimédia.

Plusieurs travaux proposent une représentation scalable des objets vidéo.

Lebuhan-Jordan [LEBU98] propose une approximation progressive des contours des objets vidéo par des polygones dont la complexité augmente en fonction de la précision spatiale requise.

Van-Beek [VANB99] propose une représentation des objets vidéo par maillage objet (voir Chapitre III section 2.2.4.d). Plusieurs niveaux de hiérarchie sont créés à partir du maillage haute résolution, par suppression des nœuds, selon un critère d'approximation du contour pour les nœuds situés sur les bords du maillage, et selon un critère de qualité de représentation pour les nœuds intérieurs.

Nous proposons dans ce chapitre une méthode de représentation multirésolution des objets vidéo présents dans une séquence à l'aide d'un maillage. Contrairement aux méthodes généralement employées, le maillage que nous utilisons recouvre toute l'image et sert à la représentation de tous les objets présents dans la scène. Il est ensuite simplifié en accord avec l'algorithme d'analyse multirésolution des maillages 3D que nous avons décrit dans le chapitre II. La nature réversible de la simplification permet donc une représentation multirésolution des objets présents dans la scène.

## **2 VERS UNE REPRESENTATION MULTIREOLUTION DES OBJETS VIDEO : METHODE PROPOSEE**

Pour illustrer notre approche, nous nous intéressons à un maillage créé pour le suivi du personnage présent dans l'image 6 de la séquence « foreman », selon l'algorithme proposé dans le chapitre II (figure 2.1).

La représentation du masque des deux objets présents dans l'image (le personnage et le fond) par un tel maillage nécessite la description :

- Du maillage en question
- De l'appartenance de chaque cellule à l'un des deux objets présents (La frontière entre les deux objets est alors un polygone).

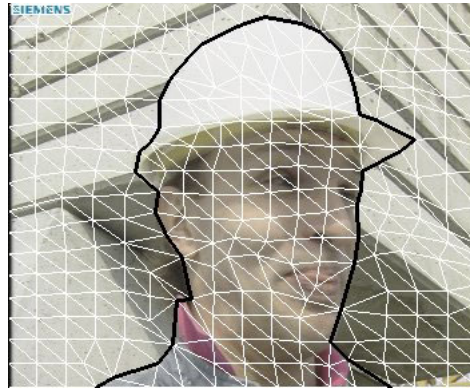


Figure 2.1 : Segmentation de la séquence « Foreman » : image 6.

On peut alors utiliser une approche de fusion irrégulière des cellules du maillages (proposée dans le chapitre II) pour avoir une représentation scalable des objets présents dans la scène. Cependant, deux contraintes de simplification supplémentaires viennent s'ajouter à celles décrites dans le chapitre II :

- La fusion des cellules du maillage ne doit plus s'effectuer en considérant uniquement la connectivité du maillage mais doit aussi prendre en compte la position des objets vidéo par rapport au maillage. En effet, la fusion de cellules appartenant à deux objets vidéo différents est à proscrire, comme le montre la figure 2.2.
- Un nœud situé sur le bord du maillage ne pourra être considéré comme nœud fils (et supprimé durant la simplification) que si ses deux nœuds pères sont aussi situés sur le bord du maillage (figure 2.3).

Ainsi, la conception d'un algorithme de fusion de cellules est plus complexe que dans le cas des maillages 3D et par manque de temps, nous n'avons pas pu adapter notre algorithme de simplification des maillages 3D au cas 2D.

Nous avons néanmoins effectué une simplification manuelle sur le maillage de la figure 2.1. La simplification est manuelle, mais l'approximation (c'est-à-dire le calcul des coordonnées des nœuds des maillages de résolution inférieure) est effectuée automatiquement par notre algorithme d'analyse multirésolution proposé dans le chapitre II, auquel on fournit les informations de subdivision irrégulière (qui sont les opérations inverses des différentes étapes de simplification).

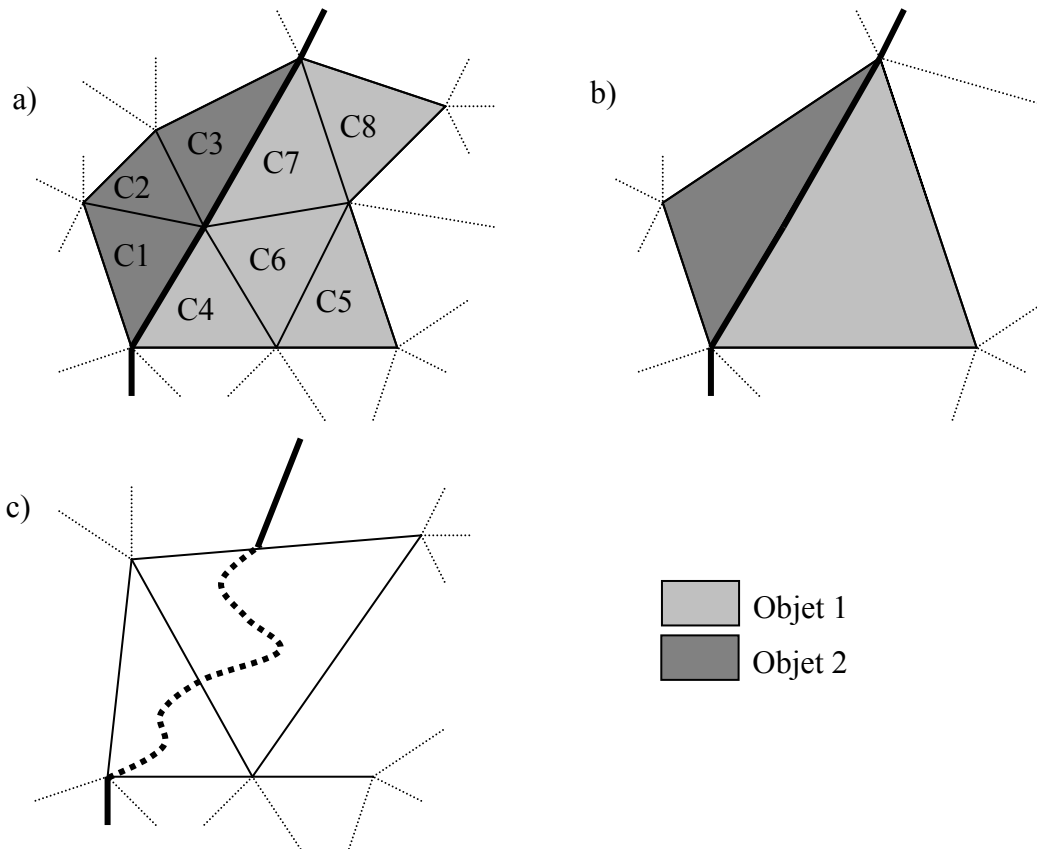


Figure 2.2 : Représentation d'une région frontière entre deux objets (a). La fusion des cellules C1, C2 et C3 ainsi que les cellules C4, C5, C6, C7 et C8 permet une bonne représentation de la frontière entre les deux objets (b). Au contraire, la fusion des cellules C1, C2 et C4 ainsi que les cellules C3, C6, C7 et C8 ne permet pas une représentation aisée de la frontière entre les objets. Ce type de fusion doit donc être interdit.

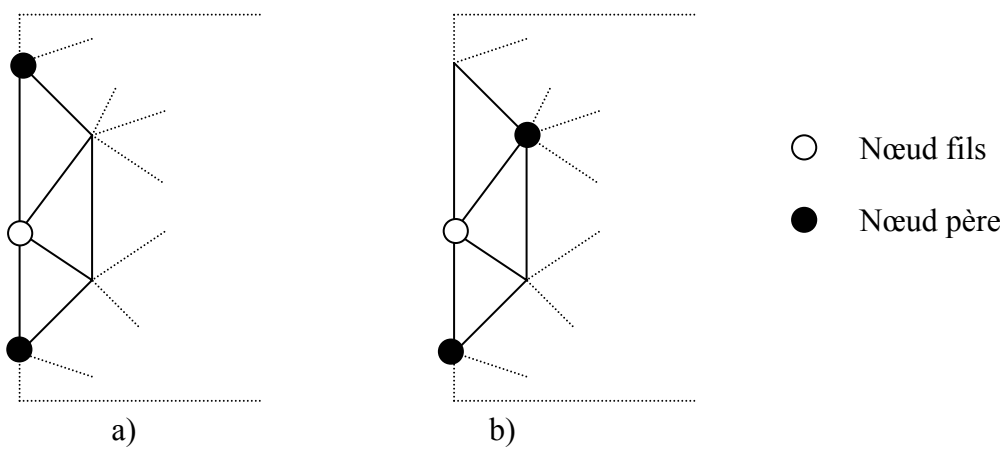


Figure 2.3 : contrainte sur les bords du maillage : exemple de configuration autorisée (a) et interdite (b)

### **3 RESULTATS EXPERIMENTAUX**

La figure 3 montre les 8 niveaux ( figure 3.a à 3.h) d'approximation obtenus à partir du maillage original. Ces maillages sont obtenus par prise en compte du maillage haute résolution uniquement, sans considérer le contenu spatial de l'image, mais la qualité des approximations est tout de même bonne. Un défaut apparaît cependant : dans les niveaux de résolution les plus bas, les contours du maillage ne correspondent plus avec les contours de l'image.

### **4 CONCLUSION ET PERSPECTIVES**

Nous avons présenté dans ce chapitre une méthode de représentation multirésolution des objets vidéo, par décomposition du maillage sur une base d'ondelettes irrégulières. Elle utilise la proposition du chapitre II. Cet algorithme permet une transmission progressive de la description des objets vidéo, mais pourrait aussi servir à estimer le mouvement dans les séquences par une approche multiéchelle, ce qui améliorerait la qualité de l'estimation du mouvement de l'algorithme proposé dans le chapitre III. Des travaux complémentaires sont néanmoins nécessaires :

- Concevoir un algorithme de simplification purement automatique, intégrant les contraintes liées à la définition des objets et aux bords du maillage.
- Modifier l'algorithme de décomposition sur la base d'ondelettes pour que les nœuds sur les bords du maillage restent sur les bords de l'image pour tous les niveaux de résolution.

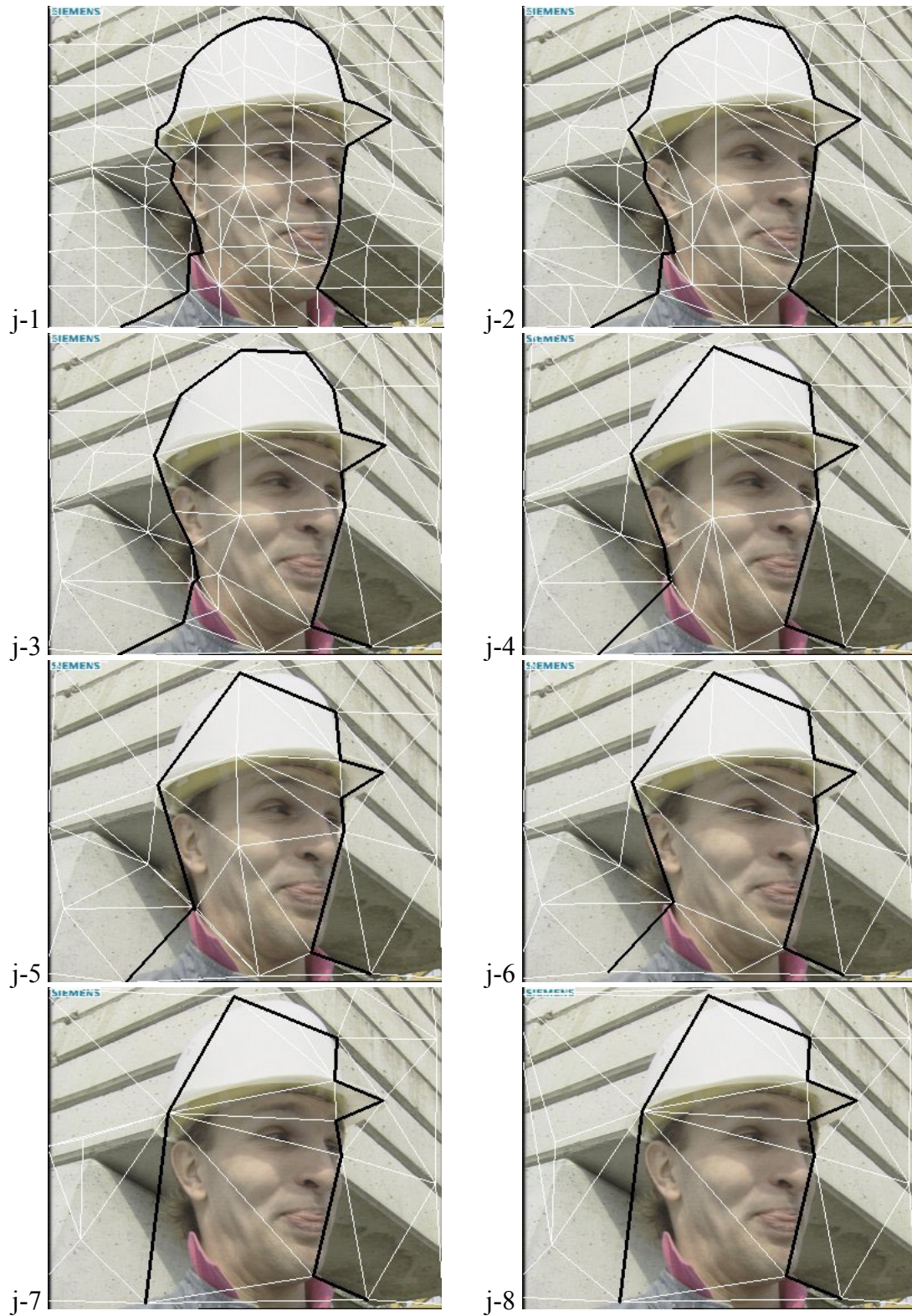


Figure 3 : approximations successives du maillage représentant la segmentation sur l'image 6 de la séquence « Foreman »

## Bibliographie

- [ALLI01a] P. Alliez, M. Desbrun, Valence-Driven Connectivity Encoding for 3D Meshes, *EUROGRAPHICS 2001*, Volume 20, No. 3, 2001, 10p.
- [ALLI01b] P. Alliez, M. Desbrun, Progressive Encoding for Lossless Transmission of 3D Meshes, *Proceedings of SIGGRAPH 2001.*, 8p.
- [ALAT98] A. Alatan, L. Onural, M. Wollborn, R. Mech, E. Tuncel, T. Sikora: Image sequence analysis for emerging interactive multimedia services - The European COST 211 framework, , *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 8, No. 7, Nov. 1998, pp. 802-813.
- [ALTU96] Y. Altunbasak, Object-scalable, content-based video representation and motion tracking for visual communications and multimedia, USA, *PhD. Thesis, University of Rochester, School of Engineering and Applied Sciences*, 1996, 238p.
- [ALTU97b] Y. Altunbasak and A. M. Tekalp, Closed-form connectivity-preserving solutions for motion compensation using 2-D meshes, *IEEE Trans. Image Processing*, vol. 6, no 9, September 1997, pp. 1255-1269.
- [ALTU98] Y. Altunbasak and A. M. Tekalp, A Hybrid Video Codec with Block-Based and Mesh-Based Motion Compensation Modes, *Int. J. Imaging Syst. and Tech.* (special issue), vol. 9, no. 4, August 1998, pp. 248-256.
- [ALTU97a] Y. Altunbasak and A. M. Tekalp , Occlusion-adaptive content-based 2-D mesh design and tracking for object-based coding, *IEEE Transactions on Image Processing*, vol. 6, no. 9, September 1997, pp. 1270-1280.
- [ALTU97b] Y. Altunbasak and A. M. Tekalp, Closed-form connectivity-preserving solutions for motion compensation using 2-D meshes, *IEEE Trans. Image Processing*, vol. 6, no 9, September 1997, pp. 1255-1269.
- [BELZ98] G. Belz, P. Pourteyron, F. Thibon et S. Valette, Compression / Décomposition et Transmission Progressive d'images 3D Représentées par des surfaces, *Projet de Fin d'Etudes, Département Génie Electrique, Institut National des Sciences Appliquées de Lyon*, Juin 1998, 50p.
- [BENO98] H. Benoit-Cattin, A. Planat-Chrétien, P. Joachimschmann, A. Baskurt, P. Clarysse, and I. E. Magnin, On the coding of active mesh, , *IEEE Int. Conf on Image Processing ICIP'98* , Chicago, USA, 4-7 October 1998, Vol. 2, pp. 895-898.
- [BENO99] H. Benoit-Cattin, P. Joachimschmann, A. Planat, S. Valette, A. Baskurt and R. Prost, Active mesh texture coding based on warping and DCT, *IEEE Int. Conf on Image Processing ICIP'99*, October 25-28, Kobe, Japan, Vol. 2, pp 830-833.
- [BONN98] G.-P. Bonneau, Multiresolution analysis on irregular surface meshes, *IEEE transactions Visualization and Computer Graphics*, Volume 4, Number 4, October-December, 1998, pp. 365-378.
- [BORS97] G. Borshukov, G. Bozdagi, Y. Altunbasak and A. M. Tekalp, Motion Segmentation by multistage affine classification, *IEEE Transactions on Image Processing*, vol. 6, November 1997, pp. 1591-1594.

- [BRUS90] H. Brusewitz, Motion compensation with triangles, *Proc. 3<sup>rd</sup> Inter. Conf. On 64kbit Coding of Moving Video*, Rotterdam, Netherlands, Sept. 1990.
- [CALD98] R. C. Calderbank, I. Daubechies, W. Sweldens, and B.-L. Yeo, Wavelet Transforms that Map Integers to Integers, *Applied and Computational Harmonic Analysis (ACHA)*, Vol. 5, No. 3, 1998, pp. 332-369.
- [CAPL01] A. Caplier, L. Bonnaud, J.-M. Chassery, Robust Fast Extraction of Video Objects Combining Frame Differences and Adaptive Reference Image, , *IEEE International Conference on Image Processing*, Thessaloniki, Greece, October 7-10, 2001.
- [CATM78] E. Catmull, J. Clark, Recursively Generated B-spline Surfaces on Arbitrary Topological Meshes”, *Computer Aided Design*, Vol. 10, No. 7, 1978, pp. 350-355.
- [CELA00] I. Celasun and A. M. Tekalp, Optimal 2D hierarchical content-based mesh design and update for object-based video, *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 10, no. 7, Oct. 2000, pp. 1135-1153.
- [CELA01] I. Celasun, A. M. Tekalp, M. H. Gokcetekin, and D. M. Harmanci, 2-D mesh-based video object segmentation and tracking with occlusion resolution, *Signal Processing: Image Communication*, vol. 16, no. 10, August 2001, pp. 949-962.
- [CHRE99] A. Chrétien-Planat, Estimation de mouvement par maillage actif multiéchelle avec prise en compte des discontinuités : Application à l'imagerie cardiaque en Résonance Magnétique, *Thèse de Doctorat, Institut National des Sciences Appliquées de Lyon*, 1999, 201p.
- [CIGN98] P. Cignoni, C. Rocchini, and R. Scopigno, Metro: Measuring Error on Simplified Surfaces. *Computer Graphics Forum* 17, Vol. 2, 1998, pp. 167-174.
- [COCQ98] J.-P. Coquerez et S. Philipp, Analyse d'images : filtrage et segmentation, *Paris : MASSON*, 1995, pp. 57.
- [COHE99] D. Cohen-Or, D. Levin and O. Remez, Progressive Compression of Arbitrary Triangular Meshes, *IEEE Visualisation 99 Conference Proceedings*, 1999, pp. 67-72.
- [DEER95] M. Deering, Geometric compression, *SIGGRAPH'95 Conference Proceedings*, August 1995, pp. 13-20.
- [DELI99] H. Delingette, General object reconstruction based on simplex meshes, *International Journal of Computer Vision*, vol. 32, no. 2, September 1999, pp. 111-146.
- [DOO78] D. Doo, M. Sabin, Behaviour of Recursive Division Surfaces Near Extraordinary Points, *Computer Aided Design*, Vol. 10, No. 6, November 1978, pp. 356-360.
- [DYN90] N. Dyn, D. Levin, J. Gregory, A Butterfly Subdivision Scheme for Surface Interpolation with Tension Control, *ACM Trans. On Graphics*, Vol.9, No. 2, April 1990, pp. 160-169.
- [ECK95] M. Eck, T. DeRose, T. Duchamp, H. Hoppe, M. Lounsbery and W. Stuetzle, Multiresolution Analysis of Arbitrary Meshes. *Computer Graphics (SIGGRAPH 1995 Proceedings)*, 1995, pp. 173-182.
- [FREY99] P. J. Frey, P.-L. Georges, Maillages, application aux éléments finis, Paris, Hermes, 1999, p. 840.
- [FU00] Y. Fu, A. T. Erdem, and A. M. Tekalp, Tracking visible boundary of objects using occlusion adaptive motion snake, *IEEE Transactions on Image Processing*, vol. 9, no. 12, December 2000, pp. 2051-2060.

- [GU98] C. Gu and M. Lee, Semantic video object tracking using region-based classification, *IEEE International Conference on Image Processing ICIP '98*, Chicago, USA, 4-7 October, 1998, pp. 643-647.
- [GUSK99] I. Guskov, W. Sweldens, P. Schröder, Multiresolution Signal Processing for Meshes, *International Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '99*, Los Angeles, California, USA, August 1999, pp. 325-334.
- [GUSK00] I. Guskov, K. Vidimce, W. Sweldens, P. Schröder, Normal meshes, *Computer Graphics Proceedings (SIGGRAPH 2000)*, pp 95-102, 2000
- [HECK97] P. S. Heckbert and M. Garland, Survey of Polygonal Surface Simplification Algorithms, *Multiresolution Surface Modeling Course, ACM Siggraph Course notes*, 1997. [en ligne] <http://www.cs.cmu.edu/~ph>
- [HOPP96] H. Hoppe, Progressive meshes. International Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '96, August 96, pp. 99-108.
- [HORN81] B. Horn, B. Schunck, Determining Optical Flow, *Artificial intelligence*, 1981, vol. 17, pp. 185-203.
- [HUAN94] C.-J. Huang and C.-Y. Hsu, A New Motion Compensation Method for Image Sequence Coding Using Hierarchical Grid Interpolation, *IEEE transactions on Circuits and Systems for Video Technology*, Vol. 4, No. 1, February 1994, pp. 42-52.
- [ISEN99] M. Isenburg and J. Snoeyink, Mesh Collapse Compression, *Proceedings of ACM Symposium on Computational Geometry*, 1999, pp. 419-420.
- [JAIN81] J.R. Jain and A.K. Jain, Displacement measurement and its applications in interframe coding, *IEEE Trans. Comm.*, Vol. COM-29, No 12, Dec 1981, pp 1799-1808.
- [JEHA01a] S. Jehan-Besson, M. Barlaud, G. Aubert, Video object segmentation using eulerian region-based active contours, *International Conference on Computer Vision*, Vancouver, july 2001, pp. 353-361.
- [JEHA01b] S. Jehan-Besson, M. Barlaud, G. Aubert, Region-based active contours for video object segmentation with camera compensation, *IEEE International Conference on Image Processing*, Thessaloniki, Greece, October 7-10, 2001, pp. 61-64.
- [JUNG98a] H.Y. Jung, T.Y. Choi, R. Prost, Rounding Transform and Its Application For Pyramid Structured Coding, *IEEE Trans. on Image Processing*, vol. 7, no. 2, 1998, pp. 234-237.
- [JUNG98b] H.Y. Jung, R. Prost, Lossless Subband Coding System Based on Rounding Transform, *IEEE Trans. On Signal Processing*, vol. 46, No. 9, Sept. 1998, pp. 2535-2540.
- [KHOD00] A. Khodakovsky, P. Schröder, W. Sweldens, Progressive Geometry compression, *International Conference on Computer Graphics and Interactive Techniques, SIGGRAPH 2000*, 2000, pp. 271-278.
- [KIM99] Y.-S. Kim, S. Valette, H.-Y. Jung and R. Prost, Local wavelets decomposition for 3D surfaces, *IEEE Int. Conf on Image Processing ICIP'99*, October 25-28, Kobe, Japan, vol.1, pp. 357-360, October 1999.
- [KIM01a] Y.-S. Kim, Analyse adaptative et synthèse multirésolution de surfaces maillées par décomposition sur une base d'ondelettes, *Thèse de doctorat, Institut National des Sciences Appliquées de Lyon*, No. d'ordre 2001 ISAL 0003, Janvier 2001.

- [KIM01b] Y.-S. Kim, S. Valette and R. Prost, Adaptive wavelets based multiresolution modeling of irregular meshes via harmonic maps, *IEEE Int. Conf. on Image Processing ICIP 2001*, Thessaloniki, Greece, October 7-10, 2001, Vol. III, pp. 210-213.
- [KIRK83] S. Kirkpatrick, C. Gellat and M. Vecchi, Optimisation by simulated annealing, *Science*, Vol. 220, 1983, pp. 671-680.
- [KUHN01] G. Kühne, J. Weickert, O. Schuster, S. Richter, A Tensor-Driven Active Contour Model for Moving Object Segmentation, *IEEE International Conference on Image Processing*, Thessaloniki, Greece, October 7-10, 2001, pp. 73-76.
- [LATZ01] M. Latzel, J. K. Tsotsos, A Robust Motion Detection And Estimation Filter For Video Signals, *IEEE International Conference on Image Processing*, Thessaloniki, Greece, October 7-10, 2001, pp. 381-384.
- [LAUR98] N. Laurent, P. Lechat and H. Sanson, Limitation Of Triangles Overlapping In Mesh-Based Motion Estimation Using Augmented Lagrangian, *IEEE International Conference on Image Processing*, Chicago, USA, 4-7 October, 1998, pp. 223-227.
- [LEBU98] C. LeBuhan-Jordan, Progressive geometrical compression of arbitrary shaped video objects, *PhD. Thesis, EPFL*, Lausanne, Suisse, 1998, 183p.
- [LECH98] P. Lechat and H. Sanson, Combined Mesh Based Image Representation and Motion Estimation, Application to Video Coding, *IEEE International Conference on Image Processing ICIP'98*, Chicago, USA, 4-7 October, 1998, 4p.
- [LECH99] P. Lechat, N. Laurent, H. Sanson, Suivi d'Objets Vidéo par Maillages Hiérarchiques, *17<sup>e</sup> colloque GRETSI 1999*, Vannes, 13-17 septembre 1999, pp. 507-510.
- [LEFL97] H. Le Floch, Echantillonnage irrégulier et interpolation : application à la représentation d'images fixes et animées, *Thèse de Doctorat, Université de Rennes-I*, No. d'ordre 191h, Décembre 1997, 201p.
- [LEE98] A. W. F. Lee, W. Sweldens, P. Schröder, L. Cowsar, D. Dobkin, MAPS : Multiresolution Adaptive Parametrization on Surfaces", *International Conference on Computer Graphics and Interactive Techniques, SIGGRAPH'98*, Orlando, Florida, USA, July 1998, pp. 95-104.
- [LORE87] W. E. Lorensen, H. E. Cline, Marching Cubes: A High Resolution 3D Surface Construction Algorithm, *Computer Graphics (Proceedings of Siggraph'87)*, Vol. 21, No. 4, July 1987, pp 163-169.
- [LOOP87] C. T. Loop, Smooth Subdivision Surfaces Based on Triangles, *Master's Thesis, University of Utah*, August 1987, P. 74.
- [LOTJ98] J. Lotjonen, P.J. Reissman, I.E. Magnin, J. Nenonen, and T. Katila, A triangulation method of an arbitrary point set for biomagnetic problem, *IEEE Transactions on Magnetics*, Vol. 34, No. 4, July 1998, pp. 228-233.
- [LOUN94] M. Lounsbery. Multiresolution Analysis for Surfaces of Arbitrary Topological Type, *PhD thesis, Dept. of Computer Science and Engineering, University of Washington*, 1994, 112p.
- [LOUT01] E. Loutas, K. Diamantaras, I. Pitas, Occlusion resistant object tracking, *IEEE International Conference on Image Processing*, Thessaloniki, Greece, October 7-10, 2001, pp. 65-68.
- [LOUN97] M. Lounsbery, T. Derose, J. Warren, Multiresolution Analysis for Surfaces of Arbitrary Topological Type, *ACM Trans. On Graphics*, Vol. 16, January 1997, No. 1, pp. 34-73.

- [NAKA91] Y. Nakaya and H. Harashima, An iterative motion estimation method using triangular patches for motion compensation, *Proc. SPIE Visual Comm. and Image Proc.*, Boston, MA, Nov. 1991, pp. 546-557.
- [NOSR01] A. Nosratinia, New Kernels for Fast Mesh-Based Motion Estimation, *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 11, No. 1, January 2000, pp. 40-51.
- [MALL89] S. G. Mallat, A Theory for Multiresolution Signal Decomposition : The Wavelet Representation, *IEEE Trans. On Pattern Analysis And Machine Intelligence*, Vol. 11, July 1989, No. 7, pp. 674-693.
- [MANS99] A.-R. Mansouri and J. Konrad, Motion Segmentation with Level Sets, *IEEE Int. Conf on Image Processing ICIP'99*, Kobe, Japan, October 25-28, 1999, 4p.
- [MARQ00a] G. Marquant, Représentation par maillage adaptatif déformable pour la manipulation et la communication d'objets vidéo, *Thèse de Doctorat, Université de Rennes-I*, Décembre 2000, 296p.
- [MARQ00b] G. Marquant, S. Pateux and C. Labit, Mesh and « Crack Lines » : Application to Object-based Motion Estimation and Higher Scalability, *IEEE International Conference on Image Processing ICIP'00*, Vancouver, Canada, September 2000, 4p.
- [MARQ00c] G. Marquant, S. Pateux and C. Labit, Mesh-based Scalable Video Coding with Rate-Distortion Optimisation, *Visual Communication and Image Processing VCIP'00*, Perth, Australia, June 2000.
- [MEYE93] Y. Meyer, Les Ondelettes : Algorithmes et Applications, *Paris : Armand Colin, 1993*, p. 196.
- [NICO01] H. Nicolas, New Methods for Dynamic Mosaicking, *IEEE Transactions on Image Processing*, Vol. 10, No. 8, August 2001, pp. 1239-1251.
- [MOST99] S. M. M. Mostafavi Gorji, Une méthode de segmentation spatio-temporelle opérant par analyse globale du mouvement, *Thèse de Doctorat, Institut National des Sciences Appliquées de Rennes*, No. d'ordre D99-20, Décembre 1999, 163p.
- [MPEG98] ISO/IEC JTC1/SC29/WG11 N2202, Information technology- Coding of audio-visual objects : Visual, *Tokyo*, March 1998.
- [PARA00] R. Parajola and J. Rossignac, Compressed Progressive Meshes, *IEEE Transactions on Visualisation and Computer Graphics*, Vol. 6, No. 1, January-March 2000, pp. 79-93.
- [PATE98] S. Pateux and C. Labit, Region-based video coder using the MDL formalism, *IEEE International Conference on Image Processing ICIP'98*, Chicago, USA, 4-7 October, 1998, 4p.
- [PATE00] PATEUX S. Tracking of video objects using a backawrd projection technique, *Proceedings of Visual Communication and Image Processing VCIP'2000*, Perth, Australia. June 2000.
- [PHAM01] Q. C. Pham, F. Vincent, P. Clarysse, P. Croisille, I.E. Magnin, A FEM-Based Deformable Model for the 3D Segmentation and Tracking of the Heart in Cardiac MRI, *Image and Signal Processing and Analysis ISPA 2001*, Pula, Croatia, June 19-21, 2001, pp. 250-254.
- [PORI01] F. M. Porikli, Y. Wang, An unsupervised multi-resolution object extraction algorithm using video-cube, *IEEE International Conference on Image Processing*, Thessaloniki, Greece, October 7-10, 2001, 4p.
- [PREC01] F. Precioso and M. Barlaud, B-Spline Active Contours for Fast Video Segmentation, *IEEE International Conference on Image Processing*, Thessaloniki, Greece, October 7-10, 2001, 4p.

- [ROSS99] J. Rossignac, Edgebreaker, Connectivity compression for triangle meshes, *IEEE Transactions on Visualization and Computer Graphics*, Vol. 5, No. 1, 1999, 16p.
- [REISS97] P.J. Reissman, Modélisation et mise en correspondance par pyramides actives: Application à l'imagerie cardiaque par résonance magnétique, *Thèse de doctorat, Institut National des Sciences Appliquées de Lyon*, No. d'ordre 97ISAL0065, Juillet 1997, 138p.
- [SALE97] P. Salembier, F. Marquès, M. Pardas, R. Morros, I. Corset; S. Jeannin, L. Bouchard, F. Meyer and B. Marcotegui, Segmentation-based Video Coding System Allowing the Manipulation of Objects, *IEEE transactions on Circuits and Systems for Video Technology*, Vol. 7, No. 1, February 1997, pp. 60-73.
- [SALE99] P. Salembier, F. Marques, Region-based representations of image and video: segmentation tools for multimedia services, *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 9, No. 8, December 1999, pp.1147-1169.
- [SHEP68] D. Shepard, A two-dimensionnal interpolation function for irregularly spaced data, *Proceedings of 23<sup>rd</sup> Nat. Conf. ACM*, pp. 517-523.
- [SOUC96] M. Soucy and D. Laurendau, Multiresolution surface modeling based on hierarchical triangulation, *Computer vision and image understanding*, volume 63, No. 1, January 1996, pp. 1-14.
- [SWEL96] W. Sweldens, The Lifting Scheme : A Custom-Design Construction of Biorthogonal Wavelets, *Applied and Computational Harmonic Analysis*, Vol. 3, No. 2, April 1996, pp.186-200.
- [STIL99] C. Stiller, J. Konrad, Estimating Motion in Image Sequences, a Tutorial on Modeling and Computation of 2D Motion, *IEEE Signal Processing Magazine*, July 1999, pp. 70-91.
- [STOL96] E. J. Stollnitz, T. D. Deroose, D. H. Salesin, wavelets for computer graphics : Theory and Application, San Francisco, Morgan Kaufmann Publishers, 1996, p. 245.
- [SULL91] G. J. Sullivan and R. L. Baker, Motion compensation for video compression using control grid interpolation, *Proceedings of ICASSP'91*, Toronto, Canada, May 1991, pp. 2713-2716.
- [TEKA99] A. M. Tekalp, A. Puri, Introduction to multimedia standards, tutorial, *IEEE International Conference on Image Processing ICIP'99*, October 25-28, Kobe, Japan, 1999, 150p.
- [TAUB98a] G. Taubin, J. Rossignac, Geometric compression Through Topological Surgery, *ACM Transaction on graphics*, Vol. 17, No. 2, April 1998, pp. 84-115.
- [TAUB98b] G.Taubin, A. Guézic, W. Horn and F. Lazarus, Progressive Forest Split Compression, *International Conference on Computer Graphics and Interactive Techniques SIGGRAPH '98*, Orlando, Florida, USA, July 1998, pp. 123-132.
- [TOKL98] C. Toklu, Object-based digital video processing using 2-D meshes, PhD. Thesis, University of Rochester, Rochester, USA, School of engineering and applied sciences, 1998, 148p.
- [TOKL00] C. Toklu, A. T. Erdem and A. M. Tekalp, Two-Dimensionnal Mesh-Based Mosaic Representation for Manipulation of Video Objects with Occlusion, *IEEE Transactions on Image Processing*, Vol. 9, No. 9, September 2000, pp. 1617-1630.
- [TOUM98] C. Touma and C. Gotsman, Triangle Mesh Compression, *Graphics Interface '98 Conference Proceedings*, June 1998, pp. 26-34.
- [VALE98] S. Valette, Décomposition en Ondelettes Géométriques de la Surface d'Images 3D, Application à la Compression et à la Transmission Progressive via Internet, *mémoire de DEA, CREATIS, INSA de Lyon*, 1998, p. 37.

- [VALE99a] S. Valette, Y. S. Kim, H. J. Jung, I. E. Magnin and R. Prost, A multiresolution wavelet scheme for irregularly subdivided 3D triangular mesh, *IEEE Int. Conf on Image Processing ICIP'99*, Kobe, Japan, October 25-28, 1999, Vol. 1, pp. 171-174.
- [VALE99b] S. Valette, F. Thibon, Y.-S. Kim, H.-Y. Jung, I. E. Magnin et R. Prost, Décomposition en ondelettes de maillages triangulaires 3D irrégulièrement subdivisés. Application à la compression, *17<sup>e</sup> Colloque GRETSI '99*, Vannes, France, 13-17 Septembre 1999, pp. 949-952.
- [VANB99] P. Van Beek, A. M. Tekalp, N. Zhuang, I. Celasun and M. Xia, Hierarchical 2D mesh representation, tracking and compression for object-based video, *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 9, No. 2, March 1999, special issue, pp. 353-369.
- [VINC00] F. Vincent, P. Clarysse, P. Croisille, I.E. Magnin, An elasticity-based region model an its application to the estimation of the heart deformation in tagged MRI, *IEEE International Conference on Image Processing ICIP 2000*, Vancouver, Canada, September 10-13, 2000, pp. 629-632.
- [WANG94a] J. Y. A. Wang and E. H. Adelson, Representing Moving Images with Layers, *IEEE Transactions on Image Processing*, Special Issue : Image Sequence Compression, vol. 3, no. 5, September 1994, pp. 625-638.
- [WANG94b] Y. Wang and O. Lee, Active mesh – a feature seeking and tracking image sequence representation scheme, *IEEE Transactions on Image Processing*, Vol. 3, September 1994, pp. 610-624.
- [WANG96a] Y. Wang and O. Lee, Use of two-dimensionnal deformable mesh structures for video coding, part I-the synthesis problem : mesh-based function approximation and mapping, *IEEE Transactions on Circuits and Systems for Video Technology*, December 1996, pp. 636-646.
- [WANG96b] Y. Wang, O. Lee and A. Vetro, Use of two-dimensional deformable mesh structures for video coding, part II-the analysis problem and a region-based coder employing an active mesh representation, *IEEE Transactions on Circuits and Systems for Video Technology*, December 1996, pp. 647-659.
- [WANG99] H.-S. Wang and R. M. Merserau, Fast Algorithms for the Estimation of Motion Vectors, *IEEE Transactions on Image Processing*, Vol. 8, No. 3, March 1999, pp. 435-438.
- [WOLB92] G. Wolberg, Digital Image Warping, *Washington, IEEE Computer Society Press*, 1992, pp. 318.
- [YOSH01] T. Yoshida, T. Shimosato, Motion Image Segmentation using 3-D Watershed Algorithm, *IEEE International Conference on Image Processing*, Thessaloniki, Greece, October 7-10, 2001, pp. 773-776.
- [ZHU00] S. Zhu and K.-K. Ma, A New Diamond Search Algorithm for Fast Block-Matching Motion Estimation, *IEEE Trans. on Image Processing*, Vol. 9, No. 2, February 2000, pp. 287-290.
- [ZHON01] D. Zhong and Z.-F. Chang, Long-Term Moving Object Segmentation and tracking Using Spatio-Temporal Consistency, *IEEE International Conference on Image Processing*, Thessaloniki, Greece, October 7-10, 2001, 4p.
- [ZORI97] D. Zorin, P Schröder, W. Sweldens, Interactive Subdivision for Meshes with Arbitrary Topology, *International conference on Computer Graphics and Interactive Techniques, SIGGRAPH '97*, Los Angeles, California, USA, August 1997, pp. 259-268.

## Liste des publications personnelles:

### Conférences internationales

- [VALE99a] S. Valette, Y. S. Kim, H. J. Jung, I. E. Magnin and R. Prost, A multiresolution wavelet scheme for irregularly subdivided 3D triangular mesh, *IEEE Int. Conf on Image Processing ICIP'99*, October 25-28, Kobe, Japan, Vol. 1, pp. 171-174, 1999.
- [VALE01a] S. Valette, Y.-S. Kim and R. Prost, The inverse problem of wavelet scheme construction for irregularly subdivided 3D triangular Meshes, *Quality Control by Artificial Vision QCAV 2001*, Le Creusot, France, may 21-23, 2001, Vol. 2, pp. 452-457.
- [VALE01b] S. Valette, I. E. Magnin and R. Prost, Active mesh for video segmentation and objects tracking, *IEEE Int. Conf. on Image Processing ICIP 2001*, Thessaloniki, Greece, October 7-10, 2001, Vol. II, pp. 77-80.
- [BENO99] H. Benoit-Cattin, P. Joachimsman, A. Planat, S. Valette, A. Baskurt and R. Prost, Active mesh texture coding based on warping and DCT, *IEEE Int. Conf on Image Processing ICIP'99*, October 25-28, Kobe, Japan, Vol. 2, pp 830-833.
- [KIM99] Y.-S. Kim, S. Valette, H.-Y. Jung and R. Prost, Local wavelets decomposition for 3D surfaces, *IEEE Int. Conf on Image Processing ICIP'99*, October 25-28, Kobe, Japan, vol.1, pp. 357-360, October 1999.
- [KIM01] Y.-S. Kim, S. Valette and R. Prost, Adaptive wavelets based multiresolution modeling of irregular meshes via -harmonic maps, *IEEE Int. Conf on Image Processing ICIP 2001*, Thessaloniki, Greece, October , 7-10, 2001, Vol. III, pp. 210-213.

### Conférences nationales

- [VALE99b] S. Valette, F. Thibon, Y.-S. Kim, H.-Y. Jung, I. Magnin et R. Prost, Décomposition en ondelettes de maillages triangulaires 3D irrégulièrement subdivisés. Application à la compression, *17<sup>e</sup> Colloque GRETSI 99*, Vannes, France, 13-17 Septembre 1999, pp. 949-952
- [VALE00] S. Valette, I. E. Magnin et R. Prost, Maillage actif de type quadtree pour la segmentation de séquences vidéo, *journées CORESA 2000*, Poitiers, octobre 2000, pp 197-204.
- [VALE01c] S. Valette, I. E. Magnin et R. Prost, Segmentation et suivi d'objets par maillage déformable : gestion des discontinuités de mouvement, *18<sup>e</sup> Colloque GRETSI 2001*, Toulouse, 10-13 Septembre 2001, Vol. 1, pp. 521-524.
- [VALE01d] S. Valette, Y.-S. Kim et R. Prost, Décomposition multirésolution et compression de surfaces triangulées à l'aide d'ondelettes géométriques, *journée thématique : « Coopération Analyse d'Image et Modélisation »*, LIGIM, Université Claude Bernard Lyon 1, Juin 2001, pp 82-86.



**FOLIO ADMINISTRATIF**

**THESE SOUTENANCE DEVANT L'INSTITUT NATIONAL DES SCIENCES APPLIQUEES DE  
LYON**

<b>Nom :</b> VALETTE	<b>Date de soutenance :</b> 7 janvier 2002
<b>Prénom :</b> Sébastien	
<b>Titre :</b> Modèles de maillages déformables 2D et multirésolution surfaciques 3D sur une base d'ondelettes	
<b>Nature :</b> Doctorat	<b>Numéro d'ordre :</b> 02ISAL0002
<b>Formation doctorale :</b> Electronique, Electrotechnique, Automatique	
<b>Code B.I.U._Lyon :</b>	<b>Classe :</b>
<b>Résumé :</b> Ce travail porte sur le développement de méthodes de traitement d'images à l'aide de maillages. Dans une première partie, nous proposons un nouveau support théorique et un algorithme pour l'analyse multirésolution des surfaces des objets 3D triangulés. Cette approche généralise la décomposition en ondelettes des maillages surfaciques, proposée par Lounsbery, dont la formulation initiale restreint son utilisation aux maillages régulièrement subdivisés. L'extension du domaine d'application de la méthode nécessite la prise en compte d'une subdivision non régulière, permettant l'analyse multirésolution de maillages quelconques par résolution d'un problème inverse. Nous démontrons également l'efficacité de notre approche pour la compression sans perte des maillages surfaciques. La deuxième partie traite de la segmentation et du suivi des objets dans les séquences vidéo 2D+t. Nous proposons un nouvel algorithme de maillage actif qui place les arêtes du maillage sur les discontinuités de mouvement. Le mouvement intérieur à chaque cellule est estimé par mise en compétition des estimations avant et arrière. Cette approche permet le déplacement des nœuds du maillage sans avoir à remailler les zones occultées ou découvertes. Une contrainte spatiale permet une meilleure déformation du maillage dans les régions homogènes. Une deuxième contrainte, de régularisation temporelle, assure la stabilité de l'évolution de la segmentation. Enfin, une conclusion générale ouvre des perspectives sur la représentation des objets vidéo représentés par des maillages, à l'aide d'une adaptation de l'algorithme de décomposition des maillages surfaciques 3D sur une base d'ondelettes.	
<b>Mots-Clés :</b> Maillage, surface, ondelette, mouvement, objets vidéo, compression, segmentation, suivi, multirésolution, 2D+t, 3D	
<b>Laboratoire de recherche :</b> Centre de Recherche Et d'Applications en Traitement des Images et du Signal (CREATIS-UMR CNRS 5515)	
<b>Directeurs de thèse :</b> Rémy PROST et Isabelle MAGNIN	
<b>Président de jury</b> Jean-Marc CHASSERY	
<b>Composition du jury :</b> Michel BARLAUD (rapporteur), Jean-Marc CHASSERY (président), Ho Youl JUNG, Isabelle MAGNIN (codirecteur), Françoise PRETEUX, Rémy PROST (codirecteur)	