



HAL
open science

Vision et Reconstruction 3D basée sur la Texture: Application à la Robotique Médicale.

Aurélien Noce

► **To cite this version:**

Aurélien Noce. Vision et Reconstruction 3D basée sur la Texture: Application à la Robotique Médicale.. Automatique / Robotique. Université Montpellier II - Sciences et Techniques du Languedoc, 2008. Français. NNT: . tel-00339769

HAL Id: tel-00339769

<https://theses.hal.science/tel-00339769>

Submitted on 18 Nov 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ MONTPELLIER II
SCIENCE ET TECHNIQUES DU LANGUEDOC

THÈSE

En vue de l'obtention du grade de

DOCTEUR DE L'UNIVERSITÉ DE MONTPELLIER II

Discipline : **Génie Informatique, Automatique et Traitement du Signal**

Formation Doctorale : **Systèmes Automatiques et Microélectroniques**

Ecole Doctorale : **Information, Structures, Systèmes**

Aurélien NOCE

VISION ET RECONSTRUCTION 3D
BASÉE SUR LA TEXTURE:
APPLICATION À LA ROBOTIQUE MÉDICALE.

Jury

Rapporteurs	Jacques	GANGLOFF	<i>Professeur, Université de Strasbourg</i>
	François	CHAUMETTE	<i>Directeur de Recherche INRIA, IRISA - LAGADIC</i>
Examineur	Claude	LAURGEAU	<i>Professeur, École des Mines de Paris</i>
Co-Directeurs de thèse	Jean	TRIBOULET	<i>Maître de Conférences, Université Montpellier II</i>
	Philippe	POIGNET	<i>Professeur, Université Montpellier II</i>
Directeur de thèse	Etienne	DOMBRE	<i>Directeur de Recherche, CNRS - LIRMM</i>

Remerciements

Je tiens à remercier chaleureusement toutes les personnes qui, par leur présence, leurs conseils et parfois même leur amitié, on fait de ces trois années de thèse des moments inoubliables.

En premier lieu, je remercie mon directeur de thèse Etienne Dombre, qui m'a accueilli dans son équipe et m'a fait confiance pendant trois ans. Je remercie également mes co-cadrants Philippe Poignet et Jean Triboulet qui m'ont suivi au quotidien. Vous m'avez accueilli dans votre équipe et m'avez énormément appris. Plus important peut-être, vous m'avez fait aimer la recherche. Merci.

Je souhaite ensuite remercier M. Claude Lurgeau, M. François Chaumette, et M. Jacques Gangloff d'avoir accepté de d'évaluer mon travail de thèse. Vos travaux ont toujours été une source d'inspiration pour moi et l'obtention de votre aval a été un grand honneur.

Je remercie également M. Michel Robert, directeur du LIRMM, qui m'a accueilli au sein de son laboratoire pendant la durée de ma thèse. Merci au personnel du laboratoire, particulièrement Céline Berger, Nadine et Philippe Tolloy, Elisabeth GreverieTilloy et Cécile Lukasik pour leur aide et leur disponibilité. De manière générale, je remercie tous les membre du LIRMM qui font de cet établissement un endroit idéal pour travailler. Un grand merci également à tous les membres du département Robotique, à commencer par son directeur actuel Philippe Fraisse. Un grand merci a Olivier Strauss, Rene Zapata, Vincent Creuze, Christine Azevedo, Frédéric Comby, Andre Crosnier, David Andreu sans oublier Sébastien Krut et Sébastien Druon (je t'ai connu thésard !) pour leur enthousiasme dans les échanges que nous avons eu durant ma thèse.

Enfin j'aimerai aussi remercier les doctorants roboticiens. Un grand merci à Walid Zarrad (je me rappellerai longtemps les soirées pizzas pendant la rédaction de thèse), Michaël Sauvée, Vincent Nabat, Vincent Bégoc, Gaël Pagès(le roi de la blague) et mes voisins José et Michel. Vous êtes les premières personnes que j'ai connu au laboratoire. Merci également à Arturo, Jean-Mathias, David Crobel (le roi du far breton), Vincent Bonnet (le camarguais du LIRMM), Rogerio (rock and roll man!) et Carla. Merci à la nouvelle génération de thésards : Peter (le collectionneur de Macs), Mourad, Milan, Olivier,

Andréa, et Kévin (un jour je comprendrais ce sur quoi tu travailles).

Je termine en remerciant ma famille. Merci à mes parents, à ma tante Joëlle et à mon oncle Jean-Louis. Merci à ma belle-maman qui m'a accueilli à la fin de ma thèse. Merci du fond du cœur pour votre soutien, je ne serais jamais arrivé là où je suis sans vous tous.

Et bien sûr merci à Camille. Tu es ce que j'ai de plus beau au monde et je te dédie ce mémoire. Je t'aime.

Introduction Générale

Depuis ses débuts, il y a 5 siècles, sous l'égide d'Ambroise Paré, jusqu'aux opérations robotisées actuelles, la chirurgie a su s'imposer comme une pratique fondamentale de la médecine moderne. Initialement pensée comme un geste médical d'urgence pratiqué à la suite d'un échec thérapeutique, elle est devenue une pratique quotidienne pour de nombreux médecins de par le monde, et son champ d'application continue de s'étendre. Encore aujourd'hui, une opération chirurgicale reste une intervention lourde et non exempte de risques.

La chirurgie cardio-vasculaire est la dernière née des spécialités chirurgicales. Opérer le cœur reste très complexe, tant au niveau des risques per-opératoires (saignement, délabrement tissulaire, infection) que des effets secondaires ressentis en phase post-opératoire (douleur, temps de récupération). C'est pourquoi la recherche continue dans ce domaine, et les dernières années ont vu naître des pratiques telles que la chirurgie mini-invasive (historiquement née en 1940 mais qui s'est surtout développée ces 20 dernières années) et les opérations à cœur battant, avec des systèmes dédiés tels que l'**Octopus**. Conjointement au développement de ces solutions, les 25 dernières années ont vu apparaître dans les blocs opératoires des systèmes d'assistance robotisée, tout d'abord dérivés de systèmes industriels dans les années 1980, puis conçus spécifiquement pour la chirurgie, tel le porte endoscope **Aesop** ou les systèmes de chirurgie laparoscopique **Zeus** et **DaVinci**.

Toutes ces solutions commencent à converger. Les enjeux sont clairs :

- d'un côté, les techniques mini-invasives et à cœur battant permettent de réduire les risques d'infection et de séquelles pour le patient, mais obligent le chirurgien à travailler dans un environnement inconfortable. Ainsi les petites incisions pratiquées dans le cadre de la chirurgie mini-invasive limitent la visibilité et privent le chirurgien d'une grande partie de sa mobilité et de ses capacités sensorielles. Les techniques d'opération à cœur battant, quant à elles, imposent l'utilisation d'outils spécialisés tels que des stabilisateurs, qui limitent encore plus la mobilité du chirurgien et qui surtout n'empêchent pas un mouvement résiduel de plusieurs millimètres (alors que le diamètre de l'artère à suturer est d'environ 2mm).
- d'un autre côté, les systèmes robotiques télé-opérés proposent un environnement

plus confortable pour le chirurgien, tout en étant parfaitement adaptés à la chirurgie mini-invasive. Solutions d'avenir, leur champ d'application est restreint par une série de limitations d'ordre technique : manque de retour tactile, dynamique trop faible, difficulté d'opérer un organe en mouvement. De plain pied dans la recherche actuelle autour de l'évolution de ces systèmes, le projet *cœur battant* de l'équipe *Robotique Médicale* du LIRMM étudie la problématique de la compensation des mouvements cardiaques en chirurgie mini-invasive robotisée. Dans le cadre de cette thèse, nous nous sommes intéressés à l'application de techniques de vision par ordinateur, afin d'évaluer les mouvements de l'organe sans contact, dans le but de réaliser un asservissement visuel du robot. Plusieurs travaux ont déjà exploré cette voie, mais tous utilisent des systèmes de marqueurs (généralement des pointeurs laser) afin de mesurer la position de l'organe. L'originalité de notre approche est de nous baser sur l'analyse de texture pour proposer un suivi des mouvements du cœur utilisant une ou deux caméras vidéo.

Nos contributions sont exposées dans le présent document, qui est organisé de la façon suivante :

- **Le premier chapitre** présente le contexte général dans lequel se situent nos travaux. Nous introduirons les principaux détails anatomiques, puis nous présenterons l'acte chirurgical considéré : le pontage coronarien. Nous présenterons ensuite le contexte de la robotique chirurgicale, puis nous nous intéresserons en priorité aux solutions utilisant la vision, afin de positionner nos travaux dans les problématiques actuelles de la communauté. Une étude critique nous permettra de justifier nos choix technologiques, et une première discussion mettra en avant nos attentes vis-à-vis de notre approche initiale et les éventuelles limitations rencontrées.
- **Le deuxième chapitre** traite de l'analyse de texture. Après une introduction aux principales méthodes d'analyse des images texturées, nous proposerons une méthodologie afin de sélectionner un jeu de descripteurs adaptés à nos images expérimentales. Puis nous présenterons les résultats obtenus sur nos images spécifiques, ainsi que sur deux jeux d'images classiques. Les résultats de ce chapitre (les descripteurs sélectionnés) serviront de socles aux approches présentées dans les chapitres 3 et 4.
- **Le troisième chapitre** propose une méthode de suivi de région d'intérêt dans les images de cœur utilisant l'information de texture. Plusieurs points seront discutés, comme le pré-traitement des images, le traitement des reflets et la fusion des données de texture avec la métrique SSD, classique en traitement d'images. Puis nous comparerons notre approche à d'autres, et montrerons qu'elle surpasse celles classiques de suivi de points, et rivalise avec des approches de suivi de motifs déformables, comme la méthode ESM. La discussion finale replacera ces résultats dans le

contexte, les mettant en perspective avec les approches présentées au chapitre 1.

- **Le quatrième chapitre** traite de la mesure 3D des mouvements du cœur. Après un rappel des différentes méthodes pour évaluer un mouvement 3D avec une seule caméra, nous expliquerons pourquoi ces méthodes ne sont pas applicables dans le cas d'un organe déformable tel que le cœur. Puis nous présenterons une approche basée sur la stéréo-vision, qui n'a pas ces limitations.
- **En conclusion** nous résumerons les contributions et terminerons par une présentation des perspectives ouvertes par ces travaux, dans le contexte de la compensation de mouvement en chirurgie mini-invasive robotisée.

Liste des publications

Revue internationale avec comité de lecture :

M. Sauvée, **Aurélien Noce**, P. Poignet, J. Triboulet and E. Dombre. 3D Heart Motion Estimation Using Endoscopic Monocular Vision System : from Artificial Landmarks to Texture Analysis. In *Biomedical Signal Processing & Control in the Special Issue : IFAC Symposia (2008)*, 2(3) : p. 199-207, July 2007.

Conférences internationales avec comité de lecture :

Aurélien Noce, J. Triboulet, P. Poignet and E. Dombre. Features Selection for Visual Servoing of the Beating Heart. In *Proceedings of the 1st IEEE/RAS-EMBS International Conference on Biomedical Robotics and Biomechatronics (BIOROB'06)*, number 183, Pisa, Italie, Feb. 2006.

Aurélien Noce, J. Triboulet, and P. Poignet. Beating Heart Tracking using Composite Texture Algorithm. In *Proceedings of the MICCAI'06 Workshop on Medical Robotics*, p. 54-65 , Copenhagen, Denmark, Oct. 2006.

Aurélien Noce, J. Triboulet and P. Poignet. Efficient Tracking of the Heart Using Texture. In *Proceedings of the 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC'07)*, Lyon, France, Aug. 2007.

Conférences nationales avec comité de lecture :

Aurélien Noce, J. Triboulet and P. Poignet. Suivi de cœur battant : Tracking 2D et 3D. In *12èmes journées d'étude et d'échange COmpression et REprésentation des Signaux Audiovisuels (CORESA'07)*, Actes du colloque, Montpellier, France, Nov. 2007.

Conférences internationales sur abstract :

M. Sauvée, P. Renaud, P. Poignet, J. Triboulet, E. Dombre, M. Karouia, N. Bonnet and **Aurélien Noce**. Mitral Valve Leaflet Motion tracking in Ultrasound Imaging. In *Proceedings of Computer-Aided Medical Interventions : tools and applications (SURGETICA '05)*, Chambéry, France, Jan. 2005.

Aurélien Noce, J. Triboulet and P. Poignet. Suivi de Cœur Battant par Analyse de Texture. In *Proceedings of Computer-Aided Medical Interventions : tools and applications (SURGETICA '07)*, p. 403-407, Chambéry, France, Jan. 2007.

Aurélien Noce, S. Petit, J. Triboulet and P. Poignet. Stereo-Vision for 3D Tracking of the Beating Heart. In *Proceedings of Computer-Aided Medical Interventions : tools and applications (SURGETICA '07)*, p. 63-65, Chambéry, France, Jan. 2007.

Table des matières

1	Contexte	1
1.1	Introduction	1
1.2	Aspects médicaux	2
1.2.1	Considérations anatomiques	2
1.2.2	Pathologies cardiaques courantes et infarctus	8
1.2.3	Chirurgie cardiaque et pontage coronarien	10
1.3	Robotique médicale	15
1.3.1	Genèse de la discipline	15
1.3.2	Avantages recherchés	17
1.3.3	Nature des solutions robotiques actuellement disponibles	17
1.4	Problématique	21
1.4.1	Objectifs médicaux	21
1.4.2	Problèmes scientifiques liés à la compensation de mouvements	21
1.4.3	Problèmes liés à l’aspect vision	23
1.5	Travaux antérieurs	24
1.6	Conclusion	27
2	Texture	29
2.1	Introduction	29
2.2	État de l’art	30
2.2.1	Modèles statistiques	30
2.2.2	Modèles spectraux	33
2.2.3	Modèles structurels	34
2.2.4	Modèle AR/ARMA de texture	35
2.2.5	Modèles par champs de Markov	36
2.2.6	Autres modèles	37
2.2.7	Discussion	38
2.3	Sélection d’attributs	42

2.3.1	Objectifs	42
2.3.2	Génération de la banque d'images-test	42
2.3.3	Étude préliminaire	45
2.3.4	Analyse en composantes principales	45
2.3.5	Analyse discriminante	51
2.4	Discussion	53
2.5	Conclusion	55
3	Suivi	57
3.1	Introduction	57
3.2	Concepts fondamentaux	58
3.2.1	Cadre théorique	58
3.3	État de l'art	61
3.3.1	Suivi de points	61
3.3.2	Suivi de région	64
3.3.3	Autres approches	69
3.4	Approche basée texture	70
3.4.1	Aspects Matériels	71
3.4.2	Suivi composite	73
3.4.3	Améliorations	75
3.5	Évaluation	77
3.5.1	Séquences artificielles	77
3.5.2	Séquences in vivo	83
3.6	Conclusion	86
4	Mouvement 3D	89
4.1	Introduction	89
4.2	Concepts fondamentaux	89
4.2.1	Géométries	89
4.2.2	Espace projectif, coordonnées homogènes et homographies	91
4.2.3	Modèle de caméra	92
4.2.4	Calibrage	93
4.3	Reconstruction 3D	94
4.3.1	Approches mono-caméra	95
4.3.2	Stéréo-vision	97
4.4	Expérimentations	99
4.4.1	Calibrage du système	99
4.4.2	Appariement de Points	104

4.4.3 Reconstruction	105
4.5 Conclusion	107
Bibliographie	111
5 Annexe	123

Table des figures

1.1	Principales causes de décès, dans le monde. On observe la prédominance des infections cardio-vasculaires. Chiffres de 2005 [World Health Organisation, 2005b].	2
1.2	Anatomie du Médiastin. (a) dans le plan sagittal, (b) dans le plan frontal. On observe la situation anatomique du cœur, confiné entre le diaphragme, le sternum et les poumons.	3
1.3	Anatomie du cœur. (source : medecine-et-sante.com, juin 2007)	3
1.4	Artères coronaires. (source : site du projet INRIA Epidaure, oct. 2007)	5
1.5	La révolution cardiaque. On observe la répartition de l'activité musculaire au sein des oreillettes et des ventricules pendant la systole, suivie du remplissage des oreillettes pendant la diastole. (source : e-s-c.fr, juin 2007)	6
1.6	Coupe anatomique du myocarde. (source : foulon.chez-alice.fr, juin 2007)	7
1.7	Interventions sur la valve mitrale. (a) détails anatomiques, (b) différents types de prothèses. On observe en (a) les cordages qui peuvent se rompre et nécessiter une intervention chirurgicale. Plusieurs projets ont été proposés pour une assistance médicale à ce geste, notamment le projet ROBEA GABIE [Sauvée et al., 2005].	9
1.8	Les causes de l'infarctus du myocarde : phénomène d'athérosclérose coronaire suivi de thrombose. (source www.zna.be, juin 2007)	10
1.9	Voies d'abord pour la chirurgie classique du coeur. Sont illustrées ici la sternotomie médiane (1) et la thoracotomie (2), généralement pratiquée au niveau du cinquième espace intercostal.(source : photocopiés en ligne de la faculté de médecine de Lyon 1, sep. 2007)	12
1.10	Les possibilités de pontage. Lorsque la sténose se trouve dans la région de l'artère coronaire droite et de l'artère circonflexe (a) on réalise un pontage à l'aide de segments de veines prélevés sur la jambe. Dans le cas où la sténose se trouve dans la région antérieure des coronaires (b), on réalise un pontage en utilisant deux artères mammaires.(source : www.prevention.ch, oct. 2007)	13

1.11	Stabilisateurs cardiaques pour la chirurgie à coeur battant. L'Octopus 4 de Medtronic (a) fonctionne par succion, tandis que l'Immobilizer de Genzyme (b) fonctionne par pression.	14
1.12	Outils de chirurgie laparoscopique. (a) divers outils pour la chirurgie mini-invasive, (b) endoscopes, (c) outils EndoWrist du robot DaVinci. (sources : www.didema.de et <i>Intuitive Surgical</i> , oct. 2007)	14
1.13	Exemples de robots médicaux. (a) ROBODOC, (b) CASPAR, (c) Zeus, (d) DaVinci, (e) Aesop, (f) BRIGIT, (g) AcuBot, (h) DERMAROB.	19
1.14	Le robot D2M2 développé au LIRMM.	22
1.15	Asservissement Visuel sur coeur battant en x et y [Nakamura et al., 2001].	25
1.16	Méthode de suivi du DLR [Ortmaier, 1999]. On observe les points suivis ainsi que la zone de recherche.	25
1.17	Techniques de vision utilisées au LSIIT. (a) approche utilisant des pointeurs laser [Ginhoux, 2003], (b) Approche par suivi ESM [Bourger et al., 2007]. (Exemple de fenêtre de suivi et de recalage en présence d'occultations).	26
1.18	Patches utilisés pour le suivi de coeur battant dans [Sauvée, 2006].	27
2.1	Construction de la matrice de co-occurrence. Illustration de plusieurs valeurs du vecteur \mathbf{t}	32
2.2	Notation du vecteur \mathbf{t}.	32
2.3	Construction de la matrice de longueur des plages. Illustration pour les principales directions (valeurs de θ).	33
2.4	Représentation structurelle d'une texture. Schéma de principe.	34
2.5	Voisinages considérés dans les modèles ARMA 2D pour l'analyse de texture. (a) Modèle ARMA causal. (b) Modèle ARMA non causal.	36
2.6	Modèle de Markov et HMM. (a) modèle de Markov observable : à chaque état correspond une valeur d'observation. (b) modèle de Markov caché ; chaque état peut émettre une valeur dans l'alphabet des observations.	37
2.7	Extraction de textons. La "structure" est représentée par la superposition de la corrélation entre le texton et l'image et des positions extraites (par seuillage et opération morphologique) pour le centre du texton.	39
2.8	Étapes de la procédure de sélection d'attributs.	43
2.9	Images extraites de nos banques d'images-tests. (a) textures de coeur, (b) textures artificielles, (c) textures de Brodatz, (d) textures Viztex.	44

2.10	Cercle de corrélation. Cercle de corrélation des deux premières composantes principales. On remarque la forte corrélation entre certains des attributs originaux.	47
2.11	Contribution aux axes des 8 meilleurs attributs texturaux, en %.	48
2.12	Détermination du nombre minimum d'attributs à retenir. (a) inertie cumulée portée par les premiers axes principaux (en %), (b) Entropie en fonction du nombre d'attributs.	49
2.13	Résultats de l'analyse discriminante.	52
2.14	Classes utilisées pour l'analyse discriminante. (a) coronaires, (b) surface de cœur.	52
2.15	ACP sur les autres banques d'images-test.	54
3.1	Modèle pour le suivi de mouvement. De l'état réel $e(t)$ à l'état estimé $\hat{e}(t)$, en passant par le modèle d'observation $o(t)$. Cette représentation permet de séparer ce qui est dû à l'acquisition (bruit de mesure, de numérisation) et ce qui est dû à l'estimation (conditionnement, biais).	59
3.2	Principe général des méthodes de suivi de région. On itère en faisant varier les paramètres \mathbf{G} de la transformation jusqu'à ce que l'erreur entre l'image de référence et l'image transformée soit nulle.	65
3.3	Capteur utilisé pour l'acquisition d'images. (a) caméra CMOS DALSA 1m75, (b) courbes caractérisant la relation entre la taille de l'image et la vitesse maximale d'acquisition (au-delà nous rencontrons un problème de "perte" d'images).	72
3.4	Principe du suivi avec la fonction de coût composite.	74
3.5	Sur-résolution par interpolation bi-linéaire.	76
3.6	Principe de création des séquences artificielles.	78
3.7	Exemples de séquences artificielles. (a) séquence simple, translation, (b) séquence texture sur texture, (c) séquence avec déformation du motif. La zone de recherche est définie centrée sur la position précédente avec une taille de trois fois les dimensions du motif recherché.	79
3.8	Résultats sur un exemple de translation. (a) exemple de trajectoire, avec illustration de l'erreur du suivi KLT (flot optique), (b) projection de l'erreur suivant l'axe vertical, (c) projection de l'erreur suivant l'axe horizontal. Cette série montre surtout la faiblesse de l'approche par flot optique en présence de réflexions spéculaires et devant la forte homogénéité de l'image. Elle illustre également l'intérêt des séries d'images générées artificiellement, qui permettent d'analyser et de quantifier facilement l'erreur induite par les différentes méthodes.	81

3.9	Un autre exemple de suivi. (a) trajectoires correspondant au mouvement, (b) extrait de la séquence correspondante. Dans cet exemple, on ajoute au mouvement du motif des déformations homographiques. Ce type de déformation fonctionnant parfaitement avec ESM (qui est conçu dans ce sens), nous n'avons représenté ici que les méthodes de suivi de point. Sur cet exemple particulier, on remarque facilement l'intérêt de l'approche par texture, qui améliore la robustesse du suivi, là où d'autres méthodes perdent la cible.	82
3.10	Images extraites de séquences expérimentales prises à des distances, résolutions et avec des éclairages différents. (a) séquence 512×512 pixels, éclairage par scialitique chirurgical, distance : 70cm de l'organe, (b) séquence 128×128 pixels (à 250Hz), éclairage par LED, distance : 25cm de l'organe, (c) séquence 512×512 pixels, éclairage hallogène, distance : 10cm de l'organe.	84
3.11	Exemple de suivi Composite sur une séquence in vivo. Le carré bleu représente le motif suivi en entier.	85
3.12	Effet des déformations du motif sur le suivi. (a) comparaison du suivi SSD et du suivi Composite : l'approche par SSD, en vert, a tendance à dévier du motif original au cours de la déformation, contrairement au suivi Composite, en bleu. (b) l'algorithme ESM a tendance à perdre la cible et diverge alors rapidement.	85
3.13	Trajectoires obtenues en suivi dans l'image. Les trajectoires verticales (trait continu) et horizontales (tirets) sont exprimées en fonction de l'indice de l'image dans la séquence.	86
3.14	Suivi de mouvements de cœur battant, tiré de [Nakamura et al., 2001].	87
4.1	Modèle de caméra et changements de repères associés.	94
4.2	Ambiguïté sur la profondeur de l'objet induite par la projection.	95
4.3	Géométrie d'un capteur stéréoscopique.	97
4.4	Contrainte épipolaire et matrice fondamentale F.	98
4.5	Extraits d'une séquence de calibrage in vivo. La mire de 5cm de côté spécialement réalisée pour cette expérimentation, est placée dans le champ du capteur en début d'expérimentation.	100
4.6	Exemple de mesure de précision en utilisant la mire de calibrage. Pour des images 512×512 , sur cette série, on obtient une précision de reconstruction de l'ordre de 0,5mm.	101

4.7	Points considérés par RANSAC comme acceptables pour l'estimation , image 128×128 , tolérance 1,5mm. Les points en verts sont conservés pour l'estimation (inliers). Les points en rouge sont rejetés par RANSAC (outliers). La plupart des rejets sont dus à des échecs du détecteur de coins.	104
4.8	Les deux approches pour la reconstruction stéréoscopique . Les étapes en sur-brillance sont celles qui sont les plus demandeuses en puissance de calcul.	104
4.9	Calcul des correspondances en utilisant la contrainte épipolaire . Exemple d'appariements de points de la mire sur l'implémentation OpenCV. Nous traçons dans l'image de droite la ligne épipolaire correspondant au point sélectionné dans l'image de gauche, et représentons en bas les pixels des régions d'intérêt considérées.	105
4.10	Erreur de re-projection . L'erreur de re-projection est la moyenne géométrique des distances entre les points relevés dans les images et la re-projection du point estimé dans ces images.	106
4.11	Exemple de reconstruction stéréoscopique . (a) images de la caméra de gauche, (b) position 3D reconstruites, (c) images de la caméra de droite.	108

Liste des tableaux

2.1	<i>Vue d'ensemble des modèles de texture.</i>	30
2.2	<i>Test KS de normalité des variables : valeur de α.</i>	45
2.3	<i>Jeu d'attributs sélectionnés par ACP. Les deux dernières étapes de sélection des attributs (jeu de 12 descripteurs puis jeu final de 8 descripteurs) sont représentés. (voir l'annexe pour la description des attributs).</i>	50
2.4	<i>Jeu d'attributs sélectionnés par AD. Sont représentés en italique les attributs communs des deux approches.</i>	53
2.5	<i>Autres attributs sélectionnés par ACP.</i>	54
2.6	<i>Jeu d'attributs sélectionnés.</i>	55
3.1	<i>Principaux filtres bayésiens.</i>	60
3.2	<i>Précision du suivi pour l'exemple présenté figure 3.8 et pour le mouvement par homographie de la figure 3.7c. Les cases marquées "X" correspondent aux cas où le suivi n'est plus possible car l'algorithme perd complètement la cible.</i>	80
4.1	<i>Hiérarchie des transformations géométriques dans le plan.</i>	90
4.2	<i>Évolution de la précision selon la résolution des images sur le test de planéité.</i>	102

Chapitre 1

Contexte des travaux

1.1 Introduction

Les maladies cardio-vasculaires sont la première cause de mortalité dans le monde, devant le cancer et le sida. Elles provoquent à elles seules environ 12 millions de décès par an (figure 1.1), parmi lesquels on compte plus de 43% d'affections coronariennes et 32% d'accidents vasculaires cérébraux [World Health Organisation, 2005b]. Ce phénomène, qui devrait encore s'accroître à l'échelle mondiale dans les prochaines années, a un coût très élevé. Ainsi, on estime que les pathologies cardiaques ont coûté 169 milliards d'euros à l'Union Européenne en 2003 [Tendera, 2006] en soins médicaux et en perte de productivité. Devant la gravité et le caractère endémique du phénomène, des actions de sensibilisation sont entreprises afin de limiter les facteurs de risque [World Health Organisation, 2005a, Mackay and Mensah, 2005] tels que le tabac, l'obésité ou le diabète. Cependant, la prévention n'est pas suffisante et la recherche pour de nouveaux traitements pour ces affections est aujourd'hui un enjeu scientifique majeur. L'Organisation Mondiale de la Santé et ses partenaires internationaux sont conscients des enjeux humains et économiques posés par le traitement des affections cardiaques et ont pris des positions claires à ce sujet [Cianflone et al., 2004].

Dans ce chapitre introductif, nous présentons la problématique de la chirurgie cardiaque, ses enjeux et ses principales méthodes. Puis nous présentons le contexte de nos travaux au sein de la communauté de la robotique médicale. Nous explicitons les besoins en termes de techniques d'imagerie et de vision pour ces applications, et situons nos réalisations par rapport à l'état de l'art. Ce chapitre s'articule de la façon suivante : nous introduisons d'abord les aspects médico-physiologiques inhérents à notre application, de l'anatomie cardiaque au §1.2.1, aux interventions de pontage au §1.2.3 ; puis nous présentons la robotique médicale au §1.3 et plus particulièrement les aspects concernant la compensation des mouvements physiologiques et la vision (§1.4).

Projections du nombre de décès dans le monde, par cause, tous âges confondus, 2005

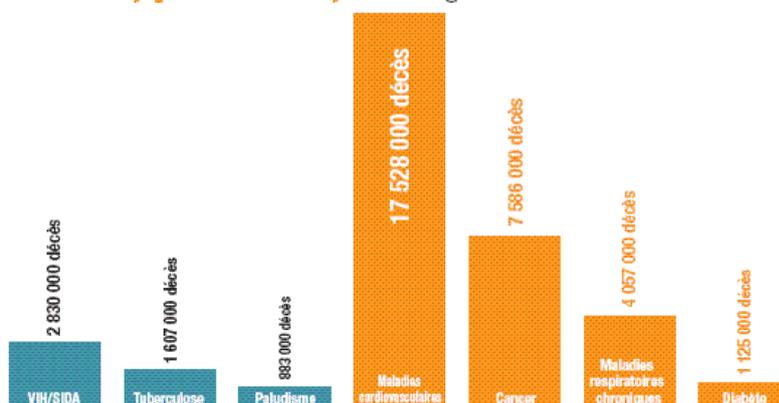


FIG. 1.1 – *Principales causes de décès, dans le monde.* On observe la prédominance des infections cardio-vasculaires. Chiffres de 2005 [World Health Organisation, 2005b].

1.2 Aspects médicaux

1.2.1 Considérations anatomiques

Le cœur est un organe musculaire creux d'un poids moyen de 270 grammes¹. Organe contractile, il contrôle la circulation sanguine et joue un rôle essentiel dans le réglage de son débit. Il est disposé au centre du thorax² où il repose sur le diaphragme. Il a la forme d'une pyramide triangulaire dont l'*apex* est orienté vers le bas à gauche. Chez l'être humain, il est situé dans la partie antérieure du médiastin³, où il est maintenu en place par ses connexions aux grandes artères et aux veines, et son confinement dans le péricarde (une enveloppe de tissus cellulaires à double paroi, dont l'une entoure le cœur et l'autre est rattachée au sternum, au diaphragme et aux membranes du thorax). La figure 1.2 présente l'anatomie du médiastin par une coupe dans le plan sagittal⁴ et une coupe dans le plan frontal⁵, et illustre le positionnement relatif du cœur.

1.2.1.1 Les cavités cardiaques

Le cœur est constitué de quatre cavités : les deux *oreillettes* (ou *atria*) et les deux *ventricules*, comme illustré sur la figure 1.3. Sur la surface du cœur, des sillons superficiels

¹ en moyenne 300g chez l'homme et 250 chez la femme.

² bien qu'on ait tendance à penser qu'il est à gauche.

³ le *médiastin* est la région de la cage thoracique située entre les deux poumons contenant le cœur, l'oesophage, la trachée et les deux bronches souches, voir figure 1.2.

⁴ ou plan médian, il sépare les parties droite et gauche du corps.

⁵ ou plan coronal, il sépare les parties avant et arrière du corps.

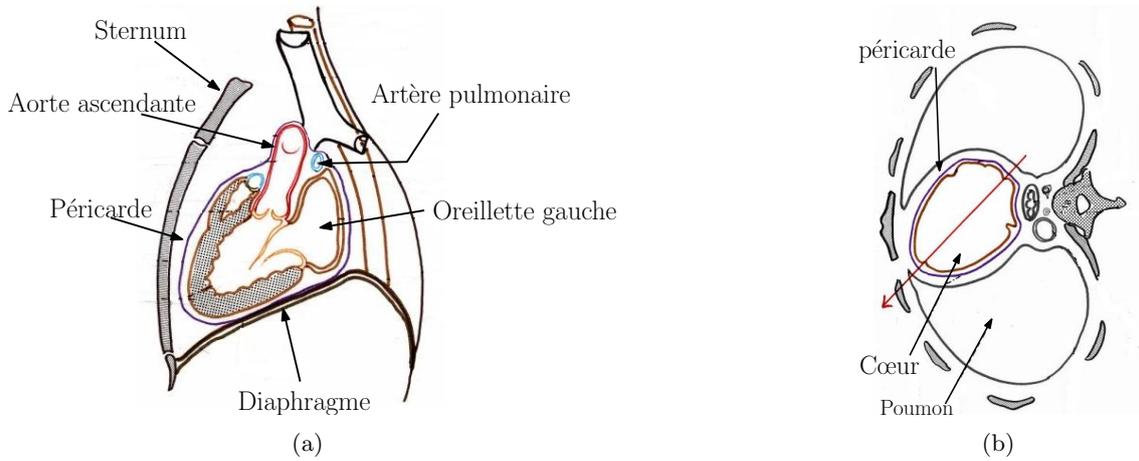


FIG. 1.2 – *Anatomie du Médiastin.* (a) dans le plan sagittal, (b) dans le plan frontal. On observe la situation anatomique du cœur, confiné entre le diaphragme, le sternum et les poumons.

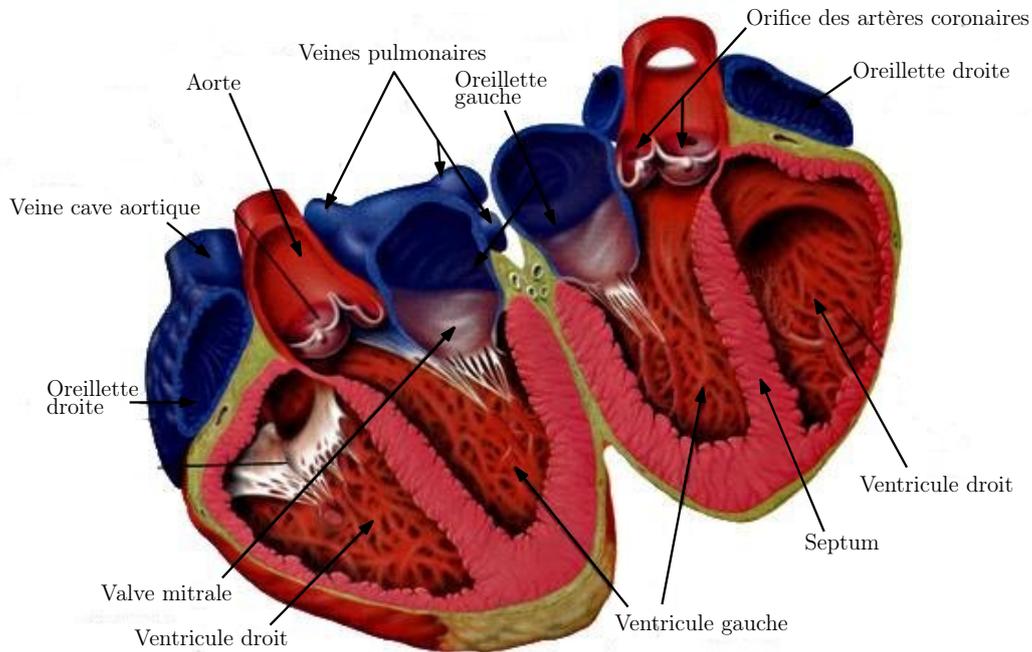


FIG. 1.3 – *Anatomie du cœur.* (source : medecine-et-sante.com, juin 2007)

délimitent les oreillettes et les ventricules. Un *sillon axial* se situe entre les oreillettes en arrière et les ventricules en avant. Un *sillon transversal* perpendiculaire à l'axe du cœur, appelé *sillon auriculo-ventriculaire*, est situé entre oreillettes et ventricules. Ces sillons sont les lieux de passage des vaisseaux nourriciers du cœur, les *artères coronaires* droite et gauche. La face antérieure de l'organe est la *face chirurgicale*, que l'on aborde en fendant le sternum.

Oreillettes et ventricules sont séparés deux à deux par une épaisse paroi musculaire longitudinale, le *septum* qui garantit l'étanchéité entre la partie droite et la partie gauche du cœur. Cette séparation est autant anatomique que fonctionnelle :

- **La partie droite** du cœur assure l'expulsion du sang pauvre en oxygène vers le circuit pulmonaire. L'oreillette droite reçoit à son pôle supérieur la veine cave supérieure, et à son pôle inférieur la veine cave inférieure par l'intermédiaire desquelles arrive le sang. Celui-ci passe ensuite dans le ventricule droit qui le propulse en direction des poumons.
- **La partie gauche** du cœur assure la circulation systémique, c'est-à-dire l'approvisionnement en oxygène de l'organisme. Le sang oxygéné par les capillaires pulmonaires passe dans les veines pulmonaires, qui se vident dans l'oreillette gauche. La contraction du ventricule gauche propulse le sang vers l'aorte, qui le diffuse vers toutes les artères de l'organisme.

1.2.1.2 Le réseau artériel et veineux

Les veines du cœur rejoignent, en presque totalité, un confluent commun, le *sinus coronaire*. Il se déverse dans l'oreillette droite. L'*aorte*, qui part, du ventricule gauche du cœur, est la principale artère du corps, chez l'homme. Elle constitue le tronc d'origine des autres artères. On la divise, à partir du cœur, en trois segments : la *croisse de l'aorte*, l'*aorte thoracique descendante* et l'*aorte abdominale* qui se ramifient pour irriguer la périphérie de l'organisme.

Deux artères coronaires, droite et gauche, naissent de la croisse de l'aorte, immédiatement au dessus des valvules sigmoïdes⁶ de ce vaisseau. Les artères coronaires alimentent en oxygène le muscle cardiaque. Elles sont terminales, c'est-à-dire que chacune d'entre elles assure la nutrition d'un territoire précis, sans possibilité de suppléance réciproque. Les artères coronaires sont représentées sur la figure 1.4.

⁶valves aortiques et valves pulmonaires.

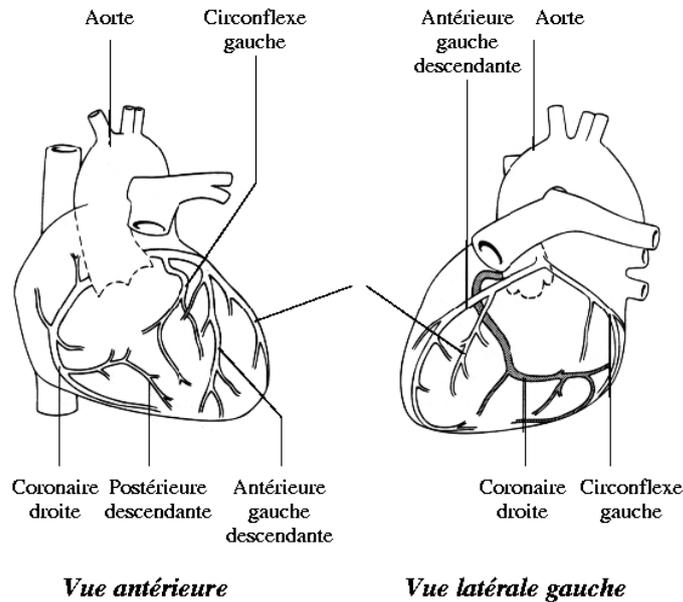


FIG. 1.4 – *Artères coronaires.* (source : site du projet INRIA Epidauré, oct. 2007)

1.2.1.3 La révolution cardiaque

Le cœur a une fréquence d'impulsion au repos de 50 (chez le sportif) à 80 battements par minute. Celle-ci s'élève au cours d'un effort jusqu'à un rythme théorique donné par l'heuristique :

$$f_{\text{théorique}} = 220 - \text{age du patient.} \quad (1.1)$$

Chacun de ces battements du cœur entraîne une séquence d'événements dont la succession est appelée révolution cardiaque. On distingue, au cours de celle-ci, trois étapes, représentées sur le figure 1.5 : la *systole auriculaire*, la *systole ventriculaire* et la *diastole*.

- **La systole auriculaire** est une phase de remplissage actif. La valve de l'aorte est fermée et la pression du sang dans les oreillettes, supérieure à celle du ventricule, entraîne l'ouverture des valves auriculo-ventriculaires. Les oreillettes se contractent et éjectent du sang vers les ventricules. Une fois le sang expulsé, les valves se ferment et évitent ainsi que le sang ne reflue vers les oreillettes.
- **La systole ventriculaire** décrit la contraction des ventricules. La paroi des ventricules, clos par les valvules sigmoïdes, développe une tension qui modifie la forme des ventricules. Le sang étant incompressible, la pression augmente rapidement jusqu'à dépasser celle de l'aorte, déclenchant l'ouverture de la valve pulmonaire à droite et de la valve aortique à gauche. La diminution de pression induite par ce transfert

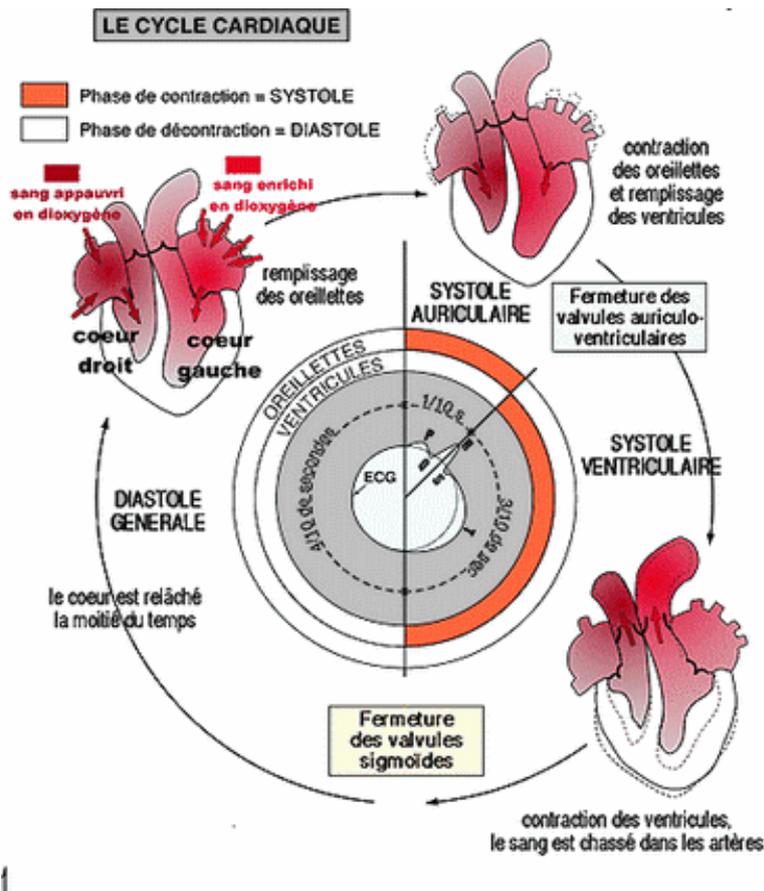


FIG. 1.5 – *La révolution cardiaque.* On observe la répartition de l'activité musculaire au sein des oreillettes et des ventricules pendant la systole, suivie du remplissage des oreillettes pendant la diastole. (source : e-s-c.fr, juin 2007)

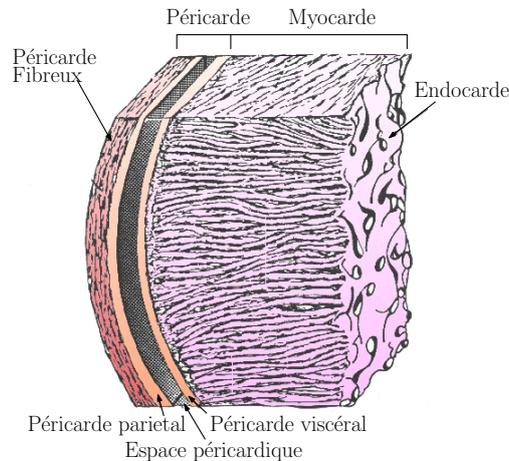


FIG. 1.6 – *Coupe anatomique du myocarde.* (source : foulon.chez-alice.fr, juin 2007)

referme les valvules sigmoïdes. En même temps les oreillettes se remplissent de sang.

- **La diastole** est la relaxation de toutes les parties du cœur, permettant le remplissage passif des ventricules et des oreillettes, et l'arrivée de nouveau sang.

1.2.1.4 Le myocarde

Les parois du cœur sont essentiellement formées par une tunique musculaire contractile : le *myocarde*, représenté sur la figure 1.6. C'est le muscle cardiaque qui donne la puissance nécessaire à la contraction de l'organe.

Il est constitué de deux membranes. A l'intérieur, une membrane continue, extrêmement mince, l'*endocardie* tapisse la cavité cardiaque afin d'éviter la coagulation du sang au contact de la paroi du cœur. À la périphérie, le myocarde est recouvert du *péricarde*, formé d'une gaine fibreuse superficielle et d'une gaine séreuse à deux feuillets. Ces deux couches glissent l'une sur l'autre librement grâce à la production d'un liquide lubrifiant.

La couche inférieure du feuillet séreux, aussi appelée *épicaarde*, constitue la surface du cœur et est en contact direct avec le myocarde situé en dessous. Le myocarde est formé de fibres musculaires striées, anastomosées en réseau. Ce tissu de fibres conjonctives⁷ forme la charpente du cœur. Son épaisseur varie considérablement : mince au niveau des oreillettes, elle augmente autour des ventricules, jusqu'à dépasser un centimètre au niveau du ventricule gauche. Cette variation obéit à une double dissymétrie : entre les oreillettes et les ventricules, dont les parois sont plus épaisses car la fonction de pompage remplie par ces dernières nécessite plus de puissance musculaire et entre la partie droite et la partie gauche du cœur, plus massique car la pression dans le circuit systémique est cinq fois plus importante que celle du circuit pulmonaire.

⁷dont la contraction échappe à toute volonté, faisant ainsi exception à la règle pour les muscles striés.

On distingue également dans le myocarde le tissu nodal, pauvre en myofibrilles, qui stimule et entretient les battements cardiaques.

1.2.2 Pathologies cardiaques courantes et infarctus

1.2.2.1 Atteintes du rythme cardiaque

Une partie importante des pathologies cardiaques concerne des dysfonctionnements au niveau du rythme cardiaque. Les causes peuvent être de nature très variée : génétiques, comme dans le cas du syndrome d'Andersen, liées à l'absorption de produits tels que l'alcool ou le tabac, liées aux séquelles d'un infarctus ou encore mécaniques après une intervention chirurgicale. D'autres causes peuvent être hormonales, comme chez les patients atteints d'hyperthyroïdie.

La forme la plus courante de ces atteintes du rythme cardiaque est la *fibrillation*, qui se manifeste par des contractions rapides et irrégulières des ventricules (ou des oreillettes dans le cas de la fibrillation auriculaire). C'est le plus fréquent des troubles du rythme cardiaque. C'est une infection grave : le risque de mortalité cardio-vasculaire est multiplié par deux chez les patients qui en sont atteints. Les risques d'accidents vasculaires cérébraux sont aussi augmentés. Le traitement est généralement médicamenteux, mais une intervention chirurgicale peut être nécessaire dans certains cas. Le cas échéant, deux interventions sont envisageables : soit l'ablation de la zone responsable du trouble (lorsque les "foyers d'activation" sont localisables) par radio-fréquence, soit la pose d'un stimulateur cardiaque définitif, un *pace-maker* qui va assurer une stimulation régulière⁸ de la cavité responsable du trouble.

1.2.2.2 Les valvulopathies

Les valves cardiaques empêchent le reflux du sang d'une cavité cardiaque vers une autre pendant le cycle cardiaque (§1.2.1.3). Le dysfonctionnement de ces valves entraîne une diminution de l'efficacité de la pompe cardiaque. Généralement situées au niveau de la partie gauche du cœur, ces affections touchent aussi bien la valve mitrale, entre l'oreillette et le ventricule, qu'aortique, entre le ventricule et l'aorte. Parmi les causes possibles, on trouve des affections microbiennes (souvent associées à une affection de l'endocarde), des causes inflammatoires, les séquelles d'un infarctus, ou encore des causes mécaniques telles que la rupture des cordages de la valve (figure 1.7a).

Lorsque les lésions le permettent, il est possible de réparer la valve par chirurgie. Souvent, lors d'une insuffisance valvulaire due à une rupture de cordages, on reconstitue

⁸depuis quelques années, on trouve également des pace-makers, qualifiés de "sentinelles", qui stimulent uniquement l'organe en cas de trouble.

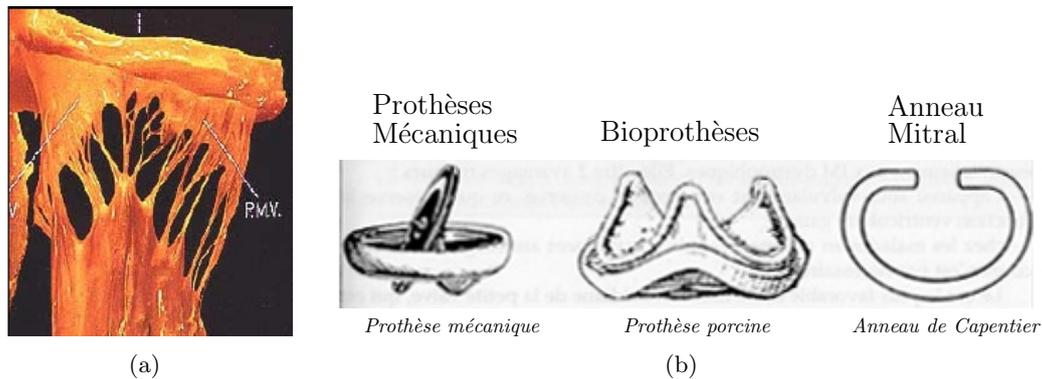


FIG. 1.7 – **Interventions sur la valve mitrale.** (a) détails anatomiques, (b) différents types de prothèses. On observe en (a) les cordages qui peuvent se rompre et nécessiter une intervention chirurgicale. Plusieurs projets ont été proposés pour une assistance médicale à ce geste, notamment le projet ROBEA GABIE [Sauvée et al., 2005].

les cordages reliant le feuillet de la valve à la paroi du cœur, puis si nécessaire on modifie la forme de la valve par la pose d'un anneau (figure 1.7b). Dans d'autres cas, le chirurgien peut modifier la forme de la valve par résection⁹, ou encore, lorsque la valve est perforée, par l'ajout d'un patch, généralement réalisé avec un bout de péricarde. Lorsque l'état du cœur ne permet pas ces interventions, il est possible de pratiquer un remplacement complet de la valve par une prothèse mécanique ou biologique. Ces différentes interventions, lourdes, imposent le passage en *circulation extra corporelle* (ou *CEC*).

Nous avons eu l'occasion de participer, au début de cette thèse, au projet ROBEA GABIE visant à développer une plate-forme d'assistance robotisée pour le remplacement des cordages de la valves mitrales, sous guidage échographique [Sauvée et al., 2005]. Ces travaux ont depuis été présentés dans plusieurs thèses [Sauvée, 2006] et ont fait l'objet d'expérimentations in-vivo.

1.2.2.3 L'infarctus du myocarde

L'*infarctus du myocarde*, couramment appelé crise cardiaque, est une nécrose ischémique du myocarde, c'est-à-dire la destruction d'une partie du tissu cellulaire du muscle cardiaque par défaut prolongé d'apport sanguin. Celle-ci fait place à une cicatrice fibreuse ayant perdu toute faculté contractile. La destruction d'une partie du myocarde fait suite à l'obstruction d'une branche d'une artère coronaire, qui freine donc l'apport de sang et donc d'oxygène.

Le plus souvent, l'infarctus est induit par une athérosclérose coronaire, c'est-à-dire un "encrassement" artériel lent, formé par un dépôt de plaques d'athérome (ou plaques

⁹c'est-à-dire par ablation d'une partie de la valve pour modifier sa forme.

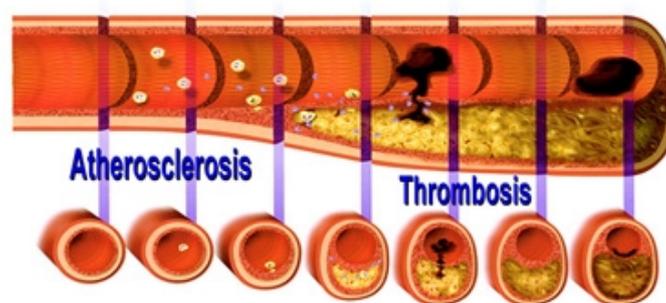


FIG. 1.8 – *Les causes de l'infarctus du myocarde : phénomène d'athérosclérose coronaire suivi de thrombose.* (source www.zna.be, juin 2007)

artérielles) riches en lipides tels que le cholestérol, qui épaississent la paroi des artères coronaires. Ce phénomène, illustré sur la figure 1.8, provoque d'abord une diminution du calibre artériel puis un ralentissement de la circulation artérielle locale, c'est l'*ischémie*. Sur ces lésions à risque peut alors survenir une thrombose : un caillot de sang se forme dans le vaisseau et complète brutalement l'obstruction. L'effet délétère correspondant à cette insuffisance circulatoire est l'*anoxie* (c'est-à-dire le défaut d'apport en oxygène au tissu cellulaire), puis très rapidement la destruction du tissu. C'est l'infarctus proprement dit.

Puisqu'il existe trois artères coronaires essentielles (inter-ventriculaire antérieure, circonflexe et coronaire droite) qui se divisent chacune en de multiples branches d'importance inégale, on peut observer plusieurs types d'infarctus, selon les artères touchées et le niveau de souffrance cellulaire. Toutefois, en règle générale, c'est le ventricule gauche qui est atteint.

Les complications de l'infarctus sont souvent redoutables. Les troubles du rythme cardiaque peuvent engendrer une insuffisance cardiaque, c'est-à-dire une incompetence de la pompe cardiaque lorsqu'une partie importante de son capital contractile est atteinte, affectant à jamais le fonctionnement de l'ensemble de l'organisme. Plus radicalement, une conséquence commune de l'infarctus est la mort subite du patient, lorsqu'une trop grande partie du muscle cardiaque est détruite.

1.2.3 Chirurgie cardiaque et pontage coronarien

1.2.3.1 Introduction à la chirurgie cardiaque

Les opérations du cœur sont les dernières nées des spécialités chirurgicales. Le cœur est l'organe qui entretient la vie et symbolise celle-ci, par ses battements qui assurent la circulation du sang. Intervenir sur le cœur, c'était risquer de le voir s'arrêter, et de

provoquer la mort en quelques minutes. Ce n'est donc qu'à partir de 1938, année qui vit se réaliser, avec succès, la première opération d'un canal artériel par le chirurgien H. Gross à Boston, que la chirurgie cardiaque pris son essor. Cinq ans plus tard, en 1943, le chirurgien américain Arthur Blalock réalisait le premier "shunt" entre l'aorte et l'artère pulmonaire, pour soulager des enfants atteints de maladies cyanogènes. Enfin, en 1953, A. Gibbon mettait au point la première machine de CEC, ouvrant la voie à la chirurgie cardiaque moderne.

Depuis, les interventions cardiaques se sont multipliées et des opérations lourdes telles que le pontage coronarien se sont banalisées. Cependant, les risques initiaux persistent, et trois facteurs principaux sont à prendre en compte lorsque l'on décide d'intervenir :

- tout d'abord la position centrale de l'organe dans la cage thoracique, qui nécessite, pour l'aborder, d'ouvrir l'une des deux hémicavités thoraciques. Cette opération provoque immédiatement un affaissement du poumon et une asphyxie partielle, qui est compensée par les méthodes classiques d'anesthésie, notamment grâce à l'intubation trachéale. La voie d'abord la plus habituelle est la section verticale du sternum (la sternotomie¹⁰) car le coeur se trouve immédiatement en arrière de cet os. Des sternotomies partielles sont parfois possibles. Elles permettent, en réduisant l'incision, d'éviter l'ouverture des deux plèvres et l'affaissement du poumon. Cependant la ventilation assistée reste nécessaire.
- le deuxième obstacle est l'abondance de l'hémorragie qui risque de survenir lorsqu'on opère le coeur et ses abords. Des transfusions sanguines, parfois extrêmement importantes sont nécessaires pour compenser ce risque.
- le troisième obstacle est la crainte de voir le coeur mal supporter l'agression opératoire et s'arrêter. En cas de traumatisme opératoire trop intense, les fibres musculaires cardiaques, au lieu de se contracter simultanément et rythmiquement, entrent en contraction séparément sans ordre défini : cette *fibrillation* rend la contraction cardiaque inefficace, mais peut être neutralisée par l'application d'un choc électrique par un défibrillateur.

Dans les paragraphes suivants, nous allons étudier les différentes approches de la chirurgie cardiaque, puis nous insisterons sur les avantages des nouvelles méthodes laparoscopiques et à coeur battant vis-à-vis de ces problématiques.

¹⁰on peut également pratiquer une thoracotomie, mais c'est une pratique rare vu les avantages qu'offre la sternotomie (accès aisé au coeur, à l'aorte ascendante et transversale, etc.). La sternotomie reste la principale voie d'abord pour la chirurgie cardiaque.

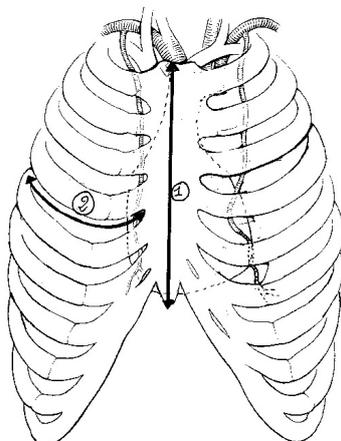


FIG. 1.9 – *Voies d’abord pour la chirurgie classique du cœur.* Sont illustrées ici la sternotomie médiane (1) et la thoracotomie (2), généralement pratiquée au niveau du cinquième espace intercostal. (source : photocopiés en ligne de la faculté de médecine de Lyon 1, sep. 2007)

1.2.3.2 L’opération de pontage coronarien

- **Approche classique :** Après une anesthésie et la mise du patient sous respiration artificielle, on pratique sur le patient une sternotomie ou une thoracotomie (voir figure 1.9). Puis on procède au prélèvement des artères mammaires nécessaires aux pontages (à raison d’un prélèvement par pontage). L’utilisation de veines saphènes¹¹ est également envisageable.

Le chirurgien incise ensuite le péricarde pour dégager le cœur, la CEC est mise en place et le cœur est stoppé.

Le chirurgien procède alors au pontage : il incise la coronaire en aval de la sténose, puis anastomose (suture) l’artère mammaire sur celle-ci. Deux cas sont possibles pour réaliser un double pontage sur la face avant du cœur, comme illustré sur la figure 1.10. La suture dure généralement de 10 à 20 minutes. Le chirurgien répète cette opération autant de fois qu’il y a de pontages à réaliser (en général il faut plus de deux pontages par intervention). Lorsque cela est nécessaire, il peut réaliser une “luxation” du cœur : le cœur est “retourné” afin d’intervenir sur sa face postérieure. Une fois les pontages réalisés, il faut remettre en fonction le cœur, en vérifiant la bonne étanchéité des sutures, puis refermer le thorax du patient.

L’opération en elle-même dure de 3 à 6 heures, mais est suivie d’une surveillance du patient dans une unité de soins intensifs pendant 24 à 72 heures.

¹¹veines superficielles de la jambe.

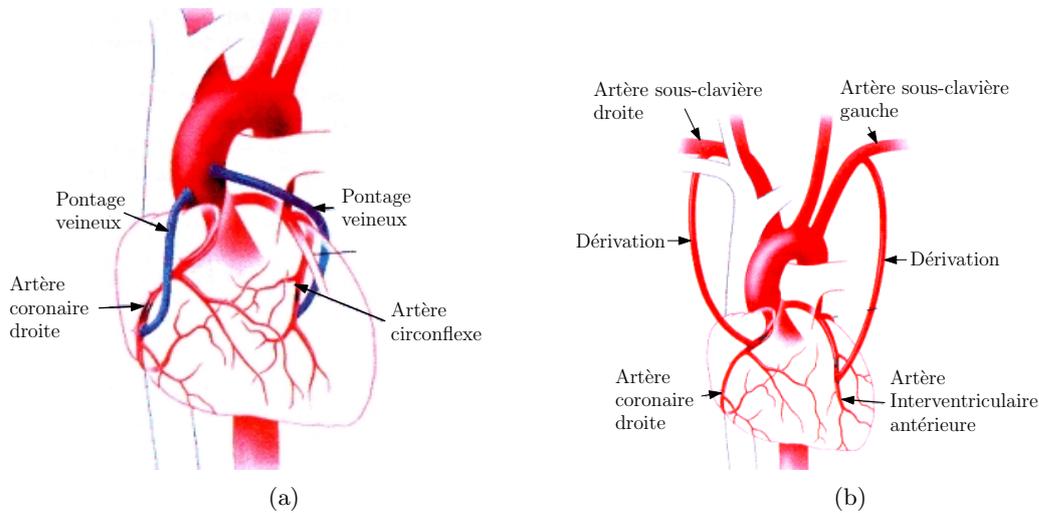


FIG. 1.10 – **Les possibilités de pontage.** Lorsque la sténose se trouve dans la région de l'artère coronaire droite et de l'artère circonflexe (a) on réalise un pontage à l'aide de segments de veines prélevés sur la jambe. Dans le cas où la sténose se trouve dans la région antérieure des coronaires (b), on réalise un pontage en utilisant deux artères mammaires. (source : www.prevention.ch, oct. 2007)

- **Approche cœur battant :** pour opérer sur cœur battant, on peut utiliser un stabilisateur mécanique tels que l'*Octopus* de Medtronic (figure 1.11a) ou l'*Immobilizer* de Genzyme [Lemma et al., 2005] (figure 1.11b). Il permet d'isoler un "champ" à la surface du muscle cardiaque dont les déplacements seront très restreints. Si cette zone stabilisée est judicieusement placée (sur le trajet de l'artère coronaire, en aval de la sténose), il devient possible d'opérer la zone en question sans arrêter le cœur.

L'opération se poursuit alors normalement, à ceci près qu'il faut clamer l'artère coronaire en amont de l'intervention afin d'éviter une perte de sang. Il est même envisageable de verticaliser le patient afin de sortir l'apex et d'accéder à la face postérieure du cœur. Un des principaux intérêts de cette méthode est le temps de réanimation très court.

Les principaux obstacles sont des risques accrus par rapport à une CEC classique. On notera aussi la plus grande technicité de l'intervention. Enfin, les effets du stabilisateur mécanique sur l'organe n'ont pas encore été évalués de façon exhaustive.

- **Approche mini-invasive :** la chirurgie mini-invasive¹², ou chirurgie laparoscopique, est une forme de chirurgie récente qui répond aux demandes actuelles de confort et d'esthétique des patients. Elle substitue aux outils chirurgicaux classiques, qui nécessitent un champ opératoire assez large pour effectuer les gestes chirurgicaux, des outils spécialisés

¹²en anglais MIS pour *Minimally Invasive Surgery*.

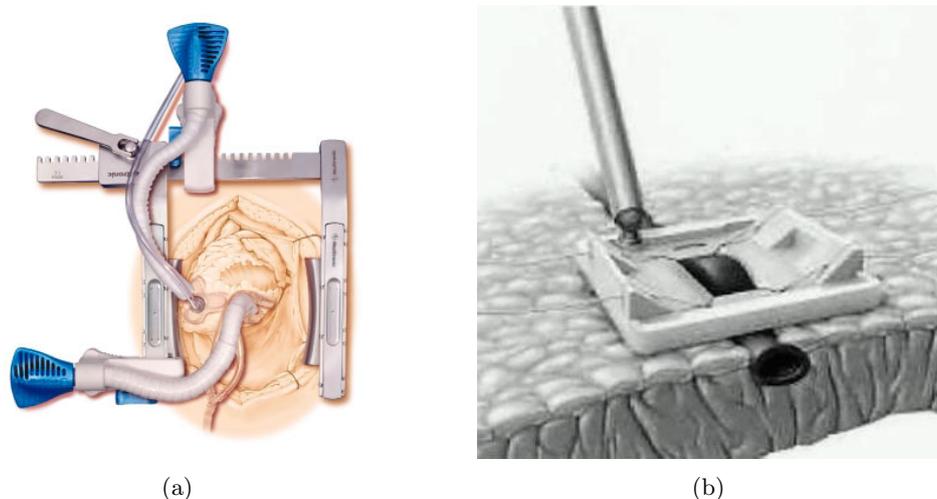


FIG. 1.11 – *Stabilisateurs cardiaques pour la chirurgie à cœur battant.* L'Octopus 4 de Medtronic (a) fonctionne par succion, tandis que l'Immobilizer de Genzyme (b) fonctionne par pression.

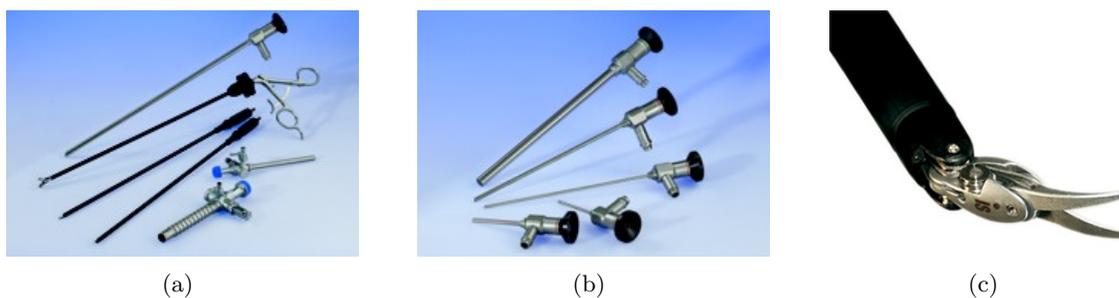


FIG. 1.12 – *Outils de chirurgie laparoscopique.* (a) divers outils pour la chirurgie mini-invasive, (b) endoscopes, (c) outils EndoWrist du robot DaVinci. (sources : www.didema.de et *Intuitive Surgical*, oct. 2007)

(figure 1.12a) qui ne nécessitent qu'une petite incision. La vision du chirurgien est assurée par une caméra reliée à un endoscope (figure 1.12b). Cet équipement présente de nombreux inconvénients pour le chirurgien :

- une mobilité réduite par rapport aux instruments “classiques”, même si les systèmes robotiques permettent désormais de restituer des mouvements complexes (figure 1.12c) ;
- une vision limitée due à l'endoscope ;
- une perte des sensations due au passage de l'instrument par le trocart ;
- une position inconfortable du chirurgien, qui peut être très éprouvante dans le cas d'une opération longue.

Les opérations en chirurgie mini-invasive sur le cœur restent rares, malgré une progression notable ces dernières années. La raison principale est qu'opérer le cœur de cette façon implique d'opérer sur cœur battant. Or l'opération sur cœur battant est déjà, en soi, un tâche très complexe. De plus, il n'est pas possible d'utiliser des systèmes de stabilisation "classiques" pour ce type d'opération ; il est cependant désormais possible d'utiliser des systèmes dérivés, adaptés aux outils laparoscopiques. Enfin, ce type d'opération reste très technique : elle est donc réservée à des chirurgiens virtuoses, et comporte une bonne dose de risque. Pour pallier ce risque, on garde à disposition durant toute l'opération une équipe médicale et du matériel de réanimation, afin de pouvoir procéder en urgence à une conversion en chirurgie classique.

1.3 Robotique médicale

La pratique de la chirurgie telle qu'on la connaît aujourd'hui est amenée à évoluer face à l'arrivée dans le bloc opératoire des technologies de l'information sous toutes leurs formes. Longtemps préservée de la vague informatique, la pratique chirurgicale commence à adopter les outils intelligents que la recherche développe depuis maintenant une vingtaine d'années. Ainsi le domaine de l'imagerie médicale est-il maintenant parfaitement intégré à la phase de planning pré-opératoire. À terme, il en sera de même pour les dispositifs d'assistance au diagnostic, de réalité augmentée ou d'assistance robotisée. C'est évidemment sur ces dernières applications que nous allons porter notre attention, afin d'introduire aussi bien les enjeux que les contraintes d'une telle évolution des instruments médicaux.

1.3.1 Genèse de la discipline

Les premiers systèmes robotisés à fin d'intervention remontent aux années 80 et reposaient sur des version modifiées de robots industriels, tel le robot PUMA 560 [Kwoh et al., 1988] utilisé en **1985** à des fins de neurochirurgie. La même année, Taylor développait un robot pour la chirurgie orthopédique basé sur la technologie *SCARA* d'IBM. Ce robot sera finalement utilisé avec succès sur animal [Taylor et al., 1989]. Suite à ces prémices de plus en plus de laboratoires se sont intéressés au domaine. Ainsi une équipe d'Imperial College a développé dès **1989** un robot pour l'ablation de la prostate, toujours sur une base de PUMA 560. Ce premier travail servira de base à la même équipe pour développer un robot de prostatectomie, le *PROBOT* [Harris et al., 1995, 1997a]. Parallèlement, le robot *Neuromate* était développé par la société IMMI en coopération avec le CHU de Grenoble de **1985** à **1991** pour améliorer la précision en *chirurgie stéréotaxique* (figure 1.9). Puis on procède au prélèvement des artères mammaires nécessaires au pontages (un prélèvement

par pontage)¹³. Il est utilisé comme porte-instrument, ce dernier étant manipulé par le chirurgien. Il peut se prévaloir d'excellents résultats pratiques, avec plus de 2000 opérations réussies¹⁴, en apportant de réels avantages au patient par rapport à une technique "traditionnelle", ce qui n'est pas toujours le cas de robots bien plus perfectionnés, comme nous le verrons. En **1992** a été lancé le système *Robodoc*, dispositif autonome (c'est-à-dire qu'il réalise le geste sans l'intervention du chirurgien) pour l'orthopédie de la hanche. Il permet de creuser très précisément le fémur du patient pour la pose de prothèse de hanche. Puis en **1993** la société Computer Motion¹⁵ a proposé un robot industriel, *AESOP 1000*, destiné à la manipulation d'endoscope. Ce robot témoigne de l'intérêt du secteur privé pour la niche industrielle que constitue la robotique médicale. Il fut l'un des premiers à obtenir l'agrément de la *Foods and Drugs Administration* américaine. Durant la fin des années 1990, les travaux de recherche se sont multipliés dans des domaines divers : l'orthopédie avec le robot *Acrobot* [Harris et al., 1997b], l'échographie 3D avec le robot *Hippocrate* [Pierrot et al., 1999a] ou la chirurgie réparatrice avec le robot *DermaRob* [Dombre et al., 2003]. Ce dernier propose une procédure totalement automatisée pour le prélèvement de peau [Dombre et al., 2003, Duchemin et al., 2004] et démontre la capacité des robots à "apprendre" les gestes des spécialistes. Du côté des industriels, on notera l'arrivée de deux imposants systèmes destinés à la chirurgie mini-invasive : le robot *Zeus* de Computer Motion en **1997** et son concurrent le robot *DaVinci* d'Intuitive Surgical en **1998**. Le second sera d'ailleurs utilisé en **2001** lors de la très médiatique opération Lindbergh au cours de laquelle le chirurgien depuis New York a enlevé la vésicule d'une patiente à Strasbourg [Marescaux et al., 2001]. Ces robots constituent une étape importante car ce sont des systèmes relativement polyvalents en comparaison avec les robots spécialisés déjà évoqués. Computer Motion ayant été repris en **2003** par Intuitive Surgical, seul le robot *DaVinci* est encore sur le marché. Environ 700 exemplaires ont été commercialisés à ce jour. Ils sont principalement utilisés en urologie et constituent par nature une base de développement pour la recherche en permettant d'avoir un retour du monde médical sur l'utilisation d'outils robotiques pour des opérations complexes. De son côté, la communauté robotique internationale s'est rapidement appropriée les nouvelles problématiques portées par ces applications. Le nombre important de publications sur l'assistance médicale robotisée témoigne d'ailleurs de cet engouement. Depuis maintenant plusieurs années, la communauté s'est regroupée autour de manifestations spécifiques telles que les conférences

¹³cette chirurgie consiste à viser très précisément une région dans le cerveau.

¹⁴chiffres de 2000Web.

¹⁵*Computer Motion*, fondée en 1989, fut une des sociétés pilotes en Robotique Médicale. Elle fut absorbée en 2003 par son principal concurrent, *Intuitive Surgical*.

*MICCAI*¹⁶, *CARS*¹⁷, *Surgetica*¹⁸, *BioRob*¹⁹ ou *EMBC*²⁰. De grandes conférences de robotique, telles que *ICRA*²¹ ou *IROS*²², proposent régulièrement des sessions sur le sujet. Des revues scientifiques se sont également intéressées au domaine, comme la revue *IEEE Transactions on Robotics and Automation* qui a consacré en **2003** un numéro spécial sur la robotique médicale [spe, 2003].

1.3.2 Avantages recherchés

Suite à l’engouement scientifique des années 1980 et 1990 pour ce domaine naissant, plusieurs chercheurs et organismes se sont interrogés sur les différents avantages pouvant motiver l’introduction de robots dans le bloc opératoire. On notera à ce propos le travail de [Davies, 2000] qui présente plusieurs raisons pouvant motiver les praticiens à “franchir le pas” :

- la capacité de placer précisément les outils dans un environnement tridimensionnel complexe ;
- la capacité de restreindre le déplacement des outils sur des trajectoires spécifiques ;
- la capacité des robots à reproduire des mouvements sans fatigue, ou de rester en position sans trembler ;
- la capacité des robots à opérer dans des environnements critiques, comme dans un scanner X ou en présence de radioactivité ;
- la rapidité et la précision (voire la micro précision) des systèmes robotisés.

Pour l’Agence Nationale d’Accréditation et d’Évaluation en Santé (ANAÉS) les raisons favorisant l’introduction des robots dans le milieu médical [Corbillon and Pouillé, 2002] sont le développement de l’imagerie médicale, l’exigence de qualité des interventions et l’émergence de la chirurgie mini-invasive.

1.3.3 Nature des solutions robotiques actuellement disponibles

Depuis la fin des années 90, le développement des robots médicaux s’est accéléré. Ainsi des disciplines comme l’orthopédie ou l’urologie ont vu l’arrivée rapide de nouvelles solutions spécialisées.

¹⁶Medical Image Computing and Computer Assisted Intervention.

¹⁷Computer Assisted Radiology and Surgery.

¹⁸Computer-Aided Medical Interventions : tools and applications.

¹⁹International Conference on Biomedical Robotics and Biomechatronics.

²⁰International Conference of the IEEE Engineering in Medicine and Biology Society.

²¹International Conference on Robotics and Automation.

²²International Conference on Intelligent Robots and Systems.

1.3.3.1 En chirurgie orthopédique

Le cas de la chirurgie orthopédique est exemplaire, car il s'agit du domaine où les applications sont apparues le plus rapidement. Nous avons déjà cité *ROBODOC* (figure 1.13a), destiné à la chirurgie du genou et de la hanche. Nous pouvons également citer *CASPAR* (figure 1.13b)²³, produit par OrthoMaquet puis URS, pour la chirurgie de la hanche. Ces deux systèmes, développés et commercialisés dès le début des années 1990 fonctionnaient comme des machines-outils : l'opération était tout d'abord planifiée, puis réalisée de façon autonome par le robot. Ces solutions n'ont pas connu de franc succès (une centaine de systèmes *ROBODOC* vendus), principalement à cause de leur coût prohibitif, et les sociétés impliquées dans leur développement ont dû déposer le bilan.

Depuis, des solutions ont été développées autour de robots laissant au chirurgien la tâche de découpe à réaliser, tels qu'*ACRODOT* déjà cité [Harris et al., 1997b] et *BRIGIT* [Maillet et al., 2005] (figure 1.13f). Les avantages de l'assistance robotique sont nombreux dans le cadre des orthoplasties, où la durée de vie de la prothèse est fonction de la qualité des plans de coupe réalisés [Kienzle et al., 1995]. De même pour certaines pathologies du rachis, on met en place des prothèses métalliques le long de la colonne vertébrale par vissage pédiculaire qui nécessite une grande précision. La solution proposée dans [Shoham et al., 2003] est originale : le robot de perçage qui se fixe directement sur les apophyses épineuses du patient, est ainsi solidaire des mouvements du patient.

On notera cependant que les robots n'ont pas encore envahi les blocs de chirurgie orthopédique. La part y est encore belle pour les systèmes de navigation, aussi appelés localisateurs, qui permettent d'exécuter un planning pré-opératoire de façon très précise. Un localisateur est constitué d'un ensemble de caméras et de cibles montées sur les instruments. De nombreuses offres commerciales sont proposées, comme par exemple les systèmes *Surgetics* de Praxim-Medivision, *StealthStation* de Medtronic ou *OrthoPilot* d'Aesculap.

1.3.3.2 Systèmes dédiés à la chirurgie mini-invasive

Afin d'améliorer le confort du chirurgien pratiquant la chirurgie mini-invasive, des industriels ont proposé des solutions à base de robots télé-opérés. Les premiers robots développés pour la chirurgie mini-invasive sont les porte-endoscope comme *Aesop* (figure 1.13e) de Computer Motion, commandé par la voix du chirurgien [Allaf et al., 1998], *EndoAssist* de Armstrong Healthcare contrôlé par les mouvements de la tête du chirurgien [Aiono et al., 1999], *Lapman* de MedSys contrôlé manuellement grâce à un boîtier électronique intégré sur l'outil [Polet and Donnez, 2004]. On a ensuite vu se développer les systèmes maître-esclave dans lesquels le chirurgien opère depuis une console maître. Les bras esclaves

²³créé en 1997 et postérieur à *ROBODOC*, il n'est resté sur le marché que deux ans.



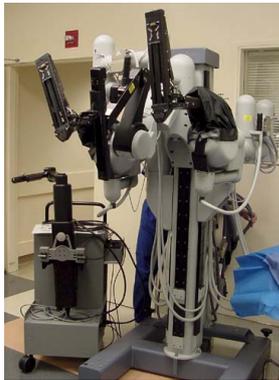
(a)



(b)



(c)



(d)



(e)



(f)



(g)



(h)

FIG. 1.13 – *Exemples de robots médicaux.* (a) ROBODOC, (b) CASPAR, (c) Zeus, (d) DaVinci, (e) Aesop, (f) BRIGIT, (g) AcuBot, (h) DERMAROB.

portent les instruments et l'endoscope comme le robot *DaVinci* [Guthart and Salisbury, 2000] déjà cité. D'autres projets restés au stade de prototypes ont proposé des alternatives à ces solutions, par exemple les systèmes *RTW* [Cavusoglu and Tendrick, 1999, Cavusoglu et al., 2001] (Robotic Telesurgery Workstation), *ARTEMIS* (Advanced Robot and Telem manipulator System for Minimal Invasive Surgery) en 1999 [Rininsland, 1999], et *Endoxi rob* [Van Meer, 2005]. De nombreuses recherches portent également sur l'amélioration de la manipulabilité, qui est réduite par le passage par le trocart des instruments de chirurgie mini-invasive [Dombre et al., 2004, Ikuta et al., 2002].

1.3.3.3 Imagerie interventionnelle

L'imagerie interventionnelle a pour objectif de réaliser des gestes chirurgicaux sous imagerie fonctionnelle : scanner, IRM, etc. L'utilisation de solutions robotisées permet de contourner les principaux défauts de ce type d'intervention : la difficulté de coordonner un mouvement avec ce type d'image et la dangerosité des rayonnements auxquels s'expose le chirurgien. Les systèmes robotiques répondant à ces problèmes utilisent une information visuelle [Jaffray et al., 2001, Hong et al., 2004]. On peut citer plusieurs robots ainsi commandés : *Acubot* [Stoianovici et al., 2003] pour les ablations de la prostate sous rayons X (figure 1.13g), *CT-Bot* qui opère dans un scanner [Maurin et al., 2004] ou *LPR* (Light Puncture Robot) [Taillant et al., 2005]. Les structures doivent être en matériaux compatibles avec les rayonnements de l'imagerie utilisée. Les actionneurs doivent être déportés (d'où par exemple l'utilisation de l'énergie pneumatique dans LPR).

1.3.3.4 Aide au diagnostic

Cette catégorie regroupe principalement des robots porte-sonde échographique comme le robot *Hippocrate* [Pierrot et al., 1999b] (actuellement à l'hôpital Broussais à Paris) et le robot développé par Salcudean et al. [Abolmaesumi et al., 2002] qui utilisent un retour d'effort pour réguler la pression de la sonde sur la peau et faire ainsi les examens cliniques de façon reproductible (par exemple, suivi dans le temps du volume des plaques d'athérome chez des patients à risque avec Hippocrate). L'utilisation de systèmes télé-opérés permet également d'envisager des diagnostics à distance. Ainsi les systèmes *TER* [Vilchis et al., 2001] et *Otelo*²⁴ [Delgorge et al., 2005] permettent à une personne non spécialiste de demander l'avis d'un expert situé dans un autre centre hospitalier qui peut piloter la sonde à distance.

²⁴maintenant commercialisé par Robosoft sous le nom d'*Estelle*.

1.3.3.5 Autres domaines d'applications

Parmi les domaines médicaux dans lesquels la robotique commence à se développer, notons la chirurgie réparatrice avec le projet *DERMAROB* (robot *SCALPP*, figure 1.13h) [Dombre et al., 2003] qui automatise le prélèvement de peau pour les grands brûlés ou encore les systèmes destinés à la chirurgie oculaire [Kumar et al., 2000, Rivière et al., 2003]. Ces solutions, déjà très performantes montrent un fois de plus la souplesse de l'outil robotique et sa capacité à s'adapter aux contraintes du milieu médical.

1.4 Problématique considérée

1.4.1 Objectifs médicaux

Depuis l'apparition des systèmes tels que le *DaVinci*, on sait pratiquer des opérations complexes de chirurgie mini-invasive par voie robotique. Cependant, certaines opérations restent hors de portée de ces systèmes, typiquement dès lors qu'elles touchent des organes en mouvement, et particulièrement les opérations à cœur battant. Opérer sur cœur battant présente de nombreux avantages par rapport à une approche plus "classique", nous l'avons vu, mais impose en contrepartie une plus grande complexité dans la réalisation des gestes chirurgicaux. Ces interventions sont donc l'apanage de chirurgiens très spécialisés, et ce malgré le développement de stabilisateurs dédiés.

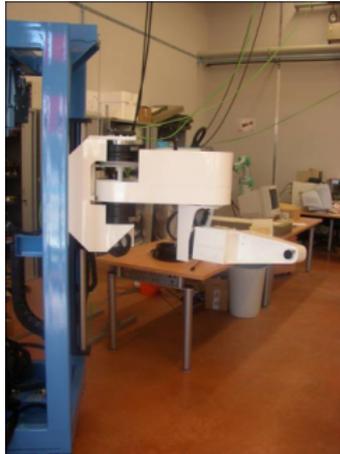
En intégrant la compensation de mouvement, il serait possible de proposer au praticien d'opérer un cœur battant dans un environnement virtuellement stabilisé, et de lever ainsi certains obstacles pratiques à la réalisation de ce type d'opération.

1.4.2 Problèmes scientifiques liés à la compensation de mouvements

Trois problèmes se posent lors de la mise au point d'un système de compensation de mouvements :

- **un problème mécanique** : le robot a-t-il une dynamique suffisante pour suivre les mouvements du cœur ?
- **un problème de capteurs** : est-on capable d'évaluer la position courante de l'organe en mouvement par rapport au robot ?
- **un problème de commande** : la loi de commande est elle suffisamment performante pour suivre le mouvement parfois très complexe de l'organe ?

Au LIRMM, le premier point a été résolu en développant un robot spécialisé, le robot D2M2 (figure 1.14) dont l'architecture de type SCARA et des moteurs à entraînement

FIG. 1.14 – *Le robot D2M2 développé au LIRMM.*

direct sur les trois premières articulations permettent un grande dynamique. D'autres expériences ont été menées par des laboratoires avec des robots industriels qui ont montré des résultats très encourageants. Quelle que soit la solution, il est clair qu'un robot puissant pose en contrepartie des problèmes de sécurité qui, s'ils ne sont pas un frein à des expérimentations de laboratoire, devront être considérés avant un éventuel déploiement en milieu médical.

Pour traiter l'aspect capteur, nous avons choisi d'explorer des techniques de traitement d'images et de vision : c'est le propos du présent document. Ce type de capteur a d'ailleurs été employé par d'autres laboratoires [Ginhoux et al., 2004, Nakamura et al., 2001], l'engouement pour cette approche s'expliquant par ses facilités de mise en œuvre :

- l'absence de contact avec l'organe limite les risques pour le patient,
- le signal vidéo peut être acquis à haute fréquence, contrairement à l'imagerie fonctionnelle 4D de type IRM ou scanner,
- les systèmes actuels disposent déjà d'un retour visuel, généralement par imagerie endoscopique,
- le matériel nécessaire pour la mise en œuvre est réduit.

Cette approche présente aussi des défauts, comme la complexité du signal, l'impossibilité de démontrer la convergence dans des cas extrêmes (comme en cas de masquage de la cible) ou encore la précision du capteur vidéo qui est limitée²⁵. En outre, les algorithmes de reconstruction sont généralement mal conditionnés. Ces problèmes peuvent être contournés en utilisant des amers artificiels dans l'image, comme des patchs posés sur le cœur ou des marqueurs laser [Ginhoux, 2003]. D'autres approches de compensation sont envisageables,

²⁵ bien que les progrès dans ce domaine soient très rapides. Pour nos expériences nous disposons d'un capteur CMOS 1000×1000 pixels, ce qui est supérieur aux résolutions rencontrées par défaut sur des systèmes tels que le *DaVinci*.

comme l'utilisation d'un capteur d'effort au contact de l'organe.

Enfin, la partie commande est très délicate à mettre en œuvre car les mouvements de l'organe sont irréguliers et complexes. Des techniques de commande prédictive ont été utilisées [Ginhoux et al., 2004, Sauvée, 2006] afin de pallier ce problème, en proposant ainsi un filtrage plus efficace du signal. Des méthodes basées sur l'analyse harmonique ont également été développées pour capter la régularité du signal [Ginhoux, 2003]. Enfin des approches par commande robuste sont actuellement en développement.

1.4.3 Problèmes liés à l'aspect vision

Les aspects vision et traitement d'image peuvent eux aussi être regroupés autour de trois grandes problématiques :

- **l'acquisition des images** qui regroupe les problèmes de vitesse d'acquisition, de qualité de l'image (déformations, résolution, réflexions), de synchronisation dans le cas où plusieurs caméras sont utilisées, etc. Il s'agit d'un problème avant tout matériel, mais des solutions logicielles sont envisageables pour rectifier certains défauts des images ;
- **le suivi des mouvements** de l'organe dans les images. Il doit être le plus robuste possible car il est la clé de voûte de toute la chaîne de traitement²⁶ ;
- **la reconstruction 3D** ou la reconstruction de pose dans le cas où la géométrie est connue. Cette étape repose sur la précédente, et de nombreuses solutions sont envisageables, comme illustré au paragraphe 1.5.

Dans notre contexte, la difficulté vient en premier lieu de la qualité des images fournies par l'endoscope, puis de la nature déformable et peu contrastée de la surface du cœur. L'analyse de ces images est complexe et nous avons dû rechercher un jeu d'outils de texture adaptés à nos images (chapitre 2), puis nous avons utilisé cette caractérisation pour effectuer le suivi de mouvement cardiaque (chapitre 3).

Comme de plus nous ne disposons pas d'un modèle fiable de la géométrie de l'organe (malgré des travaux initiés en ce sens), il n'est pas possible d'utiliser un algorithme de reconstruction de pose (tel que [Dhome et al., 1989, DeMenthon and Davis, 1995]) pour évaluer le mouvement par une approche mono-caméra. C'est pourquoi nous présenterons au chapitre 4 une approche basée sur la stéréo-vision qui lève cette ambiguïté.

²⁶notre approche, développée au chapitre 3, se base sur un schéma extraction de primitive → suivi. Il existe d'autres approches qui ne font pas cette distinction, telles que celle développée dans [Benhimane and Malis, 2007].

1.5 Travaux antérieurs

1.5.0.1 Travaux de Nakamura

Le premier travail de compensation de mouvement cardiaque par vision est dû à [Nakamura et al., 2001], lorsqu’il proposa un suivi du cœur à haute fréquence (955 images/sec.) par asservissement visuel (figure 1.15). Cette étude poursuivait les mêmes objectifs que nous, à savoir l’asservissement visuel pour la compensation de mouvements en chirurgie mini-invasive.

Le robot à 4 degrés de liberté, spécialement développé pour cette application, se monte sur l’écarteur sternal du patient. Il s’agit d’une solution ingénieuse qui élimine de fait la perturbation induite par la respiration.

Le suivi est assuré par un système d’acquisition à 995Hz, constitué de deux caméras montées devant un prisme semi-réfléchissant (une caméra couleur à fréquence vidéo pour le retour visuel et une caméra haute vitesse DALSA CA-D6 pour l’acquisition monochrome à haute vitesse destinée à l’asservissement) . La méthode de suivi n’est pas détaillée, mais il semble que les auteurs suivent un petit marqueur à la surface de l’organe, probablement par seuillage²⁷. Deux applications sont proposées :

- la stabilisation “visuelle” : la caméra monochrome détecte le déplacement du marqueur, et en déduit la translation de l’image nécessaire pour garder celui-ci “fixe” à l’écran. Le programme modifie l’image de la caméra couleur, ainsi l’opérateur voit l’image immobile ;
- l’asservissement du robot, parallèlement au plan image (pas d’asservissement en z).

Cette approche souffre de plusieurs limitations. Tout d’abord les auteurs asservissent le robot en 2D : la distance entre la caméra et l’organe n’est pas évaluée ; de plus, le suivi est réalisé en utilisant un marqueur, ce qui paraît difficilement réalisable en pratique lors d’une intervention.

1.5.0.2 Plate-forme du DLR

Dans son travail au DLR [Ortmaier, 1999], Tobias Ortmaier a développé un système de vision permettant le suivi de points sur cœur battant (avec stabilisateur). Le suivi est basé sur la SSD (§3.3.1.2), illustré sur la figure 1.16, assisté d’un algorithme d’élimination des réflexions.

Les réflexions sont détectées par seuillage, puis une érosion morphologique est appliquée. Ces zones sont alors reconstruites par interpolation linéaire sur les valeur de luminosité selon des directions déduites du tenseur de structure. Cette méthode, elle-même

²⁷c’est très probable vu la vitesse de traitement.

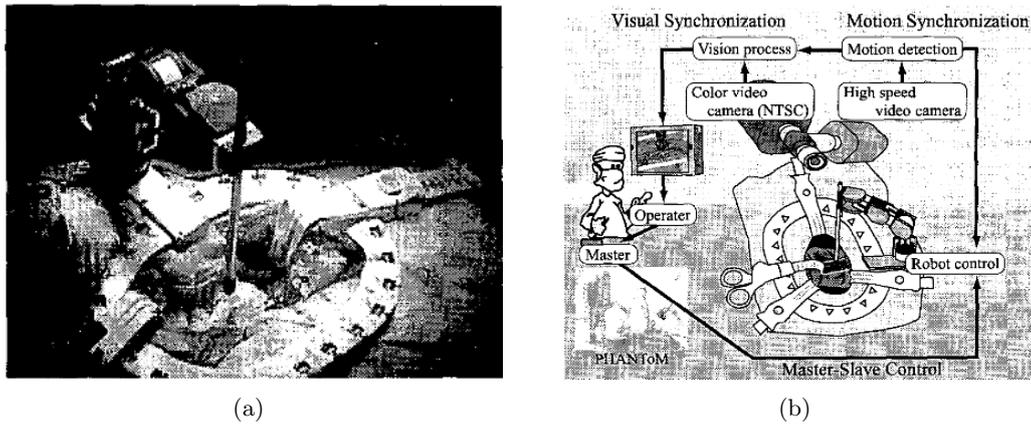


FIG. 1.15 – *Asservissement Visuel sur cœur battant en x et y* [Nakamura et al., 2001].

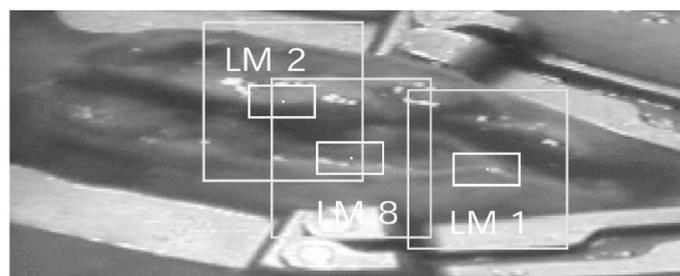


FIG. 1.16 – *Méthode de suivi du DLR* [Ortmaier, 1999]. *On observe les points suivis ainsi que la zone de recherche.*

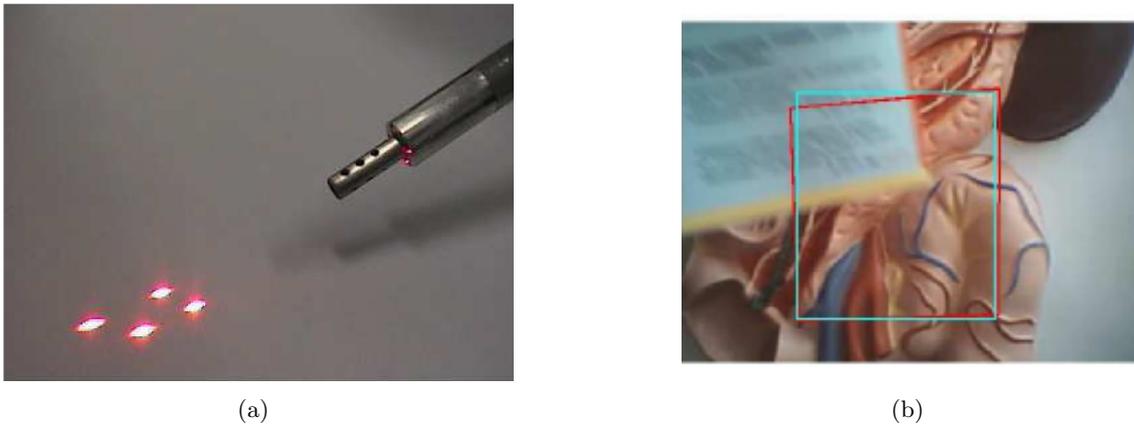


FIG. 1.17 – *Techniques de vision utilisées au LSIIT.* (a) approche utilisant des pointeurs laser [Ginhoux, 2003], (b) Approche par suivi ESM [Bourger et al., 2007]. (Exemple de fenêtre de suivi et de recalage en présence d’occultations).

discutée dans [Gröger et al., 2001], est très rapide, mais a tendance à lisser fortement l’image si les réflexions sont trop importantes.

L’auteur propose enfin d’utiliser le signal ECG pour prédire le déplacement des marqueurs occultés afin de rendre le suivi plus robuste.

1.5.0.3 Plate-forme du LSIIT

Les travaux du LSIIT [Ginhoux, 2003, Gangloff et al., 2006, Cuvillon, 2006, Salle, 2005] sont assez similaires à ceux de Nakamura : quatre pointeurs lasers sont utilisés afin de mesurer la distance entre l’outil et la surface du cœur qui est la donnée la plus importante et la plus critique pour la compensation de mouvement. De plus, une loi de commande très élaborée utilisant un contrôleur prédictif associé à un modèle du mouvement cardiaque a été développée.

Les résultats sont très intéressants, surtout dans le plan tangent. Cependant la compensation de mouvements selon l’axe de l’outil présente des erreurs importantes, probablement dues à la bande passante du robot. Dans sa thèse, Alexandre Krupa [Krupa, 2003] propose des solutions à base de traitement d’images pour rendre plus robuste l’asservissement visuel du robot *AESOP*. Ces différents travaux utilisent un outil équipé de pointeurs laser, illustré sur la figure 1.17a.

Le traitement d’image bas niveau est basé sur une segmentation, puis sur une détection du centre des taches obtenues, afin d’obtenir une localisation précise des marqueurs laser (car la surface du cœur a tendance à diffuser la marque laser). Puis la reconstruction effective utilise la méthode de Tsai [Tsai, 1986], adaptée au cas des points coplanaires.

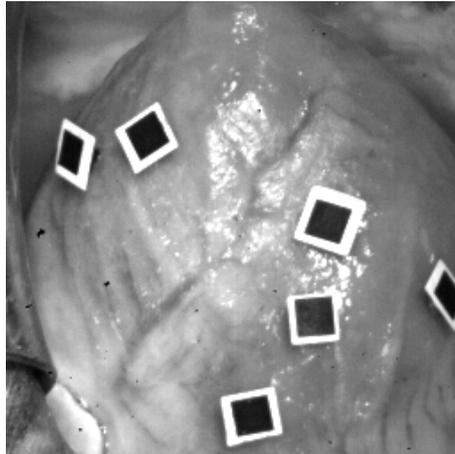


FIG. 1.18 – Patches utilisés pour le suivi de cœur battant dans [Sauvée, 2006].

Une autre méthode, illustrée figure 1.17b, a été développée plus récemment au sein du LSIIT pour la compensation de mouvements [Bourger et al., 2007] sans utiliser de marqueur. Cette technique utilise la méthode *ESM*²⁸ pour évaluer l’homographie entre les images consécutives et en déduire le déplacement nécessaire de l’outil. La partie traitement d’images utilisée comporte notamment une gestion des occultations afin de rendre le suivi plus robuste.

1.5.0.4 Travaux antérieurs au sein du LIRMM

Enfin, citons les travaux effectués au LIRMM par M. Sauvée [Sauvée, 2006] sur la reconstruction de mouvements cardiaques en utilisant des patchs rigides. Cette approche est illustrée sur la figure 1.18.

Ces résultats utilisent la méthode de suivi ESM et la décomposition de l’homographie estimée pour calculer la pose des patchs. Cette méthode est inspirée de la méthode de Zhang [Zhang, 2000] qui permet de calibrer une caméra à partir d’un mire plane, puis d’estimer les différentes poses correspondant aux différentes vues de la mire. Les résultats obtenus montrent une bonne précision dans la reconstruction.

1.6 Conclusion

Dans ce chapitre, nous avons introduit la problématique de la chirurgie robotisée, et plus particulièrement de la chirurgie cardiaque robotisée. C’est un problème ouvert, qui

²⁸une méthode récente de suivi de mouvements dans les séquences d’images proposée par Malis [Malis, 2004].

connaîtra certainement de très nombreux développements dans les années à venir. La problématique de la chirurgie à cœur battant est particulièrement stimulante sur un plan scientifique, car elle est au centre de plusieurs domaines de pointe : la robotique, avec les problèmes de dynamique liés aux robots, l'automatique, avec par exemple la commande prédictive dans les travaux de [Ginhoux, 2003] ou [Sauvée, 2006], le traitement du signal et la vision avec les problèmes de suivi et de reconstruction, et même la médecine, car les opérations de chirurgie mini-invasive et sur cœur battant sont clairement parmi les plus complexes de la chirurgie actuelle. Le chapitre suivant présente la première étape de notre travail, une étude des approches d'analyse de texture, qui a permis par la suite de développer la méthode de suivi présentée au Chapitre 3.

Chapitre 2

Analyse de texture

2.1 Introduction

L'analyse de texture est une problématique ancienne en traitement d'images. Depuis les premiers travaux de Julesz [Julesz, 1975], jusqu'aux derniers développements théoriques [Peyré, 2005], la notion même de texture est restée diffuse, et des modèles très différents ont été proposés pour la caractériser (§2.2). Pourtant, l'objectif de toutes ces approches est sensiblement le même : caractériser l'homogénéité présente dans les images et qui découle de la superposition d'éléments similaires, mais non identiques (tels que les brins d'herbe d'une pelouse, les fibres d'un bois, ou encore la structure alvéolaire des os). Les domaines de prédilection de l'analyse de texture sont, à l'heure actuelle, la segmentation d'images ou les systèmes experts. Parmi ceux-ci nous trouvons des applications cartographiques [Lorette, 1999] ou médicales [Materka et al., 1999, Ayache, 1999, Bonilha et al., 2003].

Dans le cadre de notre application, nous avons considéré l'analyse de texture pour pallier les insuffisances de notre algorithme de suivi (à base de *pattern matching*, voir chapitre 3) sur des images de cœur battant. Nous cherchions une caractérisation plus robuste que la comparaison pixel à pixel des motifs, et nous verrons à cet égard que l'analyse de texture permet de définir une série de descripteurs, dits *attributs texturaux*, aux performances très intéressantes. L'objectif central de cette thèse était l'étude de ces approches et leur intégration pour le suivi des mouvements du cœur.

Ce chapitre introduit les différents modèles de texture que nous avons comparés, les méthodes utilisées dans le cadre de cette thèse, et les résultats obtenus. La description mathématique des différents attributs texturaux a été repoussée en annexe pour privilégier la discussion et l'illustration des résultats. Le §2.2 présente les différentes approches de l'analyse de texture. Puis la discussion expose les raisons qui nous ont amenés à nous focaliser sur les approches statistiques. Le §2.3 détaille les méthodes d'analyse de données utilisées. Ce paragraphe introduit également notre jeu de textures expérimentales ainsi

	Attributs texturaux	Type d'information extraite
Modèles purement statistiques	attributs statistiques, matrice de co-occurrence, matrice de longueur des plages...	descripteurs locaux
Modèles spectraux	énergie, attributs statistiques sur les composantes spectrales.	coefficients spectraux.
Modèles structurels		texton + répartition du texton dans l'image
Champs de Markov		modèle de Markov (HMM).
Modèles AR/ARMA		coefficients du modèle.

TAB. 2.1 – *Vue d'ensemble des modèles de texture.*

qu'une série de jeux de données classiques utilisée dans un but de validation de nos résultats. Enfin le §3.5 expose les résultats obtenus sur les différents jeux de données et les met en parallèle.

2.2 Modèles pour l'analyse de texture : état de l'art

Le tableau 2.1 résume les principales caractéristiques des approches que nous avons évaluées dans notre étude. Si il n'est pas possible de comparer la totalité des approches disponibles dans la littérature scientifique, ce tableau résume néanmoins les principales familles rencontrées. On remarque une prédominance des représentations statistiques (modèles statistiques, champs de Markov et modèles auto-régressifs). Dans le tableau 2.1, nous avons séparé d'un côté la caractérisation par attributs texturaux, qui a finalement retenu notre attention, et les autres types de caractérisation que nous avons mis de côté. Les raisons de ce choix seront explicitées au §2.2.7.

2.2.1 Modèles statistiques

2.2.1.1 Introduction

Les approches statistiques pour l'analyse de texture sont les plus anciennes et sûrement les plus utilisées à l'heure actuelle. Elles constituent un domaine à part : Haralick, dans son article [Haralick, 1979], ne distinguait d'ailleurs que deux types d'analyse de texture : *structurelle* et *statistique*. Ces approches ne proposent pas d'analyser la texture par

sa structure spatiale, mais envisagent la région uniformément texturée comme un signal aléatoire possédant certaines caractéristiques statistiques¹ stables.

Afin de déterminer quels sont les paramètres statistiques utiles en perception humaine, des expériences psycho-visuelles ont été menées dans les années 1970 [Julez, 1975]. Les sujets se sont vus présenter des textures générées artificiellement et ayant certaines caractéristiques communes. Le but était de voir quelles textures étaient perçues *spontanément* comme “similaires” ou “différentes”. Ces tests ont montré la prédominance des relations inter-pixels d’ordre 1 (i.e. sur l’histogramme de l’image) et d’ordre 2 (i.e. sur les paires de pixels) par rapport à des statistiques d’ordre supérieur. Ainsi, pour discriminer des textures ayant leurs moments d’ordre 1 et 2 identiques, il faut un effort conscient supplémentaire [Julez, 1975]. On peut donc supposer que l’utilisation de statistiques d’ordre 1 et 2 puissent permettre une discrimination proche de celle de l’œil humain.

Ces approches ont de nombreuses applications et se montrent très performantes du point de vue de la discrimination, qui pour nous est le critère le plus important. Les matrices de co-occurrence [Haralick, 1979] par exemple sont des statistiques d’ordre 2 encore très utilisées [Lepistö et al., 2003] et développées [Clausi and Zhao, 2001]. Elle se montrent souvent plus performantes que des approches par transformation [Oja and Valkealahit, 1998] pour une complexité globalement moindre. De plus leurs applications au domaine médical sont nombreuses [Wei et al., 2006, Jafari-Khouzani et al., 2003], principalement pour des outils de diagnostic.

2.2.1.2 Matrices de co-occurrence

Les **matrices de dépendance spatiale de niveaux de gris**, ou **matrices de co-occurrence** contiennent les moyennes d’espace du second ordre. Elle sont construites comme un histogramme des couples de pixels, séparés par un vecteur de translation \mathbf{u} , défini suivant une distance et une orientation² (r, θ) . L’équation (2.1), tirée de [Cocquerz et al., 1995], illustre la constitution de la matrice de co-occurrence³ :

$$C_{\mathbf{u}}(a, b) = \text{card} \left\{ (\mathbf{p}, \mathbf{p} + \mathbf{u}) \in \mathbb{R}^2 \text{ tq } \mathcal{I}(\mathbf{p}) = a, \mathcal{I}(\mathbf{p} + \mathbf{u}) = b \right\} \quad (2.1)$$

La construction de la matrice de co-occurrence est illustrée sur la figure 2.1 et les différentes notations du vecteur \mathbf{t} sont illustrées sur la figure 2.2.

Longtemps écartée des applications industrielles (et principalement temps réel) par son coût de calcul prohibitif, cette méthode a gagné de l’intérêt grâce au progrès de

¹le problème étant alors de déterminer lesquelles : c’est ce que nous étudierons au §2.3.

²on trouve souvent la notation polaire (r, θ) dans la littérature. Cependant, dès que l’on passe à l’implémentation, la notation vectorielle $(x, y)^t$ est plus pratique. La figure 2.2 illustre ces deux notations.

³ \mathbf{p} désigne un pixel de l’image (par ses coordonnées (x, y)), et $\mathcal{I}(\mathbf{p})$ la valeur de luminance correspondant à ce pixel dans l’image.

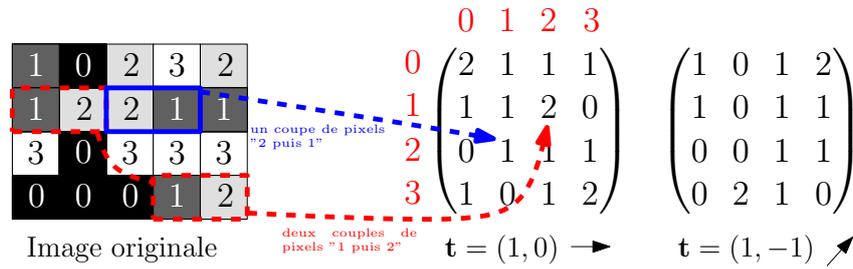


FIG. 2.1 – *Construction de la matrice de co-occurrence. Illustration de plusieurs valeurs du vecteur \mathbf{t} .*

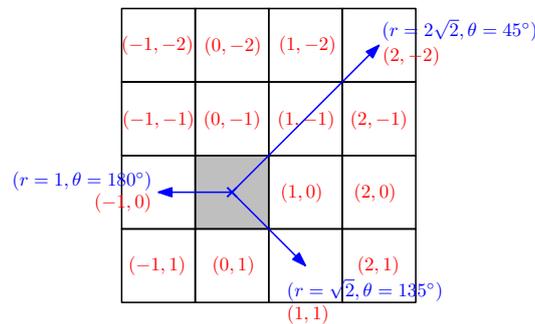


FIG. 2.2 – *Notation du vecteur \mathbf{t} .*

l'informatique. De plus, des avancées du point de vue algorithmique et une implémentation efficace (matrices symétriques, limitation du nombre de niveaux de gris, voir annexe) rendent désormais possible le calcul rapide de ces matrices.

2.2.1.3 Matrices de longueur des plages

Les **matrices de longueur des plages** (ou **run-length matrices** dans la littérature anglo-saxonne), illustrées figure 2.3, caractérisent l'homogénéité d'une région d'intérêt. Cette matrice se construit en associant à chaque niveau de gris le nombre de "plages" d'une longueur donnée, selon une direction⁴ θ , analogue à celle définie pour les matrices de co-occurrence. La figure 2.3 illustre la construction de la matrice de longueur des plages.

Il s'agit de statistiques d'ordre supérieur à 2 (les relations lient plus de 2 pixels), cependant les performances de cette approche ne sont pas forcément meilleures que celles liées par exemple aux matrices de co-occurrences (§2.3). De nombreux attributs statistiques peuvent être évalués à partir de ces matrices de longueur des plages (annexe).

⁴généralement $\theta = 0^\circ, 45^\circ, 90^\circ, -45^\circ$, cf. figure 2.3

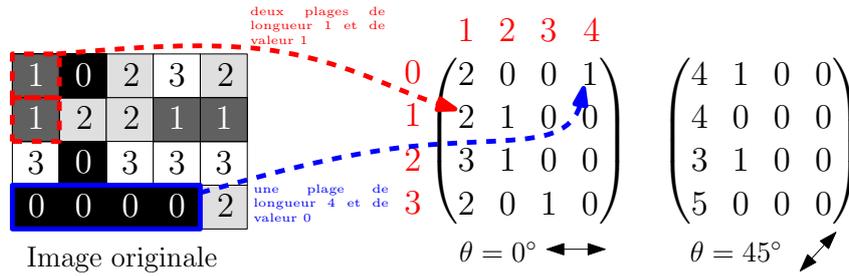


FIG. 2.3 – *Construction de la matrice de longueur des plages.* Illustration pour les principales directions (valeurs de θ).

2.2.1.4 Autres attributs statistiques

D'autres approches sont envisageables, comme le calcul de descripteurs statistiques sur les gradients de l'image, la méthode des différences de niveaux de gris (qui correspond à un histogramme construit sur le gradient bi-dimensionnel de l'image) ou le calcul d'attributs sur le spectre d'auto-corrélation ou Fourier de l'image. Ces méthodes sont détaillées en annexe.

2.2.2 Modèles spectraux

L'analyse spectrale est une voie intéressante pour l'analyse de texture par plusieurs aspects. Tout d'abord, c'est une représentation adaptée à l'analyse de signaux présentant une forte répétitivité spatiale, ce qui est le cas de nombreux motifs texturés. De plus, la pluralité des décompositions offre de nombreuses possibilités de "réglage" au niveau des filtres utilisés. Enfin, certaines propriétés, telles que l'énergie ou le contraste, s'expriment plus naturellement dans le domaine spectral que dans le domaine spatial.

Parmi les décompositions utilisées en analyse de texture, citons l'analyse de Fourier [Hsu et al., 1993], les filtres de Gabor [Mengko and Tjandra Pramudito, 2002], les transformées en ondelettes [Mallat, 1999, Scheunders et al., 1998] ou plus récemment les transformées en bandelettes (qui sont une évolution des transformées en ondelettes) [Peyré, 2005]. Dans de nombreux travaux récents, les approches multi-résolutions de type ondelettes sont privilégiées car elles permettent de capturer à la fois l'information globale et locale. De plus les ondelettes portent également une information spatiale, contrairement à la transformée de Fourier discrète⁵.

L'analyse de Fourier est surtout utilisée lorsque le signal est très répétitif. Il est alors possible d'obtenir une représentation simple (peu de coefficients) et plus facilement in-

⁵il faut aussi noter l'existence de la transformée de Fourier sur fenêtre glissante (ou transformée de Fourier Locale) qui permet d'étudier la périodicité des textures. Cependant cette approche souffre d'un temps de calcul important, qui la rend incompatible avec l'objectif recherché.

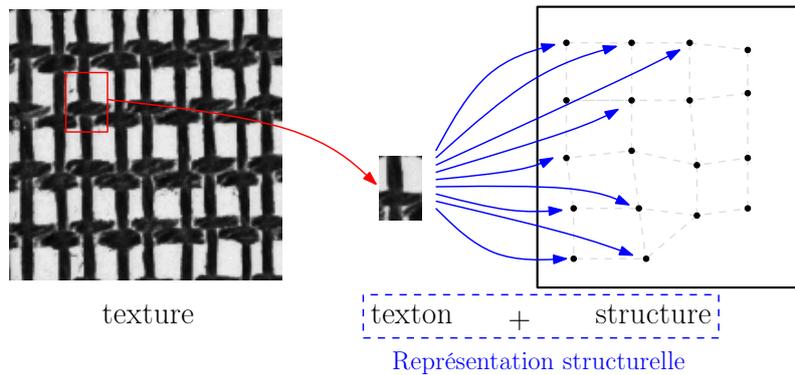


FIG. 2.4 – *Représentation structurelle d’une texture. Schéma de principe.*

terprétable. Or nos textures souffrent d’un manque de régularité, même locale, qui limite ce type d’approche. Les filtres de Gabor permettent quant à eux d’isoler une partie du spectre, de par leur réponse fréquentielle très localisée. Le principal obstacle à leur utilisation est la difficulté de trouver un filtre adapté à la taille caractéristique de la texture. De plus, ils ne portent pas d’information spatiale, ce qui les rend difficilement utilisables pour une tâche de suivi.

Les ondelettes présentent de nombreux atouts, dont une double signification dans le domaine spatial et temporel. Elles ont prouvé leur efficacité pour des applications à la segmentation⁶. Nous avons étudié quelques descripteurs classiques (maximum d’amplitude, énergie, etc.), mais nous ne voulions pas nous baser sur une approche reposant entièrement sur les ondelettes, car ce serait l’objet d’une autre thèse. De plus, le calcul des coefficients d’ondelettes reste coûteux en temps de calcul⁷.

2.2.3 Modèles structurels

Les modèles structurels [Asano et al., 2000] représentent la texture par une répartition hiérarchique de *micro-textures* ou *textons* selon un arrangement spatial particulier : la *macro-texture* (figure 2.4). La texture est alors intégralement définie par la donnée du texton et des règles nécessaires au placement des occurrences de ce texton : c’est une représentation de la texture à plusieurs échelles. Elle est également très proche du concept intuitif de texture comme étant la superposition de nombreux petits éléments constitutifs (comme les brins d’herbe formant le gazon). Un aspect particulièrement intéressant de

⁶même si le fait que les ondelettes ne soient pas invariantes par translation est souvent un facteur limitant. Dans ce cas des décompositions alternatives comme les “steerable pyramids” Portilla and Simoncelli [2000], Do and Vetterli [2002] sont envisageables

⁷théoriquement [Mallat, 1999], la *transformée en ondelettes rapide* est autant voir plus rapide que l’algorithme FFT pour la transformée de Fourier discrète. Cependant, en pratique, les approches matérielles (DSP) et logicielles (bibliothèques dédiées) donnent nettement l’avantage à l’approche par transformée de Fourier, au passage moins gourmande en espace mémoire.

cette approche est la capacité d'utiliser la décomposition micro/macro-texture pour des algorithmes de *shape from texture*, comme les travaux de Loh [2006] par exemple.

Cependant, une série de problèmes tant théoriques que pratiques se posent lors de la mise en place de ces méthodes :

- la détermination de l'élément constituant, qui peut elle-même se découper en deux étapes :
 1. Déterminer la taille du texton,
 2. Extraire le texton proprement dit ;
- la localisation des occurrences du texton dans chaque nouvelle image ;
- la reconstruction de la surface.

De plus, si cette technique se montre intéressante sur des textures artificielles, elle est peu adaptée à l'analyse de textures naturelles.

2.2.4 Modèle AR/ARMA de texture

Les modèles auto-régressifs, ou modèles AR, sont très utilisés en traitement du signal pour modéliser un signal aléatoire. Un tel signal peut s'écrire comme l'ajout d'un bruit blanc U dans un filtre purement récursif caractérisé par ses paramètres $\mathbf{A} = [a_1 \dots a_p]^t$. Pour un signal X , l'échantillon X_n est obtenu par l'équation :

$$X_n = a_1 X_{n-1} + a_2 X_{n-2} + \dots + a_p X_{n-p} + U_n \quad (2.2)$$

ou dans le cas d'une image :

$$X_n = \sum a_k X_{\phi(k)} \quad (2.3)$$

où $\phi(k)$ est une fonction décrivant un voisinage de X_n . L'analyse de texture par modèles AR consiste à identifier le vecteur A des paramètres pour la région d'intérêt considérée. Les modèles AR causaux permettent de générer facilement de nouvelles textures. Ils peuvent également être utilisés à des fins de segmentation [Alacam et al., 2003] en réalisant une classification sur le vecteur \mathbf{A} . D'autres modèles sont également envisageables, comme les modèles AR non causaux, ou les modèles auto-régressifs à moyenne mobile (dits modèles ARMA) très utilisés en automatique ou dans l'étude de modèles économiques. La figure 2.5 illustre la différence entre les modèles causaux et non-causaux.

Là encore, il est possible d'estimer les paramètres du modèle (méthode de Yule-Walker [Lacroix, 2004]) et même d'identifier le nombre de paramètres nécessaires au modèle (méthodes de Box-Jenkins ou méthode du "coin") sans connaissance *a priori* du processus étudié.

Cependant, ces méthodes ont aussi des limitations, la plus importante pour nous étant

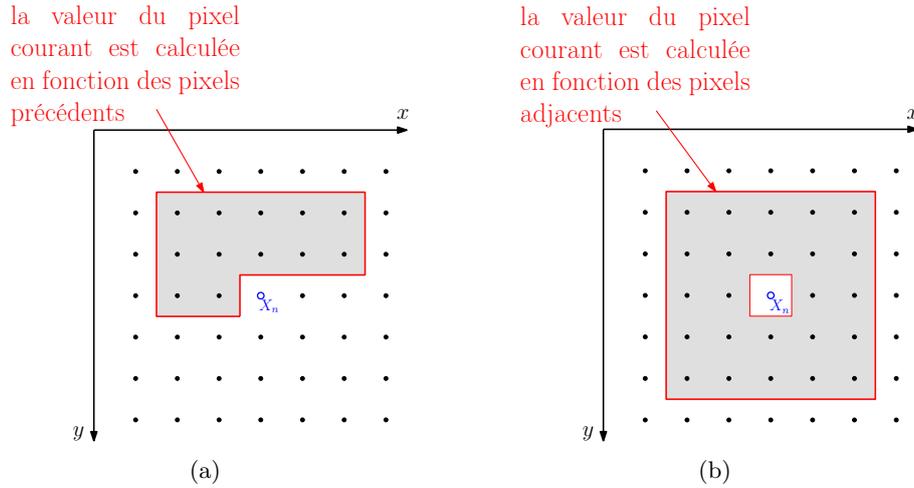


FIG. 2.5 – *Voisinages considérés dans les modèles ARMA 2D pour l'analyse de texture.* (a) *Modèle ARMA causal.* (b) *Modèle ARMA non causal.*

qu'il faut considérer un nombre important de points pour obtenir une bonne estimation des paramètres du modèle.

2.2.5 Modèles par champs de Markov

Un champ de Markov est un processus aléatoire sans mémoire, c'est-à-dire que son évolution n'est dictée que par des relations *locales*. La valeur du champ en un point donné n'est donc fonction que de ses voisins et des probabilités associées à ce voisinage : le comportement global du signal est directement fonction des interactions locales.

Du point de vue de l'analyse d'images, on rencontre deux types de solutions : à base de modèles de Markov observables⁸, ou à base de modèles de Markov cachés (*HMM* pour *Hidden Markov Models*) dont on estime les paramètres à partir du signal. Dans un modèle observable, la séquence des observations $O = O_1, O_2, \dots, O_n | O_k \in \{o_1, o_2, \dots, o_p\}$ correspond directement à l'état courant $E = E_1, E_2, \dots, E_n | E_k \in \{e_1, e_2, \dots, e_p\}$ du modèle (o_i est émis par e_i). Dans le cas d'un modèle caché (HMM), chaque état peut engendrer plusieurs observations $O_k \in \{o_1, o_2, \dots, o_p\}$, et la correspondance n'est plus directe. La figure 2.6 illustre cette différence.

Pour cela, une méthode classique est l'utilisation de l'algorithme récursif de Viterbi, utilisé également dans les communications (pour les codes correcteurs d'erreurs) ou en traitement automatique du langage. Une fois les paramètres de texture identifiés, on peut

⁸de manière informelle et pour reprendre le langage de l'informatique, on peut considérer un modèle de Markov comme un automate évoluant de façon aléatoire, les transitions étant les probabilités de passer d'un état au suivant.

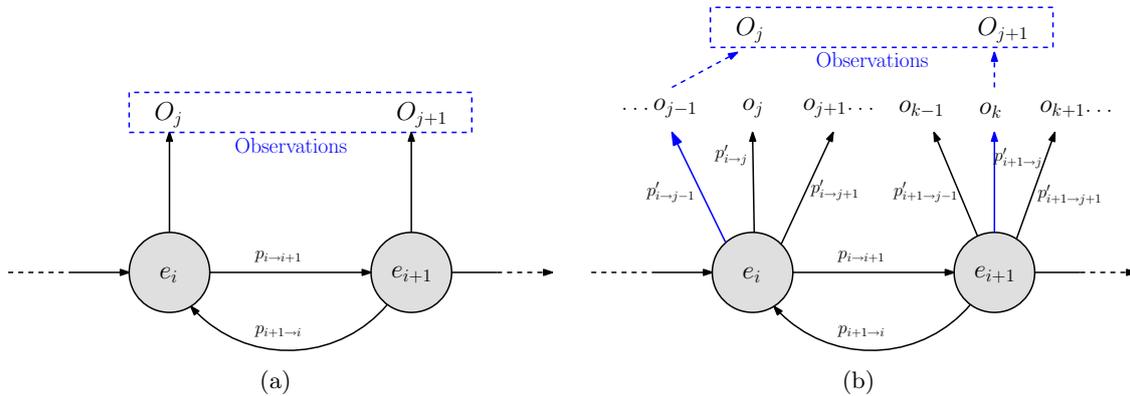


FIG. 2.6 – **Modèle de Markov et HMM.** (a) modèle de Markov observable : à chaque état correspond une valeur d'observation. (b) modèle de Markov caché ; chaque état peut émettre une valeur dans l'alphabet des observations.

réaliser la classification ou la segmentation de l'image, selon ces paramètres⁹. Pour les modèles cachés, il faut utiliser l'algorithme de Baum-Welch¹⁰, qui est décrit en détail dans [Cornuéjols et al., 2003].

On trouve encore beaucoup d'autres modèles de champs de Markov en traitement d'images, comme les champs de Markov gaussiens [Rellier, 2002], les champs de Markov couplés [Pieczynski and Tebbache, 2000], les champs de Markov triplets [Benboudjema, 2005], etc.

Ces approches sont proches de la modélisation par modèle AR, mais sont généralement plus exigeantes en termes de calculs, surtout par rapport aux modèles AR déterministes. De plus, elles partagent avec ces dernières la limitation qu'il faut considérer un nombre important de réalisations pour obtenir une estimation précise.

2.2.6 Autres modèles

D'autres modèles pour l'analyse de texture ont été proposés. Plusieurs sources¹¹ présentent des modèles basés sur la théorie des fractales. Les fractales permettent de représenter des modèles hautement irréguliers, mais possédant une homogénéité globale. Elles sont donc, a priori, bien adaptées à la représentation des textures. Le modèle "2D fractional Brownian motion" (ou *fBm*) a été proposé récemment [Jennane et al., 2007] et semble prometteur. La propriété d'invariance d'échelle est aussi une caractéristique des

⁹l'ouvrage [Winkler, 1991] propose une excellente introduction à la classification de texture par modèles de Markov observables (ainsi qu'une introduction remarquable à l'optimisation par recuit simulé...).

¹⁰c'est une version de l'algorithme *EM* d'estimation de modèles cachés appliquée aux *HMM*. Voir [Cornuéjols et al., 2003] p.403 et p.567 pour plus de détails.

¹¹notamment [Cocquerez et al., 1995].

images fractales. La *dimension fractale* est l'approche classique pour mesurer cette propriété d'invariance, et plusieurs méthodes peuvent être utilisées pour la mesurer, comme décrit dans [Levy Vehel, 1992, Chen et al., 1993, Jingaposao and Anni, 1999, Cocquerez et al., 1995]. Notons que des modèles dits "multifractals" ont été proposés pour améliorer les performances de ces techniques [Levy Vehel et al., 1992].

2.2.7 Discussion

Il existe un grand nombre d'approches au problème de l'analyse de texture. Afin de choisir parmi ces méthodes, il faut étudier les caractéristiques des images considérées :

- la surface est de texture très homogène par endroits ;
- la surface est déformable et sera observée à différentes échelles ;
- la caractérisation doit être la plus "locale" possible (on cherche à caractériser des points dans l'image) ;
- l'analyse doit être rapide pour permettre (à terme) une implémentation temps réel.

2.2.7.1 Approche structurelle

L'approche structurelle a été envisagée, car il s'agit d'une représentation très naturelle des textures. Idéalement, nous espérons être capable d'extraire un motif général (un texton) puis d'effectuer un suivi de ce "réseau" de textons à la surface du cœur. Nous nous sommes heurtés à une série de problèmes pratiques à chacune des étapes de l'algorithme : extraction de texton, reconstruction du maillage, suivi.

Plusieurs méthodes ont été envisagées pour l'extraction du texton :

- l'extraction "manuelle" ;
- l'utilisation de la fonction d'auto-corrélation. C'est une méthode utilisée en traitement du signal (détection d'écho), mais elle ne montre de bons résultats que si les textons sont peu déformés. Nous calculons la fonction d'auto-corrélation par le spectre de Fourier, puis par seuillage nous cherchons les maxima locaux. La figure 2.7 montre des exemples d'extraction de textons par cette méthode sur des textures simples et des textures de cœur.

Cette étude a mis en exergue plusieurs problèmes inhérents à ces méthodes. Tout d'abord, l'extraction automatique du texton est un problème difficile, d'autant plus difficile que l'image est irrégulière. Nous voyons sur la figure 2.7 que dès que la régularité du motif diminue, les pics d'auto-corrélation sont moins marqués. Pour la plupart de nos images de cœur, les résultats renvoyés par cette méthode ne sont pas utilisables. La sélection manuelle (avec ou sans création d'un modèle générique) reste la méthode la plus efficace.

Pour l'extraction, nous utilisons :

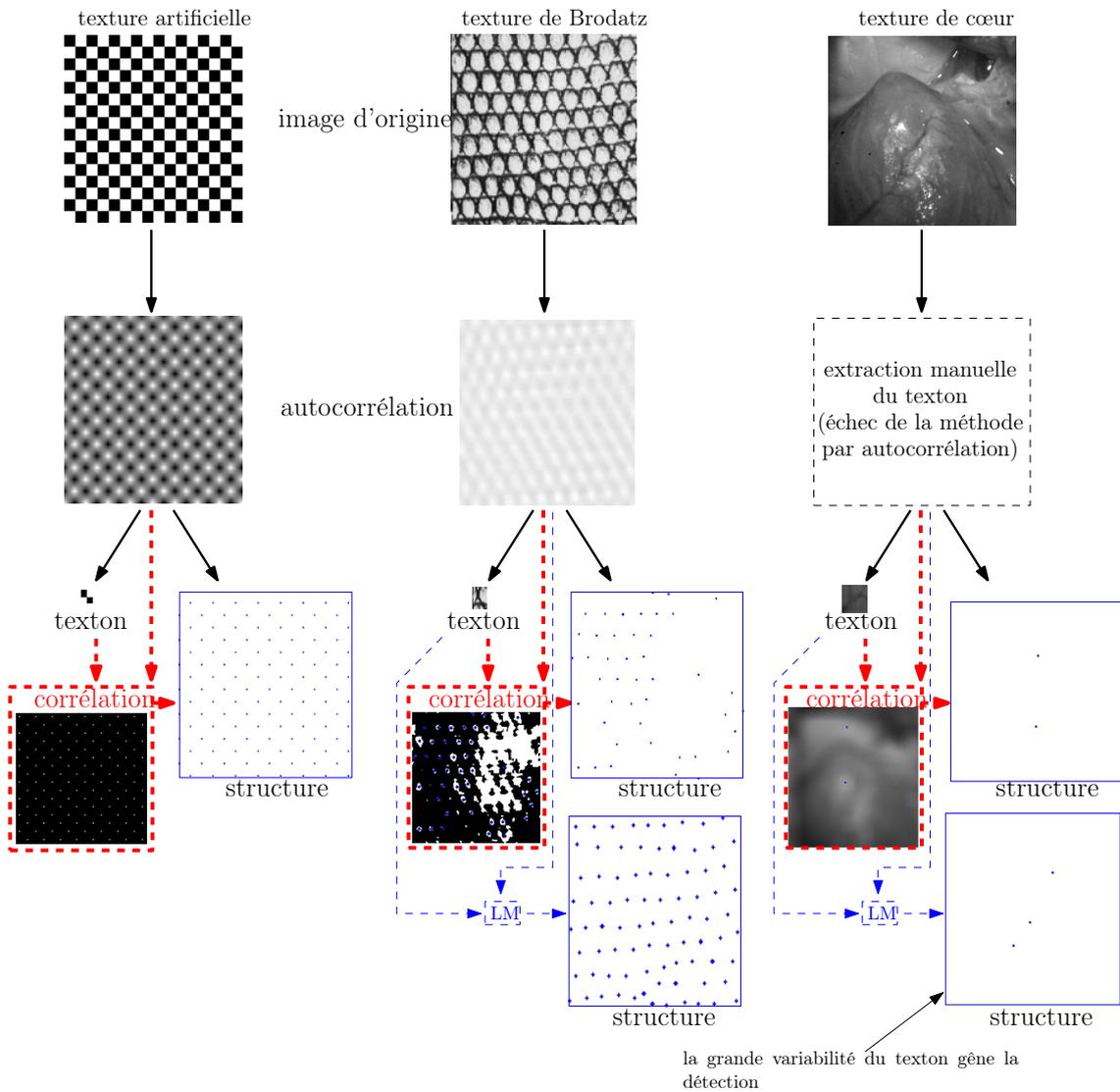


FIG. 2.7 – *Extraction de textures.* La “structure” est représentée par la superposition de la corrélation entre le texton et l’image et des positions extraites (par seuillage et opération morphologique) pour le centre du texton.

- le calcul de la corrélation croisée entre le texton et l'image d'origine, puis le seuillage de cette fonction afin d'extraire les maxima de corrélation (qui correspondent aux instances du motif). Afin d'améliorer la localisation des textons, nous utilisons également une opération d'amincissement morphologique¹². La figure 2.7 donne des exemples d'extraction de textons ;
- nous avons également testé la méthode décrite dans [Loh, 2006], qui consiste à rechercher le texton par la méthode de Levenberg-Marquardt. Cette méthode est très semblable aux méthodes de suivi de motifs présentées au §3.3.2.

L'approche par corrélation est très rapide¹³, et permet d'évaluer toutes les correspondances en une seule passe. Cependant elle est très sensible aux changements d'orientation de la texture : les pics de corrélation deviennent moins marqués et la détection est plus approximative. L'approche de [Loh, 2006] est plus lente : de l'ordre de 4s de calcul par image sous Matlab¹⁴. Techniquement, elle revient à faire une série de recalages suivant un modèle de déformation affine. Son avantage est de pouvoir trouver facilement les textons qui ont "tourné". Cependant, le temps de calcul est rédhibitoire : il n'est pas possible de descendre aux fréquences que nous recherchons (fréquences vidéo et plus) avec ce type de méthode.

Toutes ces considérations nous ont décidé à nous tourner vers d'autres approches. Les méthodes structurelles nous apparaissent comme peu indiquées pour suivre des textures naturelles. Elle peuvent probablement se révéler très intéressantes sur des motifs créés industriellement (tissus par exemple), mais ce type d'application ne relève pas de notre champ d'étude.

Nous nous sommes alors focalisés sur l'extraction de descripteurs locaux, afin de développer une méthode de suivi de type *feature tracker*.

2.2.7.2 Modèle ARMA, champs de Markov et approches spectrales

Ces approches proposent une paramétrisation de la texture, sous la forme d'un jeu de paramètres pour les modèles stochastiques (modèle ARMA et champs de Markov) ou de la décomposition dans une base de fonctions (pour les approches spectrales).

Dans le cas des approches ARMA et de Markov, l'estimation du modèle est d'autant plus précise que le nombre d'observations considérées est grand (il faut une phase "d'apprentissage" de la texture). Nous voulions autant que possible définir des descripteurs

¹²voir [Cocquerez et al., 1995] p.392-394.

¹³1ms en C pour des images 128 × 128 avec la librairie *FFTW* (et même probablement moins, la fonction `time` ne permettant pas d'être plus précise).

¹⁴la toolbox d'optimisation de Matlab étant relativement performante, il paraît illusoire d'espérer gagner plus d'un ordre de grandeur dans le temps de calcul.

locaux : en ce sens cette approche n'est pas idéale. De plus, la représentation obtenue est anisotrope, et, pour les modèles auto-régressifs, le choix d'un voisinage est un problème essentiel et difficile.

2.2.7.3 Les approches spectrales¹⁵

Les approches basées sur la transformée de Fourier souffrent de deux défauts fondamentaux :

- la représentation est adaptée à des motifs très réguliers (même après fenêtrage, le spectre reste bruité et difficile à interpréter) ;
- elle ne porte pas (directement) d'information spatiale.

Le second point est problématique : cette représentation est surtout utile pour filtrer le signal. Par exemple, si la texture est très localisée en fréquence, l'utilisation d'un (ou généralement d'une combinaison de) filtre de Gabor permet de ne conserver que la partie correspondante de l'image. Cependant cette approche n'est pas adaptée à la localisation d'un motif.

Les ondelettes posent le même problème, même si des approches telles que les histogrammes d'ondelettes [Heeger and Bergen, 1995] pourraient peut-être aller dans le sens d'une caractérisation locale des textures. Cependant, une grande partie des méthodes par ondelettes se focalisent plus sur les problèmes de synthèse de texture que sur les problèmes d'analyse.

2.2.7.4 Modélisation statistique

Les modèles statistiques présentent, selon nous, plusieurs avantages par rapport aux approches précédemment citées :

- elles font peu d'hypothèses sur la nature de l'image ;
- il existe un grand nombre d'attributs texturaux statistiques dans la littérature ;
- elle peuvent être utilisées sur des régions de taille quelconque ;
- la représentation obtenue, sous la forme d'un vecteur de descripteurs, est comparable aux méthodes basées sur des descripteurs déjà utilisées en traitement d'image et en vision [Lowe, 1999].

De plus, les principales méthodes d'analyse de texture peuvent être rendues isotropes en moyennant les valeurs obtenues selon plusieurs directions (par exemple en calculant les matrices de co-occurrences pour $\theta = 0^\circ, 45^\circ, 90^\circ, \dots, 315^\circ$, puis la matrice de co-occurrence moyenne, et en calculant les attributs texturaux sur cette nouvelle matrice "isotrope").

¹⁵nous ne considérons pas ici la génération d'attributs statistiques sur les représentations spectrales, qui sera englobée dans le paragraphe **Modélisation statistique**.

2.2.7.5 Conclusion

Suite à ces considérations, nous avons choisi de nous concentrer sur les méthodes statistiques, plus adaptées à notre objectif. Il est certain que les approches structurales, que nous laissons ainsi de côté, possèdent des avantages : elle permettent de générer de nouvelles textures ou encore d'utiliser des estimateurs statistiques (typiquement le Maximum A Posteriori), très efficaces en segmentation. Cependant, ces avantages nous apparaissent secondaires au vu de notre application, et nous préférons la souplesse de l'approche statistique (couplée à un choix judicieux des descripteurs).

Dans la section suivante, nous présenterons une étude de divers attributs texturaux. Cette étude a pour but de réduire l'espace des descripteurs afin d'améliorer la représentativité et le temps de calcul du vecteur d'attributs texturaux, que nous utiliserons au chapitre 3 pour réaliser le suivi.

2.3 Sélection d'attributs texturaux

2.3.1 Objectifs

Le paragraphe 2.2 a montré qu'il existe de nombreux attributs texturaux dans la littérature. Vu l'application visée (le suivi de cœur battant) et ses contraintes (le temps réel) il nous était impossible d'utiliser un nombre trop important dans cette application. Par contre, la sélection d'un petit nombre d'attributs décorrélés permet d'augmenter la qualité de la représentation obtenue, réduisant le bruit, et en limitant l'effet d'échelle dû aux attributs corrélés (qui rendent compte du même phénomène).

La figure 2.8 résume la procédure de sélection utilisée dans notre étude. Dans un premier temps nous utilisons une *Analyse en Composantes Principales* pour réduire l'espace de variables en éliminant la corrélation et en maximisant la variance portée par ces variables. Puis nous utilisons l'*analyse discriminante* de Fisher pour étudier le pouvoir discriminant de ces variables.

2.3.2 Génération de la banque d'images-test

Afin de comparer les solutions disponibles, il nous fallait disposer d'un jeu de données représentatives des régions d'intérêt que nous cherchons à discriminer. Il existe plusieurs jeux de textures utilisées par la communauté pour comparer leurs outils. Les plus utilisées sont les textures de *Brodatz* [Noce et al., 2006], issues d'un livre destiné aux artistes, et la base *VisTex*, disponible librement sur internet, qui propose à la fois des textures de référence et des scènes texturées pour évaluer les algorithmes de segmentation. Nous avons également généré plusieurs séries de textures artificielles répétitives.

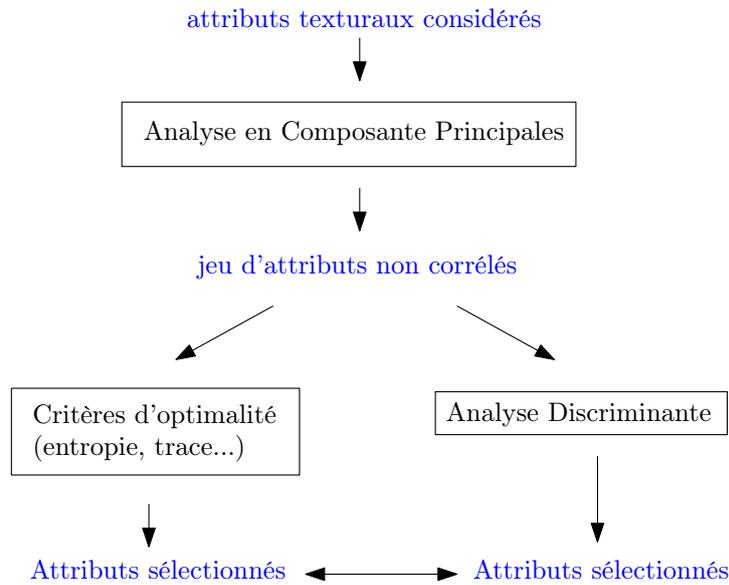


FIG. 2.8 – *Étapes de la procédure de sélection d'attributs.*

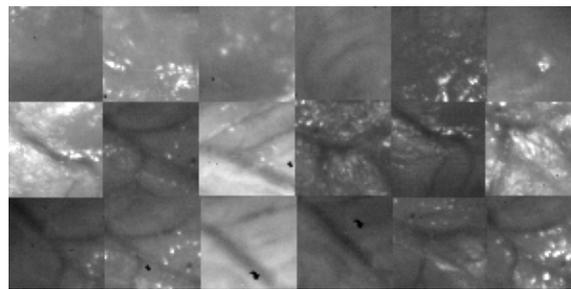
Cependant, nous avons concentré nos efforts sur plusieurs séries d'images de cœur battant réalisées lors d'expérimentations in vivo. Il s'agit d'images prises à haute vitesse (jusqu'à 300Hz) grâce à une caméra rapide DALSA 1m75 et un logiciel spécialement développé. Afin de pouvoir disposer d'une base de textures de cœur conséquente, nous avons extrait manuellement des échantillons de textures classées en plusieurs catégories : faisceaux artériel (classées par tailles et par régions du cœur d'où ils étaient extraits), différentes zones de la surface du cœur. Comme nous avons plusieurs séries prises sur des cœurs différents (deux à l'époque de l'article [Noce et al., 2006]) nous avons pu créer une base mixte. Notre base contenait initialement 200 images, puis fut enrichie de 100 nouvelles images par la suite¹⁶.

Notre approche se concentre principalement sur les textures de cœur. Les autres bases n'ont été utilisées qu'au début et à la fin de l'expérimentation :

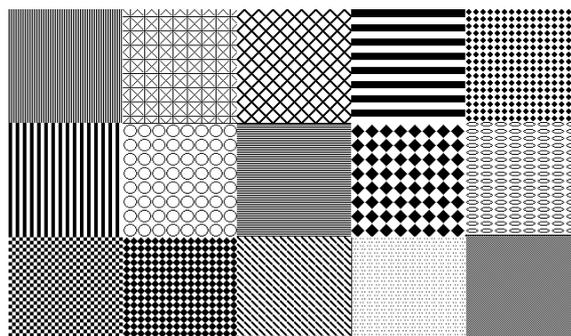
- dans un premier temps pour l'analyse en composantes principales afin de conserver un ratio Attributs/Images faible ;
- après la sélection des attributs pour étudier la qualité intrinsèque de notre jeu d'attributs.

La figure 2.9 présente des textures extraites de ces différentes banques.

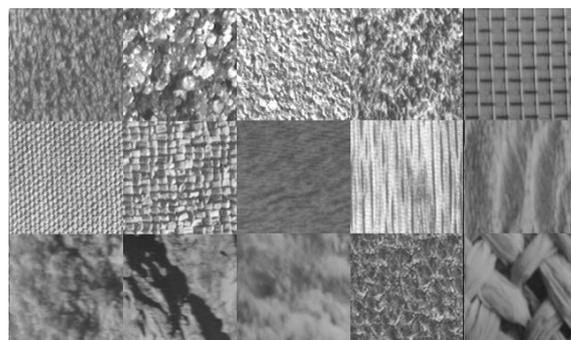
¹⁶lorsque nous avons pu assister à une nouvelle expérimentation.



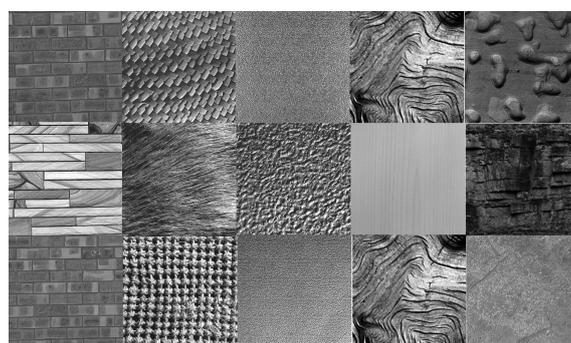
(a)



(b)



(c)



(d)

FIG. 2.9 – *Images extraites de nos banques d'images-tests. (a) textures de cœur, (b) textures artificielles, (c) textures de Brodatz, (d) textures Viztex.*

	textures de cœur	textures artificielles	textures de Brodatz	textures VisTex
moyenne	0,091	0,109	0,095	0,095
médiane	0,09	0,1	0,09	0,1
maximum	0,19	0,17	0,17	0,25
minimum	0,04	0,06	0,06	0,08

TAB. 2.2 – *Test KS de normalité des variables : valeur de α .*

2.3.3 Étude préliminaire

Avant d'appliquer les méthodes d'analyse statistique multidimensionnelle, nous devons vérifier que les variables sont normales (ou proches d'une distribution normale). Pour cela, nous utilisons un test statistique, **le test de Kolmogorov-Smirnov**, qui fonctionne en définissant une variable gaussienne (de distribution cumulative F_0) de même moyenne et écart-type que la variable considérée (dont la fonction de distribution \hat{F} est déterminée expérimentalement), et en comparant les deux distributions. Pour un échantillon de taille n , on considère que le seuil de confiance α est atteint si :

$$\max_x |\hat{F}(x) - F_0(x)| < \sqrt{\frac{-1}{2n} \log \frac{\alpha}{2}}$$

Ce test a montré qu'en règle générale, on peut considérer les distributions comme gaussiennes avec un risque d'erreur de l'ordre de 10%. Le tableau 2.2 donne les minima, maxima, moyenne et médiane du seuil de confiance¹⁷, pour chacune des banques d'images-tests. On remarque que le maximum se situe autour de 20% d'erreur, mais il ne concerne que quelques attributs isolés (la médiane est autour de 10%).

Ces résultats sont satisfaisants et valident l'approche statistique des paragraphes suivants.

2.3.4 Analyse en composantes principales

2.3.4.1 Principe

L'*analyse en composantes principales* (ou *ACP*) est une méthode de statistique descriptive multidimensionnelle permettant de mesurer et de visualiser les corrélations existantes entre plusieurs séries de mesures. En un sens, l'ACP généralise la notion de coefficient de corrélation utilisée en statistique classique pour comparer deux phénomènes et extraire une relation linéaire entre ceux-ci. Cette méthode est principalement utilisée pour la réduction de la dimension de l'espace originel, lorsque celle-ci est trop importante et que les variables qui décrivent cet espace sont redondantes. On utilise également l'ACP dans un

¹⁷le script utilisé réalise le test KS pour les valeurs de α incrémentées de 0,01 (soit 1%) à chaque itération, et retient la plus petite valeur qui passe le test.

but descriptif, puisqu'elle fournit, d'un ensemble étendu de données, une représentation graphique et une information synthétique sur l'organisation des individus en classes.

Si l'on considère n variables aléatoires (correspondant ici aux images-tests) et m réalisations conjointes de ces variables (les valeurs des attributs texturaux), les données peuvent être vues comme les coordonnées de n points dans un espace vectoriel à m dimensions, dans la base canonique de cet espace. La principale opération de l'ACP est de déterminer les axes principaux d'inertie du nuage autour de son centre de gravité. Ce sont les axes qui prennent le mieux en compte la dispersion du nuage au sens de la distance euclidienne \mathcal{R}_p .

Le critère que cherche à maximiser l'ACP est la *variance projetée* sur ces axes. Ainsi, en réduisant l'espace initial aux N premiers axes factoriels, on obtient le sous-espace qui porte la plus grande inertie possible. C'est un très bon critère pour l'application que nous visons, car nous cherchons les attributs les plus discriminants possibles.

L'ACP impose au départ la définition d'une métrique euclidienne \mathcal{R}_p . Étant donné que les variables sont hétérogènes, nous les avons d'abord centrées et réduites. Ceci rend les résultats de l'analyse indépendants des unités de mesure des variables et permet de réduire considérablement l'*effet taille* de l'ACP sur les variables.

Classiquement, on observe les données issues de l'ACP selon deux représentations privilégiées :

- **la projection selon deux axes principaux**, qui permet une représentation aisée des relations entre les individus ainsi qu'une analyse plus facile des relations entre ces individus ;
- **le cercle de corrélation**, défini pour deux axes principaux, qui permet d'analyser la corrélation entre les variables. Si celles-ci apparaissent "loin" de 0 et proches l'une de l'autre, alors elles sont significativement corrélées. C'est un outil très pratique pour éliminer l'information redondante.

2.3.4.2 Mise en oeuvre sur l'ensemble des bases d'images-tests

Nous avons tout d'abord programmé nos attributs texturaux : 117 au total. Ce jeu comprend plusieurs attributs similaires tels que des matrices de co-occurrence avec des paramètres d'orientation et de distance différents. Dans un premier temps, nous avons évalué ces attributs sur un jeu de textures très large comprenant simultanément toutes les bases de texture. Le but était d'éliminer les attributs les moins discriminants¹⁸.

Nous avons d'abord centré et réduit nos variables :

¹⁸il ne serait pas cohérent d'évaluer 117 attributs avec une banque d'images comprenant seulement 300 images-tests. Toutes banques confondues, notre jeu s'élève à 600 images ce qui est plus réaliste.

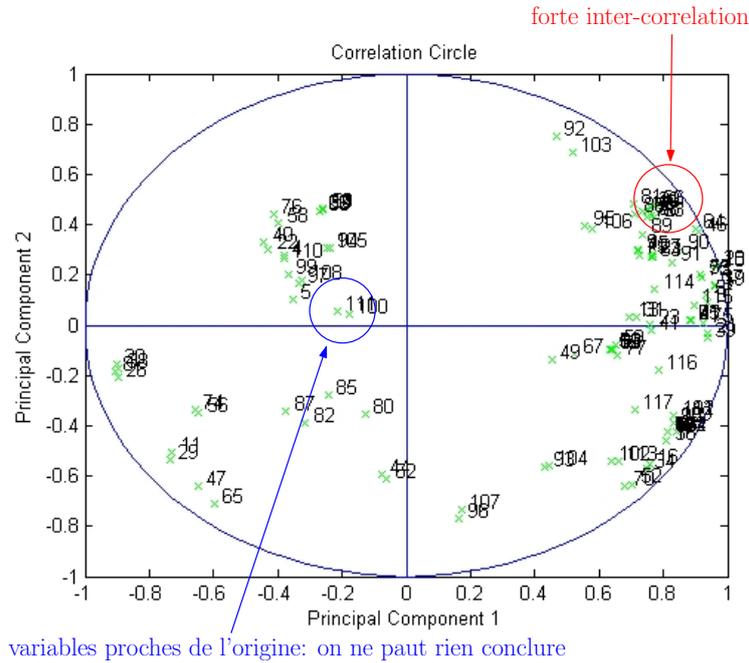


FIG. 2.10 – **Cercle de corrélation.** Cercle de corrélation des deux premières composantes principales. On remarque la forte corrélation entre certains des attributs originaux.

$$\mathbf{M} = \begin{bmatrix} \frac{X_{1,1}-\bar{X}_1}{\sigma(X_1)} & \dots & \frac{X_{1,m}-\bar{X}_1}{\sigma(X_1)} \\ \vdots & \ddots & \vdots \\ \frac{X_{n,1}-\bar{X}_n}{\sigma(X_n)} & \dots & \frac{X_{n,m}-\bar{X}_n}{\sigma(X_n)} \end{bmatrix}, \tag{2.4}$$

avec pour chaque variable aléatoire $X_k = (X_{k,1}, \dots, X_{k,m})$, \bar{X}_n sa moyenne et σ_{X_n} son écart type. Puis nous réalisons l'ACP en diagonalisant la matrice d'inter-corrélation¹⁹ :

$$C = M^t M$$

Chaque vecteur propre correspondant à un axe factoriel, l'ordre des axes est fixé par la valeur propre associée (le premier axe factoriel est associé à la valeur propre la plus élevée, etc.).

La sélection est faite en étudiant la *contribution des variables* aux axes factoriels, c'est-à-dire le produit scalaire entre le vecteur directeur de l'axe factoriel et la variable. Nous avons ainsi pu facilement réduire le nombre de nos attributs à 40, une grande disparité existant entre les meilleurs attributs et les plus mauvais.

¹⁹si nous n'avions pas centré et réduit, nous aurions ici la matrice de variance-covariance, classique en automatique.

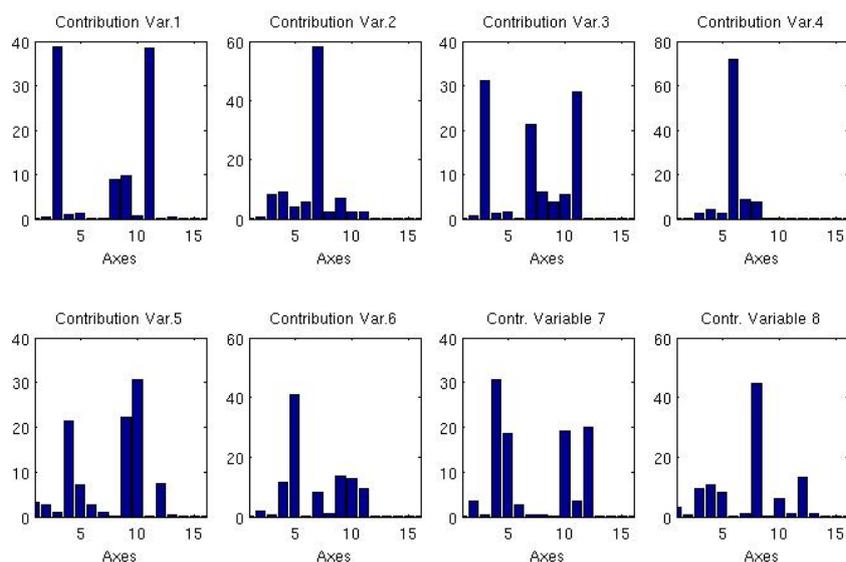


FIG. 2.11 – *Contribution aux axes des 8 meilleurs attributs texturaux, en %.*

2.3.4.3 Mise en œuvre sur les textures de cœur

Nous avons réalisé une nouvelle ACP, sur les textures de cœur uniquement. Nous avons alors utilisé conjointement le critère de contribution aux axes factoriels et les *cercles de corrélation*. La procédure mise en œuvre est alors la suivante :

1. pour chaque couple d'axes principaux, tracer le cercle de corrélation correspondant ;
2. sélectionner alors un groupe d'attributs proches les uns des autres et proches du cercle. Si des attributs apparaissent proches de 0, ils ne peuvent être analysés et il faut utiliser une autre projection ;
3. pour chaque attribut, on évalue sa contribution à la variance, c'est-à-dire la variance de la série projetée sur l'axe. Nous étudions également la corrélation à partir des cercles de corrélation, comme illustré sur la figure 2.10 ;
4. on choisit les meilleurs attributs du groupe en éliminant les moins significatifs ;
5. on change de projection et on continue.

La figure 2.10 résume la méthode d'interprétation d'un cercle de corrélation. Par cette méthode, nous avons réduit le jeu d'attributs jusqu'à obtenir 16 attributs, peu corrélés et représentant bien les 16 premiers axes principaux. La figure 2.11 illustre la contribution de 8 variables aux axes principaux.

La dernière question était de déterminer le nombre d'attributs minimum nécessaires pour une discrimination optimale. En effet, le nouveau jeu de 16 attributs était déjà

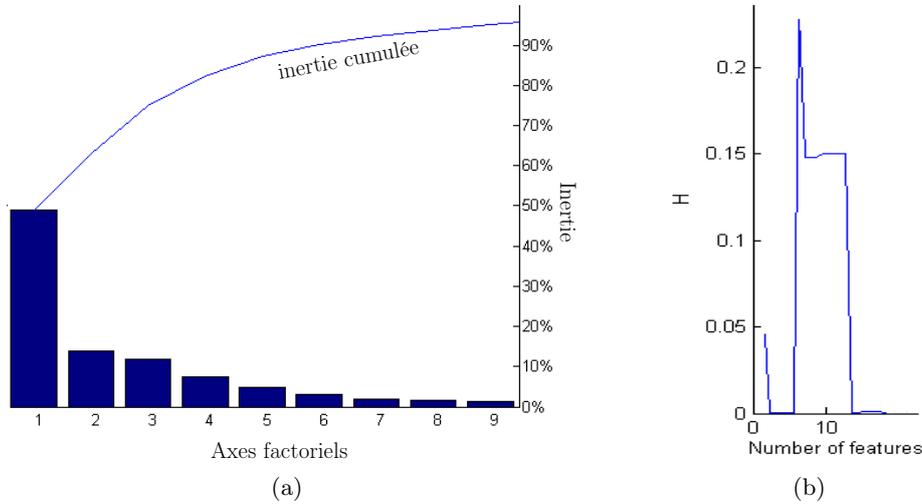


FIG. 2.12 – *Détermination du nombre minimum d'attributs à retenir.* (a) inertie cumulée portée par les premiers axes principaux (en %), (b) Entropie en fonction du nombre d'attributs.

beaucoup plus restreint que le jeu initial, mais pouvait-on le réduire encore ? Plusieurs critères peuvent être utilisés afin de décider du nombre d'axes à conserver :

- **le critère de Kaiser** qui préconise de retenir les composantes principales dont les valeurs propres sont supérieures à 1^{20} . Ce critère nous incite à conserver 7 axes factoriels
- **l'étude de l'inertie** (cumulée) portée par les axes factoriels. Conserver les 8 premiers axes factoriels permet de conserver plus de 95% de l'inertie totale, comme illustré sur la figure 2.12a, ce qui valide l'estimation fournie par le critère de Kaiser.
- **l'étude de l'entropie** : l'entropie se calcule également à partir des valeurs propres de la matrice d'inter-corrélation, en analysant les valeurs λ_i de cette matrice (classées de la plus forte à la plus faible : $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$)

$$H = \bar{\lambda}_i \log \bar{\lambda}_i \quad (2.5)$$

$$\text{avec } \bar{\lambda}_i = \frac{\lambda_i}{\sum_{j=1}^n \lambda_j}.$$

Elle renseigne sur l'indépendance de l'information portée par les axes. Comme l'illustre la figure 2.12b, la meilleure entropie est obtenue sur les 8 premiers axes factoriels.

En suivant les indications des critères de Kaiser, de l'inertie et de l'entropie, nous avons étudié les contributions aux 8 premiers axes factoriels (de manière analogue à la méthode présentée figure 2.11). Nous en avons déduit 12 attributs texturaux intéressants, présentés

²⁰la limite 1 n'est valable que pour des données centrées réduites !

Approche	Jeu de 12 attributs	Jeu de 8 attributs
Matrice de co-occurrence	Énergie	Énergie
Matrice de co-occurrence	Contraste	Contraste
Matrice de co-occurrence	Entropie	Entropie
Matrice de co-occurrence	Cluster Prominence	Cluster Prominence
Matrice de co-occurrence	Corrélation	
Matrice de co-occurrence	Homogénéité	
Méthode des différences de NG	Entropie	
Matrice de longueur des plages	Non Uniformité	Non Uniformité
Matrice de longueur des plages	HGRE	HGRE
Matrice de longueur des plages	SRE	
Statistiques	Symétrie	Symétrie
Statistiques	Kurtosis	Kurtosis

TAB. 2.3 – *Jeu d'attributs sélectionnés par ACP. Les deux dernières étapes de sélection des attributs (jeu de 12 descripteurs puis jeu final de 8 descripteurs) sont représentés. (voir l'annexe pour la description des attributs).*

dans le tableau 2.3.

Cependant l'espace discriminant n'étant que de dimension 8 (il y a huit axes factoriels), nous aimerions nous limiter à huit attributs. Pour cela, nous appliquons le **critère de la trace**²¹ :

$$\max_{\mathbf{v}}(\text{trace}(T_{\mathbf{v}}^{-1}B_{\mathbf{v}})) \quad (2.6)$$

où \mathbf{v} est un jeu de variables (pris ici parmi tous les ensembles possibles de 8 attributs texturaux), $T_{\mathbf{v}}$ et $B_{\mathbf{v}}$ les matrices de variance totale et de variance inter-classes correspondant à ce jeu d'attributs.

Ce critère est combinatoire, c'est-à-dire qu'il faut tester toutes les possibilités pour déterminer les meilleures, attributs. Il n'est donc pas envisageable de l'utiliser sur un jeu d'attributs trop important, car le temps de calcul devient exponentiel²². Finalement, les huit attributs retenus sont présentés sur le tableau 2.3.

²¹il existe d'autres critères, comme le critère de *Wilks*, pour la discrimination à n classes (et encore plus si nous avons considéré le cas à deux classes). Nous avons choisi le critère de la trace car il est très utilisé et son efficacité est démontrée en pratique.

²²notons cependant que l'article Nasri and El Hitmy [2002] présente une méthode basée sur les algorithmes génétiques pour calculer un jeu d'attributs optimal pour ce critère dans un temps raisonnable.

2.3.5 Analyse discriminante

2.3.5.1 Principe

Pour affiner la sélection des attributs texturaux, nous avons utilisé une seconde méthode d'analyse factorielle : l'*Analyse Discriminante*. Celle-ci est très proche de l'ACP, mais optimise la séparabilité des classes : les "groupes" de textures similaires apparaissent plus éloignés après la transformation.

C'est une méthode utilisée classiquement en classification automatique afin de réduire l'espace des descripteurs. Dans notre approche, son utilisation est motivée par plusieurs considérations :

- le problème de classification est proche du problème d'appariement, central en stéréovision ;
- cette approche donne un autre point de vue (en termes d'optimalité) sur nos attributs, et permet donc de comparer nos résultats avec ceux obtenus en ACP ;
- elle est proche du critère de la trace, qui est un bon étalon pour l'efficacité des variables en termes de discrimination.

2.3.5.2 Mise en œuvre

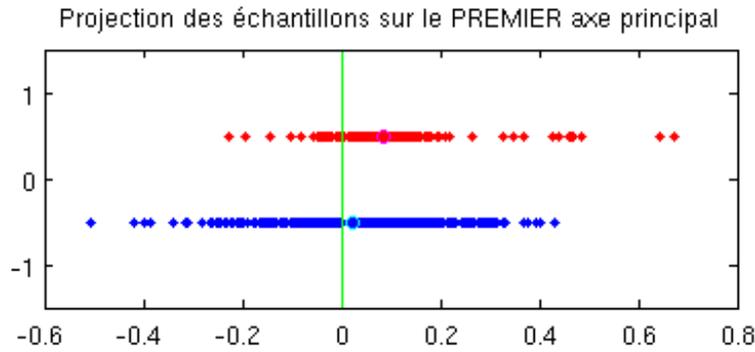
Nous avons appliqué l'analyse discriminante sur les attributs décorrélés par ACP. En effet, cette méthode ne prend pas en compte la corrélation entre les individus et sans cette précaution nous pourrions nous retrouver après sélection avec plusieurs variables équivalentes.

Soit T , B et W les matrices de variance totale, variance interclasse et variance intra-classe associées aux individus, b une forme linéaire sur \mathbb{R}^p . Sachant que $T = W + B$, on a :

$$b^t T b = b^t B b + b^t W b. \quad (2.7)$$

On voit alors qu'il est équivalent de chercher à éloigner les différentes classes (minimiser $\frac{b^t B b}{b^t T b}$) ou de chercher à rapprocher les individus d'une même classe (maximiser $\frac{b^t W b}{b^t T b}$). L'analyse cherche b ayant une variance égale à 1, et le problème s'énonce ainsi : maximiser $b^t B b$ sous contrainte $b^t T b = 1$. De manière analogue à l'ACP, la solution est fournie par les plus grandes valeurs propres λ_i de $T^{-1}B$. Les formes linéaires discriminantes b_i sont alors les vecteurs propres associés :

$$T^{-1} B b_i = \lambda_i b_i, \quad (2.8)$$

FIG. 2.13 – *Résultats de l'analyse discriminante.*FIG. 2.14 – *Classes utilisées pour l'analyse discriminante. (a) coronaires, (b) surface de cœur.*

et les axes factoriels discriminants sont définis par les vecteurs u_i :

$$u_i = Tb_i. \quad (2.9)$$

Cependant, contrairement à l'ACP, il y a un nombre restreint d'axes discriminants : $k - 1$, où k est le nombre de classes utilisées pour la classification²³. Dans notre cas, nous avons défini 2 classes, illustrées sur la figure 2.14, afin de différencier les zones où une artère était présente des autres. On n'a donc qu'un axe discriminant²⁴, ce qui représente une réduction d'espace beaucoup trop forte. La figure 2.13 illustre la répartition des individus selon cet axe factoriel²⁵ : on voit un "décalage" des nuages mais pas de séparation nette, ce qui montre une probable difficulté pour discriminer les deux ensembles, avec certains individus très proches.

Nous pouvons étudier la contribution des différents attributs à cet axe, afin de déterminer les attributs les plus discriminants. Si nous choisissons les 8 attributs les plus discriminants, nous retrouvons des résultats (tableau 2.4) très proches de ceux obtenus en

²³cela vient du fait que B est de rang k .

²⁴qui correspond au critère de Fisher.

²⁵l'axe est horizontal, l'axe vertical ne servant qu'à mieux visualiser les points.

Approche	Jeu de 8 attributs	Jeu de 8 attributs (précédemment sélectionnés par ACP)
Matrice de co-occurrence	<i>Énergie</i>	<i>Énergie</i>
Matrice de co-occurrence	<i>Contraste</i>	<i>Contraste</i>
Matrice de co-occurrence	Cluster Shade	Entropie
Matrice de co-occurrence	Corrélation	Cluster Prominence
Matrice de longueur des plages	<i>Non Uniformité</i>	<i>Non Uniformité</i>
Matrice de longueur des plages	SLGLRE	HGRE
Statistiques	Skewness	Symétrie
Statistiques	<i>Kurtosis</i>	<i>Kurtosis</i>

TAB. 2.4 – *Jeu d’attributs sélectionnés par AD*. Sont représentés en italique les attributs communs des deux approches.

ACP (tableau 2.3). Cela est dû à la qualité intrinsèque de certains descripteurs, définitivement supérieurs sur les textures de cœur.

2.4 Discussion

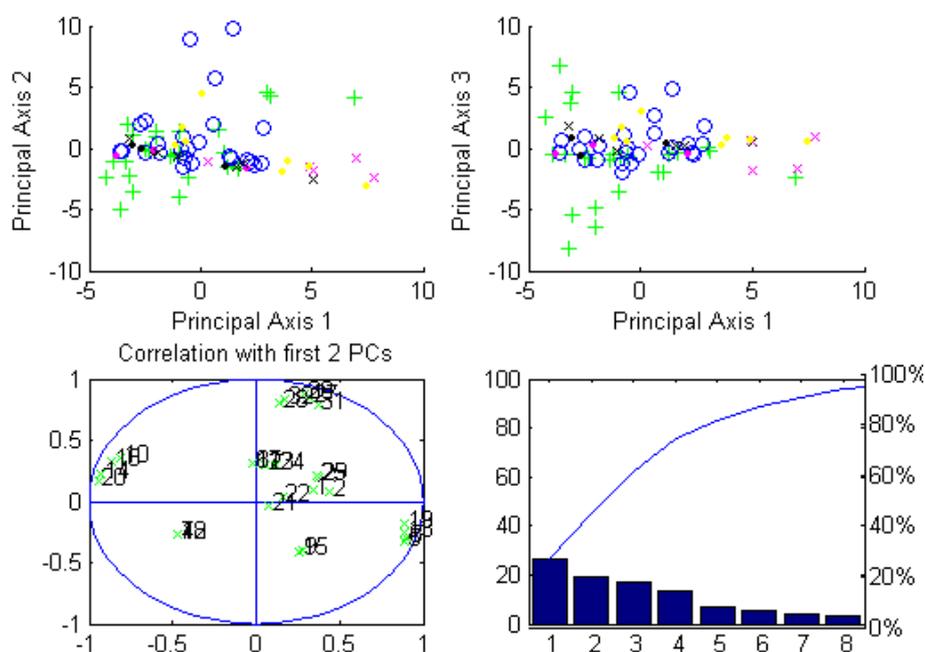
Notre approche a pris le parti de réduire au maximum le nombre d’attributs tout en conservant un maximum d’inertie. Cela permet d’améliorer le temps de calcul, tout en conservant l’information contenue dans les attributs. Nous avons volontairement centré notre étude sur les textures de cœur, afin d’obtenir une représentation optimisée pour nos textures.

Cependant, comme nous disposions d’autres banques d’images, nous avons décidé de les utiliser afin de comparer les résultats.

Autres jeux de textures

Nous avons effectué une nouvelle classification en utilisant les trois autres jeux de texture dont nous disposons. La figure 2.15 présente quelques illustrations de cette analyse :

- deux projections des images selon des *plans factoriels*. Il s’agit d’une représentation des images-tests (un point par image, les différents types de rond représentant des textures de Brodatz et les différents types de croix des textures de la base VisTex) dans le plan formé par deux axes principaux ;
- le cercle de corrélation des variables, très similaire à celui de la figure 2.10 ;
- l’histogramme cumulé de l’inertie des 8 premiers axes. Nous constatons que dans ce cas il faut également 8 axes pour représenter 95% de l’inertie totale du nuage.

FIG. 2.15 – *ACP sur les autres banques d'images-test.*

Après sélection selon la même méthode que précédemment, nous obtenons le jeu d'attributs du tableau 2.5. On remarque la prédominance de l'approche par matrices de co-occurrence, qui sont donc les grandes gagnantes de cette étude. D'autres approches, comme les matrices de longueur des plages, semblent moins efficaces sur ces textures.

Ainsi, si certains attributs restent très efficaces, il semble intéressant de développer un jeu d'attributs spécifiques à chaque application, afin d'obtenir une discrimination suffisamment efficace. En un sens, ce choix modélise les propriétés visuelles des textures et fournit une connaissance a priori.

Approche	Jeu de 8 attributs
Matrice de co-occurrence	Énergie
Matrice de co-occurrence	Contraste
Matrice de co-occurrence	Corrélation
Matrice de co-occurrence	Cluster Prominence
Matrice de co-occurrence	Contraste
Matrice de longueur des plages	Non Uniformité
Matrice de longueur des plages	RPC
Statistiques	Variance

TAB. 2.5 – *Autres attributs sélectionnés par ACP.*

Approche	Jeu de 8 attributs
Matrice de co-occurrence	Énergie
Matrice de co-occurrence	Contraste
Matrice de co-occurrence	Cluster Shade
Matrice de co-occurrence	Cluster Prominence
Matrice de longueur des plages	Non Uniformité
Matrice de longueur des plages	SLGLRE
Statistiques	Skewness
Statistiques	Kurtosis

TAB. 2.6 – *Jeu d’attributs sélectionnés.*

2.4.0.3 Jeu d’attributs finalement retenu

Nous avons finalement choisi 8 attributs, représentés sur le tableau 2.6, en considérant les attributs les plus performants de l’analyse en composantes principales et de l’analyse discriminante.

Une autre possibilité était de garder les deux jeux de descripteurs : le jeu d’attributs issu de l’ACP pour le suivi (car il concentre la variance) et le jeu issu de l’analyse discriminante pour l’appariement en stéréo-vision (car il différencie mieux les différents types de textures). Cependant le gain en terme de robustesse n’est pas évident en pratique, et cette approche nous oblige à calculer plus d’attributs (lorsque qu’on couple le suivi et la reconstruction). Nous avons donc décidé de n’utiliser que le jeu d’attributs présenté sur le tableau 2.6 dans les chapitres suivants.

2.5 Conclusion

Dans ce chapitre, nous avons recensé les principales approches d’étude des textures disponibles dans la littérature. Ces méthodes ont un large spectre d’applications, mais paradoxalement leur utilisation à des fins de suivi de cœur est une idée originale qui implique une étude préalable. Des deux méthodes qui ont été envisagées, méthode structurelle et statistique, seule la dernière s’est révélée être adaptée et a été retenue pour ce projet. Nous avons également choisi une méthodologie pour la réduction de dimension, afin d’isoler les attributs les mieux adaptés à notre application et obtenir une caractérisation à la fois rapide, car il n’y a que peu d’attributs à calculer, et plus précise, car on élimine les bruits apportés par les attributs moins performants. Le chapitre suivant présente la problématique de suivi de mouvements dans les images, ainsi que la méthode basée sur la texture que nous avons développée puis évaluée par rapport aux méthodes classiques dans l’optique de notre application.

Chapitre 3

Suivi d'indices visuels dans les séquences de cœur

3.1 Introduction

Afin de commander un robot en utilisant un système de vision (avec une, deux ou n caméras) il est nécessaire de pouvoir extraire des images issues du capteur un certain nombre de caractéristiques, dont la nature peut être très hétéroclite : points, primitives géométriques, informations de forme ou de texture, etc. Certains travaux utilisent une combinaison de ces différentes informations afin d'obtenir un résultat plus précis.

Cette étape de *traitement d'image* est fondamentale car elle est le socle sur lequel se fonde le système de vision. Dans notre application, la principale difficulté était de suivre les mouvements du cœur malgré ses déformations et la rapidité de ses mouvements. Nous proposons dans ce chapitre une solution basée sur l'analyse de texture afin d'assurer un suivi des mouvements de l'organe. Dans la suite de ce chapitre, le paragraphe 3.2.1.3 introduit les notions fondamentales du suivi d'objets dans les images en mouvement. Puis le paragraphe 3.3 présente les principales approches actuellement utilisées en suivi. Enfin le paragraphe présente notre approche, qui sera évaluée au paragraphe 3.5.

3.2 Concepts fondamentaux

3.2.1 Cadre théorique

Nous cherchons à reconstruire l'état $\mathbf{e}(t)$ de l'objet à l'instant t . Typiquement, on a, si l'on cherche à suivre seulement un (ou des) point(s) dans l'image¹ :

$$\mathbf{e}(t) = \begin{bmatrix} x(t) \\ y(t) \end{bmatrix} \quad (3.1)$$

et dans le cas d'un point dans une scène en trois dimensions :

$$\mathbf{e}(t) = \begin{bmatrix} x(t) \\ y(t) \\ z(t) \end{bmatrix}. \quad (3.2)$$

Souvent, en vision, on cherchera par des méthodes de reconstruction de pose à reconstruire la position 3D d'un objet connu par CAO. Dans ce cas, le vecteur d'état à au moins 6 composantes².

Pendant que l'objet est en mouvement, d'autres paramètres du système de suivi évoluent et doivent être pris en compte. C'est le cas du système d'observation. Chaque valeur de l'état $\mathbf{e}(t)$ est extraite d'une observation $\mathbf{o}(t)$. Or la méthode d'observation influe directement sur la méthode de suivi : on ne suit pas de la même façon un objet dans une scène vidéo, une image échographique ou une image catadioptrique. De plus, il y a toujours des erreurs d'observation qui influent sur la précision du suivi et peuvent avoir des conséquences importantes sur l'estimation de l'état $\mathbf{e}(t)$. Ces erreurs peuvent être filtrées ou compensées.

Enfin, le suivi peut prendre en compte la composante temporelle. Il est classique, par exemple, d'utiliser un modèle bayésien du mouvement. Dans ce cas un filtre de Kalman (ou d'autres méthodes telles que les filtres particulaires) permet d'estimer l'état en prenant en compte l'historique des observations. Cela permet d'induire un comportement prédictif sans posséder un modèle complet du mouvement. La figure 3.1 résume les différents modèles considérés.

3.2.1.1 Modèles d'observation usuels

Plusieurs modèles d'observation sont envisageables :

¹c'est notre cas, car nous avons développé des "blocs" indépendants acquisition-suivi-reconstruction. C'est un choix qui à notre avis favorise la réutilisabilité, mais d'autres approches existent, telles que [Benhimane and Malis, 2007].

²typiquement la position du centre de masse et les angles d'Euler définissant l'orientation du modèle

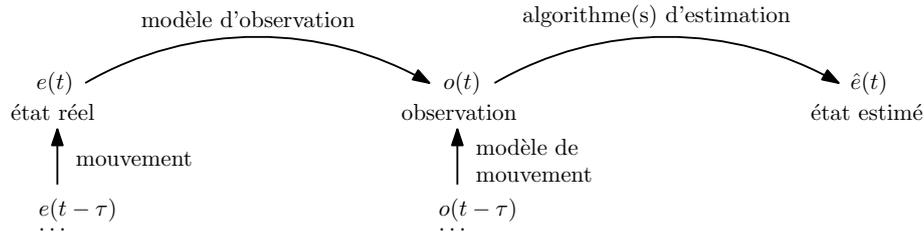


FIG. 3.1 – **Modèle pour le suivi de mouvement.** De l'état réel $e(t)$ à l'état estimé $\hat{e}(t)$, en passant par le modèle d'observation $o(t)$. Cette représentation permet de séparer ce qui est dû à l'acquisition (bruit de mesure, de numérisation) et ce qui est dû à l'estimation (conditionnement, biais).

- soit on représente l'objet comme l'ensemble de ses pixels dans l'image. C'est l'approche classique en *pattern matching* [Hutchinson et al., 1996, Malis, 2004];
- soit on extrait de l'image une série de points caractéristiques. C'est l'approche utilisée dans les *feature trackers* (et dans notre approche) [Lucas and Kanade, 1981, Noce et al., 2006, Bosch et al., 2007, Tran and Marchand, 2007];
- enfin de nouvelles approches proposent d'évaluer directement la position 3D [Malis, 2007].

Les méthodes exploitant directement l'image sont généralement plus rapides que les méthodes nécessitant un algorithme de traitement d'image. Cependant, avec les méthodes utilisant des descripteurs, la phase de recalage peut être très rapide [Bosch et al., 2007].

3.2.1.2 Modèles de mouvement

Dans ce travail, nous ne disposons pas de modèle précis des mouvements cardiaques. Bien que des travaux aient été initiés en ce sens³, nous n'avons pu obtenir un modèle utilisable à des fins de suivi.

Cependant, il existe plusieurs modèles génériques applicables pour prendre en compte la contrainte temporelle, et qui permettent un filtrage efficace du bruit de mesure : les modèles bayésiens. Ils se déclinent en trois grands modèles⁴ suivant la nature de la chaîne d'observation, et sont résumés sur la tableau 3.1.

La méthode la plus sûre est évidemment la filtre particulaire, mais l'algorithme *condensation*, pour CONDITIONAL DENSITY PROPAGATION [Blake and Isard, 1998], utilisé est

³avec la collaboration de Sébastien Auch-Roy, étudiant ingénieur en stage de fin d'études au Lirmm.

⁴il existe bien sûr de nombreuses variantes de ces estimateurs, les plus connues étant le *Filtre de Kalman Rapide* (ou *FKF* pour *Fast Kalman Filter*) Lange [2001] très efficace sur les grandes matrices et utilisé en traitement d'images; et le *filtre de Kalman sans biais* (*Unscented Kalman Filter*) qui est une version améliorée (et plus précise) de l'*EKF* Julier and Uhlmann [1997]. D'autres approches sont présentées dans Negenborn [2003].

	Filtre de Kalman [Kalman, 1960, Cuevas et al., 2005]	Filtre de Kalman étendu [Cuevas et al., 2005, Negenborn, 2003]	Filtre particulière [Blake and Isard, 1998]
modèle de processus	linéaire	non linéaire	non linéaire
modèle de bruit	gaussien	gaussien	quelconque (approché par un modèle “mélange de gaussiennes”)

TAB. 3.1 – *Principaux filtres bayésiens.*

beaucoup plus gourmand en puissance de calcul⁵. En pratique, un filtre de Kalman est suffisant dans beaucoup de cas.

3.2.1.3 Formalisation du suivi en termes d’optimisation

La phase d’estimation, fondamentale dans l’algorithme de suivi, se pose aisément en termes d’optimisation. De fait, une très large majorité des algorithmes de suivi se basent sur la minimisation d’une métrique de similarité entre le modèle connu de l’objet et l’image courante fournie par le capteur.

On considère une image \mathcal{I} . C’est une matrice de taille $n \times m$ (on a souvent $n = m$) associant à un ensemble de points \mathbf{p} les valeurs de luminance⁶ correspondantes $\mathcal{I}(\mathbf{p})$. On se donne également une fonction $d(\mathcal{I}, \mathcal{I}^*)$ qui mesure la similarité entre deux images. Idéalement, d se comporte comme une distance (elle est positive et s’annule lorsque les deux images sont identiques) et est, au moins localement, unimodale⁷ afin d’assurer la convergence des algorithmes d’optimisation. Réaliser le suivi, c’est trouver un nouvel ensemble de points \mathbf{p}^* dans l’image \mathcal{I}^* tels que

$$d(\mathcal{I}(\mathbf{p}), \mathcal{I}^*(\mathbf{p}^*)) = 0. \quad (3.3)$$

En pratique on cherchera à résoudre :

$$\mathbf{p}^* = \operatorname{argmin} d(\mathcal{I}(\mathbf{p}), \mathcal{I}^*(\mathbf{p}^*)) \quad (3.4)$$

⁵même si des progrès ont été fait dans l’amélioration de cette méthode.

⁶ou un vecteur dans le cas d’une image couleur, mais nous avons toujours travaillé avec des images noir et blanc.

⁷c’est une condition nécessaire en théorie, mais en pratique on ne la vérifie jamais car on ne dispose pas de suffisamment d’information a priori sur le signal.

ce qui n'implique pas que d s'annule. En introduisant l'état $\mathbf{e}(t)$ du système dans l'équation, on obtient :

$$\begin{aligned}\mathbf{p} &= f(\mathbf{e}(t)), \mathbf{p}^* = f(\mathbf{e}(t + \tau)), \\ d(\mathcal{I}(f(\mathbf{e}(t))), \mathcal{I}^*(f(\mathbf{e}(t + \tau)))) &= 0,\end{aligned}\tag{3.5}$$

d'où finalement

$$\mathbf{e}(t + \tau) = \operatorname{argmin} d(\mathcal{I}(f(\mathbf{e}(t))), \mathcal{I}^*(f(\mathbf{e}(t + \tau)))).\tag{3.6}$$

Cette formulation est la plus générique, mais n'est pas la plus utilisée en pratique. En effet, on cherche généralement la position de l'objet dans l'image avec l'équation (3.4), puis on en déduit la nouvelle valeur de l'état, plutôt que d'utiliser directement l'équation (3.6).

Cette modélisation fait ressortir les deux points critiques en suivi :

- le choix d'une métrique d bien adaptée à l'objet suivi et suffisamment robuste au bruit,
- le choix d'un algorithme d'optimisation suffisamment performant.

3.3 Principales méthodes de suivi : état de l'art

3.3.1 Techniques de suivi de points

Le suivi de points est la composante fondamentale du suivi de mouvement. C'est à la fois l'approche la plus simple (seulement 2 degrés de liberté) mais aussi celle qui rencontre le plus d'applications : pour le suivi d'objets inconnus, pour le suivi d'objets CAO à partir de points caractéristiques, pour la reconstruction 3D en stéréo-vision (chapitre 4) où l'on cherche à retrouver le même point sur plusieurs vues, etc. Bien que très étudié et utilisé, ce type de suivi laisse encore beaucoup de champ à des développements. De nouvelles méthodes [Lowe, 1999, Malis, 2004], sont régulièrement proposés. Dans ce paragraphe, nous introduisons les méthodes les plus courantes pour réaliser le suivi de points en mouvement, dont certaines nous ont directement inspiré dans le développement de l'algorithme présenté au §3.4.

3.3.1.1 Hypothèses fondamentales

Un point de l'image est caractérisé par une paire de coordonnées en pixels⁸. Le tracking de points consiste à trouver le déplacement (un vecteur de translation \mathbf{u}) entre un point d'origine \mathbf{p} et un point d'arrivée \mathbf{p}^* , soit

$$\mathbf{p}^* = \mathbf{p} + \mathbf{u}.\tag{3.7}$$

⁸en réalité on peut obtenir une précision inférieure au pixel, mais en considérant les coordonnées comme des réels le modèle reste valable.

que l'on résout par la minimisation introduite dans les équations (3.4) et (3.6). Afin de garantir l'existence de solutions et d'appliquer les méthodes présentées ci-dessous, nous ferons en plus deux hypothèses fondamentales :

- **l'image est consistante** , c'est-à-dire que l'objet que nous cherchons à repérer se trouve dans l'image et reste dans l'image tout au long du suivi⁹. En d'autres termes, l'équation

$$\mathbf{u} = \operatorname{argmin} d(\mathcal{I}(\mathbf{p}), \mathcal{I}^*(\mathbf{p} + \mathbf{u})) \quad (3.8)$$

admet toujours une solution t , unique, correspondant à l'objet recherché.

- **les déplacements entre deux images consécutives sont limités**, et il en est de même des différents phénomènes induits par ces déplacements : déformation, changement de pose, variation d'illumination. Cette hypothèse justifie l'utilisation d'approches par optimisation, dont la convergence n'est possible que dans un voisinage local du minimum recherché. Elle nous permet également de limiter la zone de recherche pour la nouvelle position du point, et ainsi d'améliorer le temps de calcul.

3.3.1.2 Suivi par corrélation/SSD

Cette méthode de suivi consiste à utiliser des mesures de similarité pour rechercher dans l'image un point qui minimise cette mesure.

C'est une méthode simple, mais qui a fait ses preuves dans des applications très diverses dont des application médicales [Ortmaier, 1999, Sauvée et al., 2005]. Son principal défaut est sa sensibilité aux variations du motif suivi tels que les déformations ou les changements dans l'éclairage. Nous nous sommes inspirés de cette méthode, en proposant d'utiliser une mesure moins sensible aux déformations (basée sur la texture) afin d'améliorer son comportement sur nos séquences in vivo.

Des métriques classiques pour ces méthodes sont les suivantes :

- **Mesure de corrélation** : Elle permet d'évaluer le degré de dépendance linéaire entre deux signaux et donc leur similarité. Numériquement, la corrélation est d'autant plus importante que les deux images sont proches et le suivi par corrélation va chercher à maximiser cette quantité. En pratique, on utilise la corrélation normalisée, donnée par l'équation¹⁰ :

$$NCC(\mathcal{I}(\mathbf{p}), \mathcal{I}^*(\mathbf{p}^*)) = \frac{\sum_{\mathbf{p}, \mathbf{p}^*} \mathcal{I}(\mathbf{p}) \mathcal{I}^*(\mathbf{p}^*)}{\sqrt{(\sum_{\mathbf{p}} \mathcal{I}(\mathbf{p}))^2 \cdot (\sum_{\mathbf{p}^*} \mathcal{I}^*(\mathbf{p}^*))^2}}. \quad (3.9)$$

⁹il existe des formes faibles de consistance, utilisées en stéréo-vision dans les cas d'occultation ou de transparence, mais nous ne les considérerons pas ici

¹⁰*NCC* pour *Normalized Cross Correlation*.

- **Mesure SSD :** La mesure *SSD* (pour *Sum of Squared Differences*) est une alternative à la mesure de corrélation. Elle exprime la distance quadratique entre les vecteurs constitués par les valeurs de luminance des pixels des régions d'intérêt considérées. On peut l'évaluer selon l'équation¹¹ :

$$NSSD(\mathcal{I}(\mathbf{p}), \mathcal{I}^*(\mathbf{p}^*)) = \frac{\sum_{\mathbf{p}, \mathbf{p}^*} (\mathcal{I}(\mathbf{p}) - \mathcal{I}^*(\mathbf{p}^*))^2}{\sqrt{(\sum_{\mathbf{p}} \mathcal{I}(\mathbf{p}))^2 (\sum_{\mathbf{p}^*} \mathcal{I}^*(\mathbf{p}^*))^2}}. \quad (3.10)$$

- **Mesure SAD :** La mesure *SAD* (pour *Sum of Absolute Differences*) est un autre exemple de norme classiquement utilisée en suivi. Elle présente l'avantage d'être moins sensible au bruit que la norme SSD, cependant elle est moins utilisée car elle se prête moins bien à l'optimisation :

$$SAD(\mathcal{I}(\mathbf{p}), \mathcal{I}^*(\mathbf{p}^*)) = \sum_{\mathbf{p}, \mathbf{p}^*} |\mathcal{I}(\mathbf{p}) - \mathcal{I}^*(\mathbf{p}^*)| \quad (3.11)$$

3.3.1.3 Le suivi Lucas-Kanade-Tomasi (par flot optique)

La méthode de Lucas, Kanade et Tomasi, ou méthode *KLT*, repose sur la notion de flot optique. Le flot optique est le déplacement effectué par chaque pixel (considéré indépendamment) entre deux instants. Il peut donc s'exprimer comme un vecteur de translation dépendant du pixel p considéré, du temps t et du temps t^* correspondant à l'image finale. Soit $z(p, t, t^*)$ le vecteur du flot optique au point p entre deux images. On a donc, par définition :

$$\mathcal{I}^*(\mathbf{p}^*) = \mathcal{I}(\mathbf{p} - z(\mathbf{p}, t, t^*)). \quad (3.12)$$

Le principe de la méthode *KLT*, et de toutes les méthodes basées sur la SSD et la mesure de corrélation, est alors de choisir une petite fenêtre autour du point considéré et de supposer, sous couvert de l'hypothèse de petits déplacements, que les déplacements de tous les pixels sont parallèles, et donc que le flot optique est constant. On a alors :

$$\mathcal{I}^*(\mathbf{p}^*) = \mathcal{I}(\mathbf{p} - z(t, t^*)) \quad (3.13)$$

soit, en considérant qu'on a toujours t et t^* correspondant à deux images successives fournies par la caméra :

$$\mathcal{I}^*(\mathbf{p}^*) = \mathcal{I}(\mathbf{p} - z). \quad (3.14)$$

¹¹il s'agit de la version normée de la SSD (d'où le *NSSD* pour *Normalized Sum of Squared Differences*), que nous utilisons pour des raisons évoquées plus loin dans ce document. De nombreux auteurs ne font pas de normalisation, ce qui permet de se ramener à un cas de minimisation quadratique.

Il faut donc résoudre¹²

$$z = \operatorname{argmin}_{q \in V(p)} \|I(q) - I(q - z)\|^2 \quad (3.15)$$

On résoud généralement ce système par la méthode de Gauss-Newton, en utilisant les gradients de l'image. Des développements récents de cet algorithme, principalement pour améliorer la rapidité de la méthode, sont présentés dans [Adelson et al., 1984] et [Sinha et al., 2006]. De plus, cette approche peut être modifiée pour le suivi de région, comme cela est présenté au paragraphe 3.3.2.

3.3.1.4 Le suivi avec la méthode SIFT

Développée par David Lowe et présentée initialement dans Lowe [1999], cette méthode propose d'utiliser des attributs invariants par rotation et par changement d'échelle. Les approches classiques par détecteur de coins (les détecteurs de Harris-Stephens ou KLT) échouent à suivre les points lors d'un zoom de la caméra. Par contre, les approches par coin sont généralement de bonnes approches pour la stéréo-vision car elle sont peu sensibles à une rotation de la vue.

La robustesse au changement d'échelle de SIFT vient de l'utilisation d'une représentation par *pyramide gaussienne* [Adelson et al., 1984] afin de chercher des descripteurs à plusieurs échelles simultanément. C'est une approche très intéressante du suivi, et des approches multi-échelles ont été développées pour d'autres méthodes, comme la méthode KLT [Adelson et al., 1984] avec succès.

Une autre application intéressante de l'approche SIFT est la reconnaissance d'objets. Il est possible de générer des "histogrammes d'orientation" pour définir une classe de descripteurs invariants par changement d'échelle. Ceux-ci sont par exemple utilisés dans le jouet Aibo [Munich et al., 2006] de Sony pour la reconnaissance d'objets.

3.3.2 Techniques de suivi de région

3.3.2.1 Hypothèses fondamentales

Pour les mêmes raisons que précédemment, on se place dans le cas d'une image consistante avec de petits déplacements. On introduit également une transformation w , le *warp*, paramétrée par un vecteur de paramètres \mathbf{G} , qui seront donc les inconnues dans le problème de suivi. Le vecteur \mathbf{G} optimal est celui qui envoie les pixels de l'objet suivi dans l'image \mathcal{I} initiale sur les pixels correspondant au même objet dans l'image suivante \mathcal{I}^* .

¹²avec $V(\mathbf{p})$ un voisinage de \mathbf{p} .

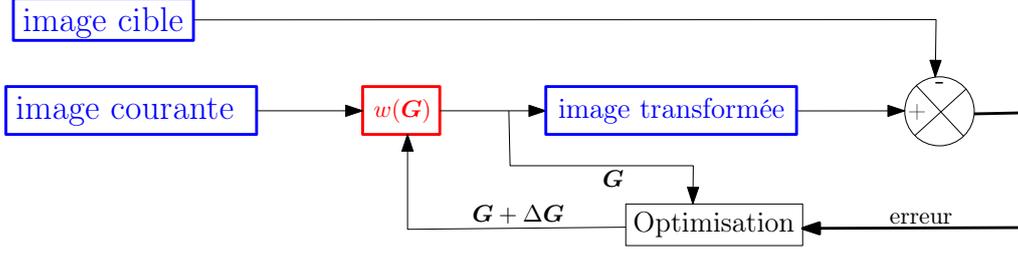


FIG. 3.2 – *Principe général des méthodes de suivi de région.* On itère en faisant varier les paramètres \mathbf{G} de la transformation jusqu'à ce que l'erreur entre l'image de référence et l'image transformée soit nulle.

Notre hypothèse de consistance s'écrit :

$$\mathcal{I}(\mathbf{p}) = \mathcal{I}^*(w(\mathbf{G})(\mathbf{p}^*)) = \mathcal{I}^*(\mathbf{p}^*) \quad (3.16)$$

et le problème de suivi s'écrit, au sens de la SSD :

$$\mathbf{G} = \operatorname{argmin} \sum_p (\mathcal{I}^*(w(\mathbf{G})(\mathbf{p})) - \mathcal{I}(\mathbf{p}))^2 \quad (3.17)$$

soit encore, en considérant les régions d'intérêt $\mathcal{I}(\mathbf{p})$ et $\mathcal{I}^*(\mathbf{p}^*)$ comme des vecteurs contenant les valeurs de niveaux de gris correspondantes :

$$\mathbf{G} = \operatorname{argmin} \|\mathcal{I}^*(w(\mathbf{G})(\mathbf{p})) - \mathcal{I}(\mathbf{p})\|_2. \quad (3.18)$$

Comme nous le verrons, toute la difficulté de ces méthodes est de trouver une façon efficace (en termes de temps de calcul) de résoudre ce problème d'optimisation. La figure 3.2 montre le schéma général des procédures d'optimisation itératives présentées dans ce paragraphe.

3.3.2.2 Méthode de Lucas Kanade

Version originale : L'approche de Lucas et Kanade consiste à développer au premier ordre le terme contenant le warp dans l'équation (3.18) puis d'utiliser une minimisation de type Gauss-Newton. La linéarisation s'écrit en exprimant l'évolution de la fonction de coût $F(\mathbf{G})$:

$$\begin{aligned} F(\mathbf{G}) &= F(\mathbf{G}_0 + d\mathbf{G}) = \|\mathcal{I}^*(w(\mathbf{G}_0 + d\mathbf{G})(\mathbf{p})) - \mathcal{I}(\mathbf{p})\|_2 \\ &= \left\| \mathcal{I}^*(w(\mathbf{G}_0)(\mathbf{p})) + \frac{\delta \mathcal{I}^*(w(\mathbf{G})(\mathbf{p}))}{\delta p} \Big|_{p_0} d\mathbf{G} - \mathcal{I}(\mathbf{p}) \right\|_2 \quad (3.19) \\ &= \|\mathcal{I}^*(w(\mathbf{G}_0)(\mathbf{p})) + J(\mathbf{G}_0).d\mathbf{G} - \mathcal{I}(\mathbf{p})\|_2 \end{aligned}$$

toute la subtilité étant alors d'exprimer la jacobienne $J(\mathbf{G}_0)$ de l'image "warpée" au point initial. Il faut pour cela décomposer le problème en utilisant le théorème de dérivation composée. En chaque point p on a

$$\left. \frac{\delta \mathcal{I}^*(w(\mathbf{G})(\mathbf{p}))}{\delta p} \right|_{p_0} = \frac{\delta \mathcal{I}^*}{\delta p}(w(\mathbf{G}_0)(\mathbf{p})) \cdot \left. \frac{\delta w}{\delta \mathbf{G}} \right|_{G_0} \quad (3.20)$$

le premier terme $\frac{\delta \mathcal{I}^*}{\delta p}(\dots)$ étant le gradient de l'image (c'est une matrice 1×2 contenant les gradients en x ou en y) exprimé aux points $w(\mathbf{G})(\mathbf{p})$, c'est-à-dire aux images des points d'origine par le warp. Le second terme $\left. \frac{\delta w}{\delta \mathbf{G}} \right|_{G_0}$ est la jacobienne du warp au point p . C'est une matrice $2 \times N_G$, avec N_G le nombre de paramètres du warp¹³. Finalement le produit (3.20) donne un vecteur $1 \times N_G$, qui est en fait la ligne de la jacobienne $J(\mathbf{G}_0)$ correspondant au point p . En effectuant ce calcul sur tous les points de l'image, on obtient finalement une jacobienne $N_p \times N_G$, où N_p est le nombre de points dans l'image. Enfin, on utilise un algorithme d'optimisation itératif, par exemple l'algorithme de Gauss-Newton, pour converger vers la valeur de \mathbf{G} recherchée

$$G \leftarrow \mathbf{G} + \Delta \mathbf{G} \quad (3.21)$$

$$\Delta \mathbf{G} = -(J(\mathbf{G}_0)^T J(\mathbf{G}_0))^{-1} J(\mathbf{G}_0)^T (\mathcal{I}^*(w(\mathbf{G})(\mathbf{p})) - \mathcal{I}(\mathbf{p})) \quad (3.22)$$

D'autres algorithmes sont possibles pour l'optimisation, comme la méthode du gradient ou la méthode de Newton. Cependant les méthodes de quasi-Newton (Gauss-Newton et Levenberg-Marquardt) offrent le meilleur compromis entre vitesse de convergence et complexité algorithmique, et sont donc les plus rapides en pratique.

Améliorations : Simon Baker [Baker and Matthews, 2002] propose une série d'améliorations à l'algorithme original de Kanade afin de diminuer les temps de calcul. Il s'agit essentiellement de trouver une nouvelle formulation qui permette de faire une partie des calculs hors ligne. L'auteur propose de composer le warp avec lui-même :

$$w(\mathbf{G})(\mathbf{p}) = w(\mathbf{G}_0) \circ w(\Delta \mathbf{G})(\mathbf{p}). \quad (3.23)$$

Pour pouvoir utiliser cette formulation, certaines conditions doivent être vérifiées :

- il faut que l'équation (3.23) admette toujours une solution. Heureusement, pour de nombreux warps courants (transformations affines, homographies) cette condition est respectée ;
- afin d'amorcer l'algorithme itératif d'optimisation, il faut que le warp de paramètre

¹³voir par exemple Baker and Matthews [2002] pour un exemple de calcul de cette jacobienne.

0 soit l'identité

$$w(O)(\mathbf{p}) = \mathbf{p}. \quad (3.24)$$

On peut alors réécrire le développement (3.20) qui devient

$$\left. \frac{\delta \mathcal{I}^*(w(\mathbf{G})(\mathbf{p}))}{\delta p} \right|_{p_0} = \frac{\delta \mathcal{I}^*(w(\mathbf{G}_0)(\mathbf{p}))}{\delta p} \left. \frac{\delta w}{\delta \mathbf{G}} \right|_0 \quad (3.25)$$

Le second terme du développement est exprimé en 0 et non plus en G_0 ¹⁴. Il est donc possible de calculer le jacobien une pour toute. C'est la méthode "forward compositional".

Une seconde amélioration a également été proposée : l'algorithme "inverse compositional". Il s'agit de composer non plus par le warp, mais par la fonction inverse :

$$w(\mathbf{G})(\mathbf{p}) = w(\mathbf{G}_0) \circ w^{-1}(\Delta \mathbf{G})(\mathbf{p}). \quad (3.26)$$

Pour comprendre l'intérêt de cette méthode, il faut réinjecter cette expression dans l'équation initiale (3.19)

$$F(\Delta \mathbf{G}) = \|\mathcal{I}^*(w(\mathbf{G}_0)(w^{-1}(\Delta \mathbf{G})(\mathbf{p}))) - \mathcal{I}(\mathbf{p})\|_2 \quad (3.27)$$

soit en posant $\mathbf{p} = w(\Delta \mathbf{G})(\mathbf{q})$:

$$F(\Delta \mathbf{G}) = \|\mathcal{I}^*(w(\mathbf{G}_0)(\mathbf{q})) - \mathcal{I}(w(\Delta \mathbf{G})(\mathbf{q}))\|_2 \quad (3.28)$$

On a donc déplacé la recherche dans l'image d'origine et non plus dans l'image courante. Après linéarisation, la fonction de coût devient, en reprenant la notation \mathbf{p} pour les pixels :

$$F(\Delta \mathbf{G}) = \|\mathcal{I}^*(w(\mathbf{G})(\mathbf{p})) - \mathcal{I}(\mathbf{p}) - \left. \frac{\delta \mathcal{I}(w(\mathbf{G})(\mathbf{p}))}{\delta p} \right|_0 d\mathbf{G}\|_2. \quad (3.29)$$

La jacobienne $\left. \frac{\delta \mathcal{I}(w(\mathbf{G})(\mathbf{p}))}{\delta p} \right|_0$ étant celle de l'image de référence, elle peut donc être calculée une seule fois pour tout le suivi.

3.3.2.3 Méthode ESM

La méthode ESM¹⁵ est très similaire à la méthode de Lucas Kanade. En fait, elle peut être vue comme une modification de l'algorithme "forward compositional" pour utiliser un algorithme d'optimisation plus performant. Au lieu d'utiliser l'approximation de

¹⁴on voit ici tout l'intérêt de la formulation par composition : la valeur initiale de $\Delta \mathbf{G}$ est toujours 0 (le warp identité) et non plus \mathbf{G}_0 comme dans l'approche additive classique : le jacobien associé est donc constant selon \mathbf{G} .

¹⁵que nous avons réimplémentée, les codes sources n'étant pas disponibles sur le site officiel.

Gauss-Newton “ $J^T J$ ” pour approximer la matrice hessienne, l’algorithme se base sur une approximation J_{esm} qui permet d’obtenir, sous certaines conditions, un taux de convergence équivalent à celui de la méthode de Newton, mais sans avoir à évaluer la matrice hessienne en chaque point. Si l’on reprend l’expression de la fonction de coût et que l’on en fait le développement limité à l’ordre 2 :

$$\begin{aligned} F(\Delta \mathbf{G}) &= \|\mathcal{I}^*(w(\mathbf{G}_0)(w(\Delta \mathbf{G})(\mathbf{p}))) - \mathcal{I}(\mathbf{p})\|_2 \\ &= F(0) + J(0).\Delta \mathbf{G} + \frac{1}{2}M(0, \Delta \mathbf{G}).\Delta \mathbf{G} + o(\|\Delta \mathbf{G}\|^3) \end{aligned} \quad (3.30)$$

en remarquant que :

$$J(\Delta \mathbf{G}) = J(0) + M(0, \Delta \mathbf{G}).\Delta \mathbf{G} + o(\|\Delta \mathbf{G}\|^3)$$

d’où

$$M(0, \Delta \mathbf{G}) = J(0) - J(\Delta \mathbf{G} + o(\|\Delta \mathbf{G}\|^3))$$

il vient

$$F(\Delta \mathbf{G}) = F(0) + \frac{1}{2}(J(0) + J(\Delta \mathbf{G})).\Delta \mathbf{G} + o(\|\Delta \mathbf{G}\|^3) \quad (3.31)$$

Tout l’art de cette méthode, décrit avec précision dans [Benhimane and Malis, 2007], est de choisir une paramétrisation permettant de calculer facilement $J(0) + J(\Delta \mathbf{G})$. Il est montré dans [Malis, 2004] que cette matrice peut s’écrire (grâce à une paramétrisation utilisant l’algèbre de Lie) comme une carte exponentielle, c’est-à-dire une exponentielle d’une somme de matrices élémentaires A_i paramétrée par un nouveau vecteur de paramètres \mathbf{x} :

$$\Delta \mathbf{G} = \Delta \mathbf{G}(\mathbf{x}) = \exp\left(\sum_{i=1}^8 x_i A_i\right). \quad (3.32)$$

On peut donc exprimer les termes de la jacobienne $J(O)$, comme dans l’équation (3.25) par :

$$\left. \frac{\delta \mathcal{I}^*(w(\Delta \mathbf{G}(\mathbf{x}))(\mathbf{p}))}{\delta p} \right|_{p_0} = \frac{\delta \mathcal{I}^*(w(\mathbf{G}_0)(\mathbf{p}))}{\delta p} \left. \frac{\delta w(\mathbf{Z})}{\delta \mathbf{Z}} \right|_{\mathbf{Z}=\Delta \mathbf{G}(0)=\mathbf{I}} \left. \frac{\delta \mathbf{G}(\mathbf{x})}{\delta \mathbf{x}} \right|_{x=0}. \quad (3.33)$$

On voit l’apparition du troisième terme découlant de la re-paramétrisation dans l’algèbre de Lie. Dans l’article original Malis [2004], ces trois matrices sont notées $J_{\mathcal{I}}$, J_w et J_G . La matrice J_w dépend du warp choisi (mais le warp est toujours le même, donc J_w est calculée une seule fois), J_G dépend de la paramétrisation (et est aussi constante) et $J_{\mathcal{I}}$ est la matrice du gradient de l’image de référence vue à travers le warp : c’est une matrice variable pendant la phase d’optimisation. De même, on peut décomposer la jacobienne $J(\Delta \mathbf{G}(x))$ en $J_{\mathcal{I}^*}$, J_w et J_{G_0} . La clé de voûte de la méthode et de la paramétrisation

(3.32) est la relation :

$$J_{G_0} \cdot \Delta \mathbf{G} = J_G \cdot \Delta \mathbf{G} \quad (3.34)$$

qui permet d'écrire :

$$(J(0) + J(\Delta \mathbf{G})) \Delta \mathbf{G} = ((J_{\mathcal{I}} + J_{\mathcal{I}^*}) J_w J_G) \Delta \mathbf{G}. \quad (3.35)$$

On note alors :

$$J_{esm} = (J_{\mathcal{I}} + J_{\mathcal{I}^*}) J_w J_G \quad (3.36)$$

et il suffit alors d'utiliser cette "jacobienne ESM" à la place de la jacobienne initiale dans l'optimisation de Gauss-Newton, soit :

1. on calcule J_{esm} ;
2. on détermine la nouvelle valeur des paramètres x à partir du vecteur d'erreur $\mathcal{I}^*(w(\mathbf{G})(\mathbf{p})) - \mathcal{I}(\mathbf{p})$:

$$x \leftarrow J_{esm}^+ (\mathcal{I}^*(w(\mathbf{G})(\mathbf{p})) - \mathcal{I}(\mathbf{p})) \quad (3.37)$$

3. on en déduit le warp correspondant avec la formule (3.32) en on recommence jusqu'à ce que la fonction de coût soit suffisamment proche de 0.

3.3.3 Autres techniques de suivi

Il existe d'autres techniques de suivi d'objets en mouvement. Notons l'exemple des contours actifs ou *snakes* qui permettent le suivi d'objets déformables. Ces méthodes permettent d'analyser la forme de l'objet suivi, fournissant ainsi une représentation d'un niveau supérieur de l'information contenue dans l'image. Elles ont été notamment utilisées dans des applications de suivi de piétons, d'analyse d'images échographiques, d'analyse d'images IRM, etc. Elles avaient d'ailleurs été envisagées dans le cadre du projet GABIE [Sauvée et al., 2005], mais ne se montraient pas suffisamment robustes en présence de bruit.

Notons également les bons résultats obtenus à Rennes avec une approche basée sur le suivi de contours des objets [Comport et al., 2003] qui se montre très rapide et ne nécessite aucun pré-traitement de l'image. Cette méthode a également été intégrée dans une approche mixte contour-texture [Pressigout and Marchand, 2006].

Les méthodes très récentes de suivi de motifs déformables [Gay-Bellile et al., 2007] sont de bons candidats pour notre application. Leur principale limitation est le temps de calcul, encore trop élevé pour le type d'application que nous visons (jusqu'à 10 secondes de calcul par image selon les auteurs) .

Enfin, l'utilisation récente du classifieur *ferns* a montré des performances intéressantes [Bosch et al., 2007]. Il s'agit d'une approche totalement différente de celles que nous avons

évoquées précédemment, plus proche de certaines méthodes de stéréo-vision que des méthodes classiques de suivi¹⁶. Le suivi se fait en deux phases : tout d’abord un traitement bas niveau afin d’extraire des points caractéristiques, puis une phase de recalage afin d’apparier certains de ces points avec des points correspondants sur le modèle¹⁷. La phase de détection de points est évidemment plus rapide que pour les méthodes de suivi, mais la phase de recalage est, du moins avec les approches antérieures, très gourmande en temps de calcul. Le classifieur *ferns*, très rapide, permet de lever cette limitation, et donc de faire du suivi à fréquence vidéo. Cette approche permet de considérer un nombre élevé de points, ce qui rend le suivi plus robuste (une partie des points peut être masquée sans biaiser l’estimation de la pose). Cependant les approches proposées par les auteurs de *ferns* sont pour l’instant focalisées sur des objets rigides, et il faudrait évaluer ce type de classifieur sur des surfaces en déformation avant d’appliquer cette méthode à notre problématique.

3.4 Suivi de mouvement cardiaque en utilisant la texture

Lorsque nous nous sommes intéressés à la problématique de compensation des mouvements cardiaques, nous nous sommes rendus compte qu’il était très difficile de réaliser un suivi efficace des mouvements de l’organe en utilisant uniquement la mise en correspondance de motifs. Cela est dû à plusieurs facteurs, inhérents à la nature de l’organe et de ses mouvements :

- le cœur a une surface relativement homogène, ce qui affecte certains algorithmes qui ont besoin de beaucoup de contraste au niveau des régions suivies ;
- les mouvements du cœur sont rapides, donc il faut augmenter la fréquence d’acquisition pour que l’hypothèse des petits déplacements soit valide ;
- la surface du cœur est extrêmement déformable et réflexive, deux facteurs qui mettent en défaut les algorithmes de suivi classiques.

Afin de résoudre ce problème, nous avons mené une série d’expériences et proposé des solutions matérielles et logicielles.

¹⁶citons également les travaux du LASMEA [E. et al., 2007a,b] sur la problématique du SLAM (Simultaneous Location and Mapping) avec des méthodes similaires. Nous reparlerons de cette distinction au §4.4.2.1.

¹⁷c’est l’approche classique de la stéréo-vision, lorsqu’on ne cherche pas à suivre un mouvement. On sélectionne alors un jeu aléatoire de points (selon une heuristique désignant de “bons” points à suivre) et on reconstruit la scène.

3.4.1 Acquisition d'images et aspects matériels

Afin de pouvoir travailler sur le plus de séquences différentes possibles, nous avons pris plusieurs séries d'images expérimentales sur des cœurs porcins¹⁸. Ces séries d'images ont été acquises à haute fréquence (entre 100 et 300Hz) et à des résolutions allant de 128×128 jusqu'à 1024×1024 . La vitesse d'acquisition est essentiellement limitée par trois facteurs :

- la résolution de l'image ;
- la vitesse de transfert du capteur vers la mémoire de l'ordinateur ;
- l'illumination de la scène (le temps d'exposition étant décroissant).

3.4.1.1 Aspects capteur

Lors des prises de séquences d'images expérimentales, nous avons développé un logiciel permettant l'acquisition d'images brutes vers le disque dur¹⁹. Le principal facteur limitant de ce logiciel est alors le taux de transfert du disque dur (de l'ordre de 20Mo/sec sur notre plateforme), mais en contrepartie il n'y a pas de limite forte sur la longueur des séquences, qui peuvent atteindre plusieurs minutes.

Nous avons également réalisé l'acquisition vers la mémoire RAM de l'ordinateur, ce qui autorise des fréquences d'acquisition plus élevées, mais limite de nombre d'images de la séquence (à 2000 pour des images de résolution 256×256). L'acquisition RAM est donc à réserver pour le traitement en ligne des images.

La figure 3.3a montre le type de caméra que nous avons utilisé : une caméra *DALSA Im75*, interfacée avec une carte d'acquisition *Matrox Meteor II* par une liaison *CamLink*. Le programme a été réalisé en utilisant les bibliothèques *Matrox MIL* et *OpenCV*²⁰, le tout piloté par une série d'utilitaires écrits en langage Perl²¹. La figure 3.3b montre l'évolution de la vitesse d'acquisition maximale par rapport à la résolution. On remarque deux phases sur ces courbes : pour des résolutions très faibles (32×32 ou 64×64), les vitesses d'acquisition sont élevées et ne sont limitées que par la durée de la boucle acquisition-stockage. Quand la résolution augmente, on remarque une évolution quasiment linéaire qui marque les limites liées au transfert des images de la caméra vers la mémoire (disque ou RAM).

3.4.1.2 Éclairage et réflexions

Le problème de l'éclairage est prépondérant en vision rapide, et encore plus lorsqu'on considère des applications médicales. Les images étant trop sombres avec un éclairage

¹⁸Nous avons réalisé 3 expérimentations in vivo et nous sommes greffés sur une expérimentation réalisée au *LRP* dans le cadre des travaux Vitrani [2006]. Nous remercions au passage le *LRP* pour cette opportunité et pour son accueil.

¹⁹un phase de décompression étant alors nécessaire afin d'exploiter les images.

²⁰bibliothèque utilisée pour certains traitements et pour la sauvegarde dans différents formats.

²¹pour l'interface et pour générer automatiquement des rapports sur les différentes acquisitions.

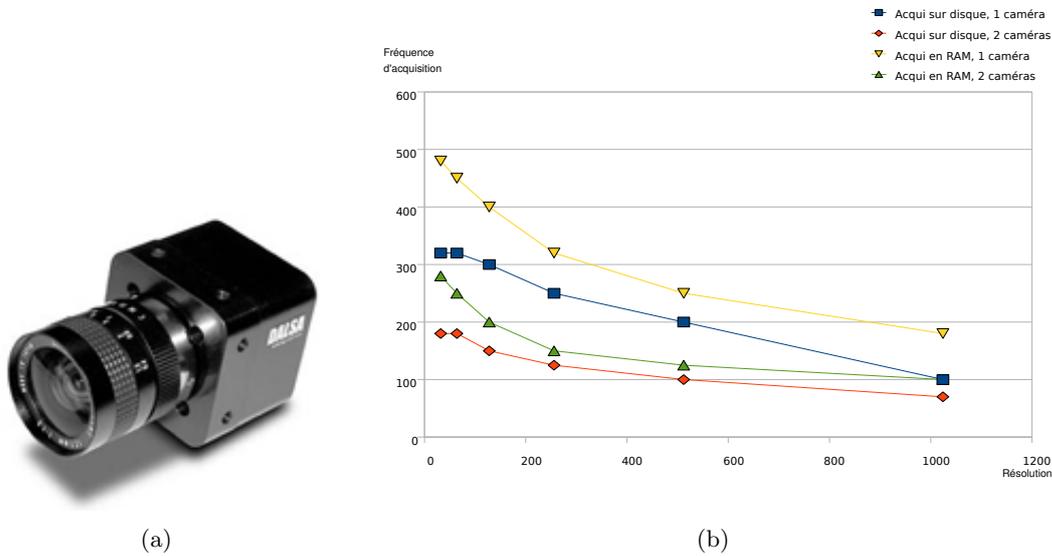


FIG. 3.3 – *Capteur utilisé pour l’acquisition d’images.* (a) caméra CMOS DALSA 1m75, (b) courbes caractérisant la relation entre la taille de l’image et la vitesse maximale d’acquisition (au-delà nous rencontrons un problème de “perte” d’images).

classique, nous avons dans un premier temps choisi un éclairage halogène. Celui-ci, très puissant, permet d’obtenir des images claires, mais il pose d’autres problèmes :

- il dégage énormément de chaleur, ce qui risque de dessécher voire d’endommager les tissus ;
- de nombreuses réflexions spéculaires apparaissent dans les images, rendant difficile le suivi.

Pour pallier ces problèmes, nous nous sommes alors tournés vers des éclairages de type LED. Ceux-ci dégagent peu de chaleur et la lumière monochromatique²² induit moins de réflexions dans l’image.

De plus, nous avons effectué un pré-traitement des images en ajustant la courbe LUT (pour *look up table* : cette fonction ajuste la réponse du capteur à certaines valeurs d’intensité lumineuse) afin d’obtenir une bonne dynamique. Cette opération étant matérielle (elle est intégrée à la caméra), elle ne pèse pas sur la boucle des traitements tout en augmentant sensiblement la qualité des séquences d’images.

²²après plusieurs essais, nous avons choisi d’utiliser une lumière rouge qui induit une meilleure réponse du capteur. Évidemment, cela implique de travailler avec des images en niveaux de gris.

3.4.2 Suivi par analyse de texture

Au chapitre 2, nous avons présenté la procédure de sélection qui a permis de déterminer les 8 attributs texturaux les plus performants pour caractériser la surface du cœur. Afin de les utiliser pour réaliser un suivi, il faut créer un descripteur à partir de ces attributs. Dans notre cas, nous avons considéré le *vecteur de texture* constitué des 8 attributs. On définit donc une fonction \mathcal{T} qui, à chaque région d'intérêt $\mathcal{I}(\mathbf{p})$, associe le vecteur de texture correspondant :

$$\mathcal{T} : \mathcal{I} \rightarrow \begin{pmatrix} \rho_1 \\ \vdots \\ \rho_8 \end{pmatrix}. \quad (3.38)$$

À chaque région d'intérêt, nous faisons correspondre un vecteur de texture. La minimisation correspondant à l'équation (3.4) se fait sur une distance entre ces vecteurs caractéristiques.

3.4.2.1 Distance inter-textures

Notre idée initiale était de réaliser un suivi basé uniquement sur l'information de texture. Pour cela, nous l'avons vu, nous associons à chaque région d'intérêt un vecteur caractéristique. L'étape suivante est de choisir une distance pour décider si deux régions sont "proches" ou "différentes". Nous avons envisagé plusieurs solutions classiques telles que :

- La distance de Manhattan

$$d_T(\mathcal{I}, \mathcal{I}^*) = \frac{1}{8} \sum_{i=1}^8 |\rho_i - \rho_i^*| \quad (3.39)$$

- La distance euclidienne

$$d_T(\mathcal{I}, \mathcal{I}^*) = \frac{1}{8} \sum_{i=1}^8 (\rho_i - \rho_i^*)^2 \quad (3.40)$$

- La distance d_∞

$$d_T(\mathcal{I}, \mathcal{I}^*) = \max_{i \in [1,8]} |\rho_i - \rho_i^*| \quad (3.41)$$

- La distance de Mahalanobis

$$d_T(\mathcal{I}, \mathcal{I}^*) = (\mathcal{T}(\mathcal{I}^*(\mathbf{p}^*)) - \mathcal{T}(\mathcal{I}(\mathbf{p})))^t C^{-1} (\mathcal{T}(\mathcal{I}^*(\mathbf{p}^*)) - \mathcal{T}(\mathcal{I}(\mathbf{p}))) \quad (3.42)$$

où C est la matrice de covariance des données.

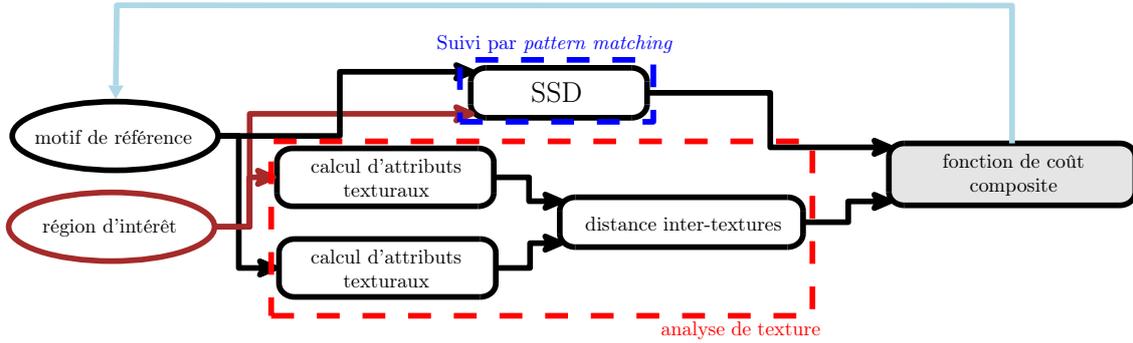


FIG. 3.4 – *Principe du suivi avec la fonction de coût composite.*

Nous avons donc testé ces différentes distances pour effectuer un suivi, mais le résultat obtenu n'était pas très "précis" (c'est un critère visuel, mais sensible, comme l'illustre la figure). Nous avons donc décidé d'utiliser la distance euclidienne et de fusionner l'approche par texture avec un suivi par SSD : c'est la *fonction de coût composite*.

3.4.2.2 Fonction de coût Composite

La fonction de coût composite est une distance mixte basée sur la SSD et sur une distance entre les vecteurs de texture des régions d'intérêt. Le principe de cette méthode, illustré sur la figure 3.4, est de combiner les avantages de deux approches :

- la **SSD** qui n'est pas robuste aux déformations de motifs mais est très précise : elle ne s'annule que lorsque les images considérées sont exactement identiques.
- la **distance inter-textures** qui est beaucoup plus robuste aux déformations mais n'est pas très précise.

En reprenant les notations introduites précédemment, on a donc d'un côté la distance inter-textures

$$\begin{aligned} d_T(\mathcal{I}(\mathbf{p}), \mathcal{I}^*(\mathbf{p}^*)) &= \frac{1}{8} \|\mathcal{T}(\mathcal{I}(\mathbf{p})) - \mathcal{T}(\mathcal{I}^*(\mathbf{p}^*))\|^2 \\ &= \frac{1}{8} \sum_{i=1}^8 (\rho_i - \rho_i^*)^2 \end{aligned}$$

et d'un autre côté la distance SSD, introduite en (3.10)

$$d_{SSD}(\mathcal{I}(\mathbf{p}), \mathcal{I}^*(\mathbf{p}^*)) = \frac{\sum_{\mathbf{p}, \mathbf{p}^*} (\mathcal{I}(\mathbf{p}) - \mathcal{I}^*(\mathbf{p}^*))^2}{\sqrt{(\sum_{\mathbf{p}} \mathcal{I}(\mathbf{p}))^2 \cdot (\sum_{\mathbf{p}^*} \mathcal{I}^*(\mathbf{p}^*))^2}}. \quad (3.43)$$

On peut alors exprimer la fonction de coût composite \mathcal{D} , comme une somme pondérée des deux distances :

$$\mathcal{D}(\mathcal{I}(\mathbf{p}), \mathcal{I}^*(\mathbf{p}^*)) = \lambda \cdot d_{SSD}(\mathcal{I}(\mathbf{p}), \mathcal{I}^*(\mathbf{p}^*)) + (1 - \lambda) \cdot d_T(\mathcal{I}, \mathcal{I}^*) \quad (3.44)$$

La valeur λ représente l'influence de l'approche par comparaison de motifs et $1 - \lambda$ celle de l'analyse de texture.

3.4.3 Améliorations

Par rapport aux schémas de base (suivi par distance texture ou composite) du §3.4.2, nous avons mis au point une série d'améliorations pour rendre le suivi plus robuste.

3.4.3.1 Ajustement dynamique de λ

Ce paramètre peut être fixé de façon statique en fonction de l'application visée, ou être variable. Une heuristique intéressante est d'évaluer en premier la similitude des motifs et de favoriser cette dernière si les images sont très proches, ou de favoriser l'analyse de texture sinon. Ainsi à chaque image on peut, en fonction de l'évolution de la déformation du motif, affiner λ pour le suivi à l'image suivante (l'importance relative de la SSD est d'autant plus importante que la déformation est faible, et inversement).

On définit alors λ à partir d'une fonction de pondération σ :

$$\begin{aligned} \sigma : [0, 1] &\longrightarrow [0, 1] \\ d_{min}(\mathcal{I}(\mathbf{p}), \mathcal{I}^*(\mathbf{p}^*)) &\longmapsto \lambda \end{aligned} \quad (3.45)$$

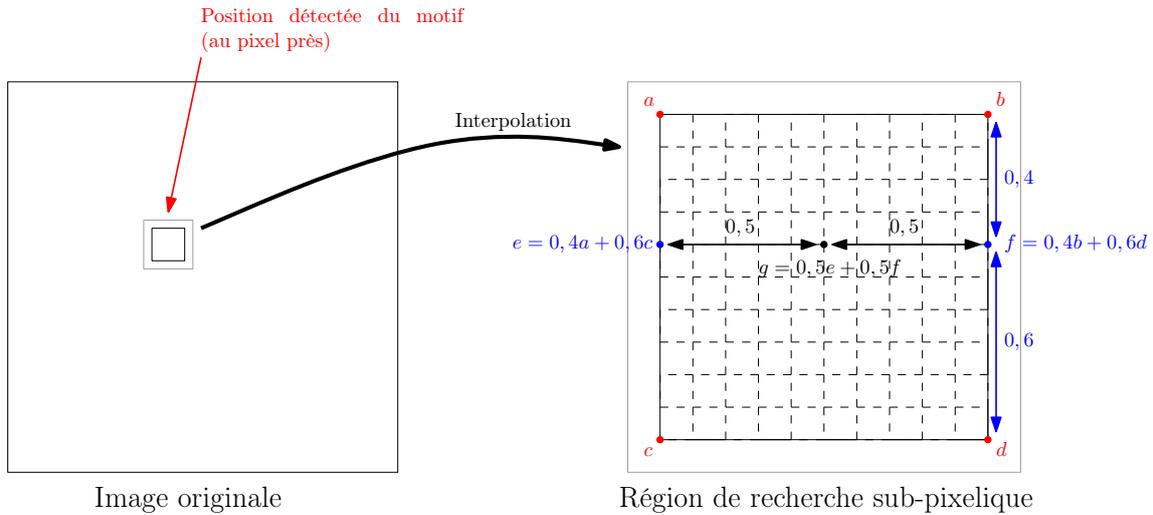
avec d_{min} la valeur minimale de la distance SSD évaluée à l'image précédente, $\sigma(0) = 1$ (la similarité des motifs est parfaite, on ne prends pas en compte l'analyse de texture) et $\sigma(1) = 0$.

Du point de vue de l'implémentation, nous définissons des valeurs seuil ϵ_{min} et ϵ_{max} en dessous et au dessus desquelles le suivi est totalement basé sur la SSD ou l'analyse de texture. Cela permet d'optimiser certaines phases du suivi (lorsque le motif est très peu déformé ou au contraire totalement "étiré").

Pendant nos tests, nous avons testé plusieurs fonctions de pondération parmi lesquelles les fonctions :

$$\sigma : x \longmapsto x - 1 \quad (3.46)$$

$$\sigma : x \longmapsto \exp \frac{-x}{1 - x} \quad (3.47)$$

FIG. 3.5 – *Sur-résolution par interpolation bi-linéaire.*

se sont montrées très performantes.

3.4.3.2 Sur-résolution

Une deuxième amélioration de la méthode consiste à rechercher une précision inférieure au pixel. Pour cela, nous utilisons un procédé en deux étapes :

- nous recherchons d’abord la position du motif, avec une précision de l’ordre du pixel ;
- puis nous réalisons une interpolation bi-linéaire (interpolation d’ordre 1) afin d’obtenir un grossissement de l’image autour de cette position, comme illustré sur la figure 3.5.

Cette méthode permet d’espérer une précision de l’ordre du dixième de pixel. Pour la rendre plus rapide, nous n’utilisons que la distance SSD pour cette phase d’ajustement.

3.4.3.3 Détection de perte d’un point

Dans certaines scènes, il est possible que le suivi échoue et “perde” une des zones d’intérêt. Pour détecter cela, nous associons aux points suivis un booléen qui désigne la validité de l’indice visuel²³. Nous considérons qu’un point devient invalide si :

- il entre dans une zone située à 5 pixels du bord de l’image (il risque de sortir de l’image) ;
- la distance composite \mathcal{D} dépasse une valeur seuil.

²³cette approche est empruntée au suivi KLT.

3.4.3.4 Filtrage

Enfin, nous avons testé le filtrage de Kalman²⁴ avec un modèle à vitesse constante, comme évoqué au §3.2.1. Cette approche permet de lisser la courbe du mouvement. Cependant, si ce type de suivi est très robuste, il n'est pas très adapté aux mouvements brusques du cœur, où il a tendance à ajouter une "erreur de traînage".

3.5 Évaluation

Afin d'évaluer les performances de notre méthode, nous avons effectué une série de tests, tout d'abord sur des images artificielles, puis sur des images naturelles. L'intérêt de commencer l'évaluation par des images générées par ordinateur est que l'on possède une vérité terrain, ce qui n'est pas le cas pour les séquences naturelles.

3.5.1 Séquences artificielles

Afin d'évaluer la capacité de notre algorithme à suivre des scènes texturées nous avons généré une série de tests selon le principe suivant : un patch texturé est déplacé suivant une trajectoire prédéfinie sur un fond également texturé. On connaît donc avec précision la trajectoire de référence, que l'on peut comparer avec la trajectoire calculée par l'algorithme de suivi.

3.5.1.1 Séquences de test

Nous avons généré plusieurs jeux de textures artificielles. Le principe utilisé pour générer ces textures est illustré sur la figure 3.6 : nous superposons à un fond (fond uni, séquence de cœur battant, etc.) un petit patch texturé que nous cherchons à suivre. L'évaluation de la précision du suivi se fait alors en mesurant la distance entre la position du motif (le centre du motif dans le cas du suivi ESM ou composite) et la trajectoire de référence, en mesurant l'écart moyen ou l'écart maximum :

$$d_{max} = \|\mathbf{P}_{ref} - \mathbf{P}_{suivi}\|_{\infty} = \max \left(\sqrt{(x_i - x'_i)^2 + (y_i - y'_i)^2} \right)$$

$$d_{mean} = \|\mathbf{P}_{ref} - \mathbf{P}_{suivi}\|_2 = \sqrt{\sum_{i=1..n} ((x_i - x'_i)^2 + (y_i - y'_i)^2)}$$

avec $\mathbf{P}_{ref} = \begin{pmatrix} x_1 & x_2 & \dots & x_n \\ y_1 & y_2 & \dots & y_n \end{pmatrix}$ la trajectoire de référence,

²⁴implémentation pour le suivi est très bien décrite dans [Cuevas et al., 2005], et des approches plus avancées sont développées dans [Negenborn, 2003].

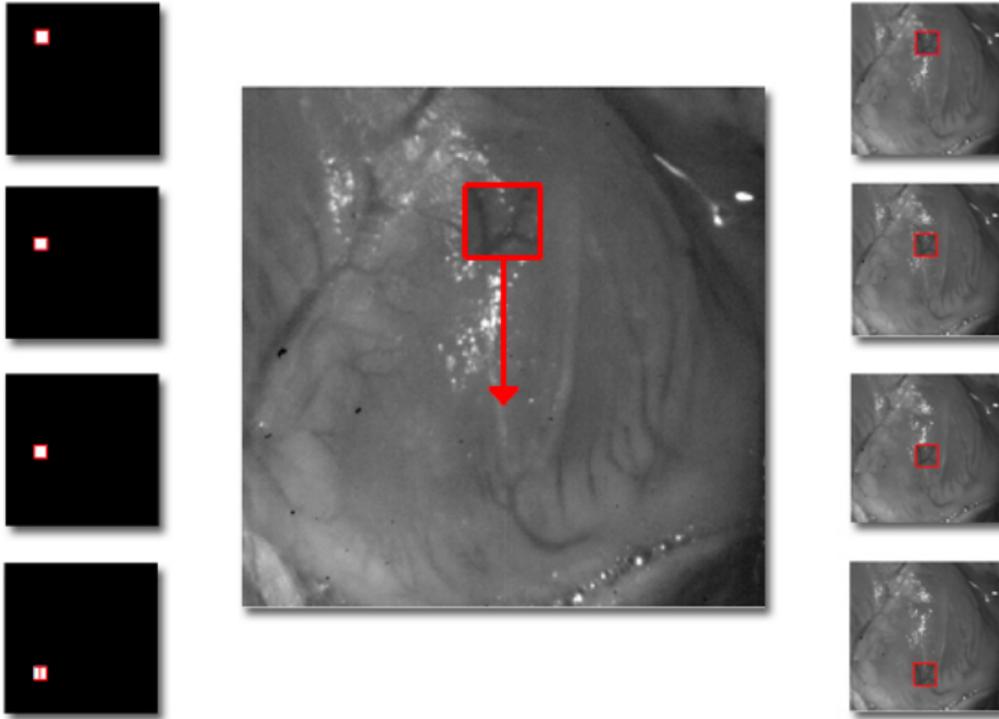


FIG. 3.6 – *Principe de création des séquences artificielles.*

et $\mathbf{p}_{\text{sui}} = \begin{pmatrix} x'_1 & x'_2 & \dots & x'_n \\ y'_1 & y'_2 & \dots & y'_n \end{pmatrix}$ la trajectoire évaluée par l'algorithme de suivi.

La figure 3.7 illustre certaines des séquences générées. On trouve des séquences très simples, telles celle de la figure 3.7a où l'on cherche à donner un mouvement de translation. Ces séquences ne posent aucun problème aux différentes méthodes. D'autres séquences sont générées en superposant des séquences de cœur (figures 3.7b), parfois en appliquant des transformations au patch suivi : rotation ou transformation homographique (figure 3.7c). Ces dernières mettent rapidement en défaut les approches classiques par flot optique ou SSD²⁵, alors que la méthode ESM, qui cherche à estimer cette transformation, est d'une excellente robustesse. Notre méthode basée sur la texture est plus robuste que les approches classiques, mais reste moins efficace sur ce type de déformations que l'approche ESM. Cependant, il faut noter que les déformations observées à la surface du cœur ne sont pas de nature homographique. Le phénomène le plus gênant en terme de suivi n'est pas la rotation du motif mais son "étirement" sous l'effet de la dilatation de l'organe. Comme nous le verrons au §3.5.2, le comportement de notre approche est bon dans ce cas.

²⁵bien que ce modèle soit limité au vu des déformations observables sur des séquences réelles de cœur, il est aisément paramétrable pour les tests.

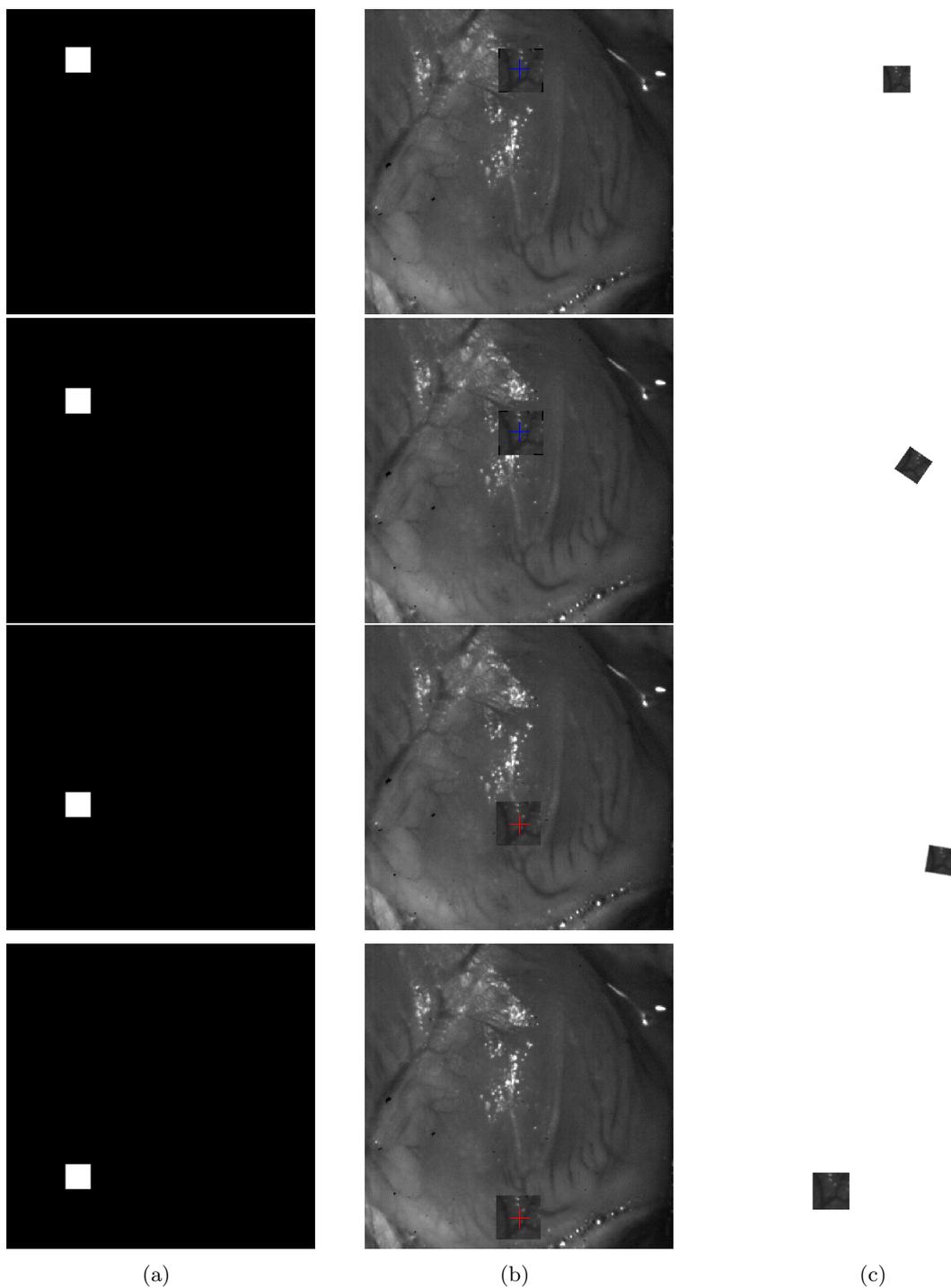


FIG. 3.7 – *Exemples de séquences artificielles.* (a) séquence simple, translation, (b) séquence texture sur texture, (c) séquence avec déformation du motif. La zone de recherche est définie centrée sur la position précédente avec une taille de trois fois les dimensions du motif recherché.

Méthode	scène en translation (figures 3.7b et 3.8)		scène avec mouvements homographiques (figure 3.7c)	
	d_{max}	d_{mean}	d_{max}	d_{mean}
SSD	1	0,2	2.3	1.2
Flot Optique (KLT)	4,1	0,9	X	X
Composite	0	0	1.6	0.7
ESM	0	0	0.3	0.1

TAB. 3.2 – *Précision du suivi* pour l'exemple présenté figure 3.8 et pour le mouvement par homographie de la figure 3.7c. Les cases marquées “X” correspondent aux cas où le suivi n'est plus possible car l'algorithme perd complètement la cible.

3.5.1.2 Évaluation

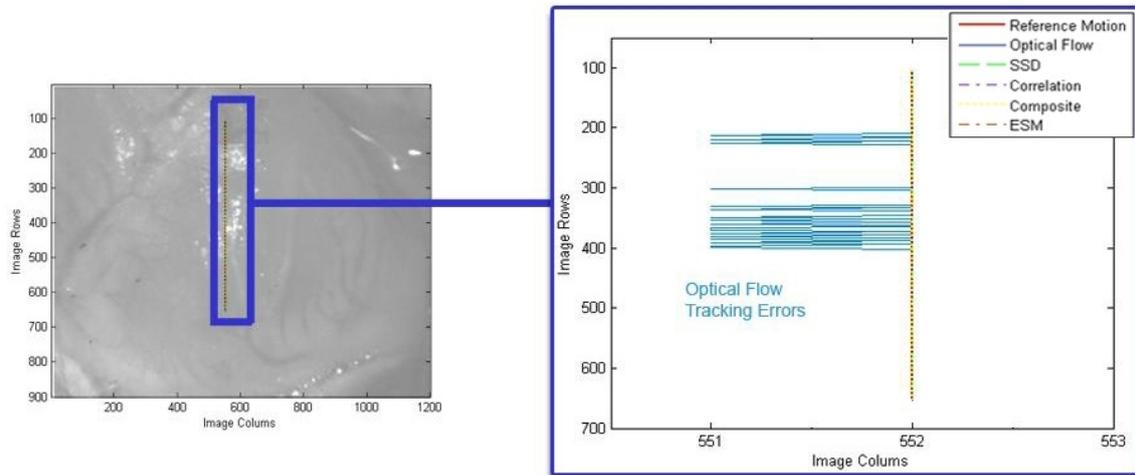
Lors des séquences de test générées avec un motif en translation pure, les différentes méthodes ont un bon comportement. La figure 3.8 illustre le type de résultats obtenus pour les trajectoires générées par l'approche décrite figure 3.6. On remarque les résultats sensiblement moins bons obtenus par le suivi KLT (flot optique) qui a tendance à “osciller” dès qu'apparaissent de petites réflexions²⁶. C'est un problème typique de cette méthode, qui est très sensible aux variations locales du gradient (d'autant plus que nos images ne présentent pas de forts contrastes). Les autres méthodes se comportent très bien, avec une erreur quasiment nulle, comme illustré sur le tableau²⁷.

Ce tableau illustre aussi les résultats obtenus avec la séquence présentée à la figure 3.7c, sur laquelle nous avons déplacé notre motif par des transformations homographiques successives sur un fond blanc. Nous utilisons une interpolation linéaire pour projeter le motif sur l'image, ce qui nous permet d'évaluer la position du centre au niveau sub-pixellique, et de tester ainsi l'approche par sur-résolution. On remarque que la méthode par flot optique n'arrive pas à assurer le suivi durant toute la séquence. À l'opposée, la méthode ESM se montre d'une efficacité remarquable sur cet exemple. Les raisons sont multiples : la méthode est conçue pour suivre ce type de transformation, et le fort contraste du motif sombre sur un fond blanc favorise encore l'approche. Les autres méthodes, la méthode SSD et notre méthode, arrivent plus difficilement à suivre les déplacements du motif, même si on remarque que l'approche par texture amène un plus.

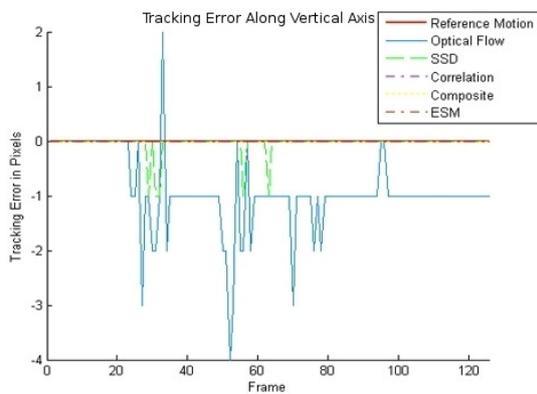
La figure 3.9 montre un cas problématique, où nous projetons notre cible sur une

²⁶ces réflexions apparaissent généralement sur la partie du cœur la plus proche de la caméra (et de la source lumineuse). Il s'agit souvent de la zone que nous cherchons justement à suivre.

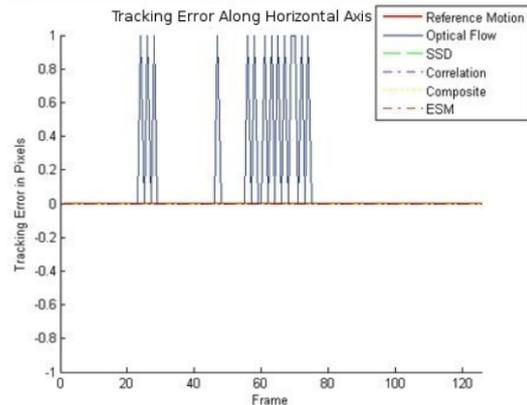
²⁷nous utilisons ici la SSD **non** normalisée, afin de comparer avec la corrélation (car les versions normalisées des deux méthodes sont équivalentes).



(a)



(b)



(c)

FIG. 3.8 – *Résultats sur un exemple de translation.* (a) exemple de trajectoire, avec illustration de l'erreur du suivi KLT (flot optique), (b) projection de l'erreur suivant l'axe vertical, (c) projection de l'erreur suivant l'axe horizontal. Cette série montre surtout la faiblesse de l'approche par flot optique en présence de réflexions spéculaires et devant la forte homogénéité de l'image. Elle illustre également l'intérêt des séries d'images générées artificiellement, qui permettent d'analyser et de quantifier facilement l'erreur induite par les différentes méthodes.

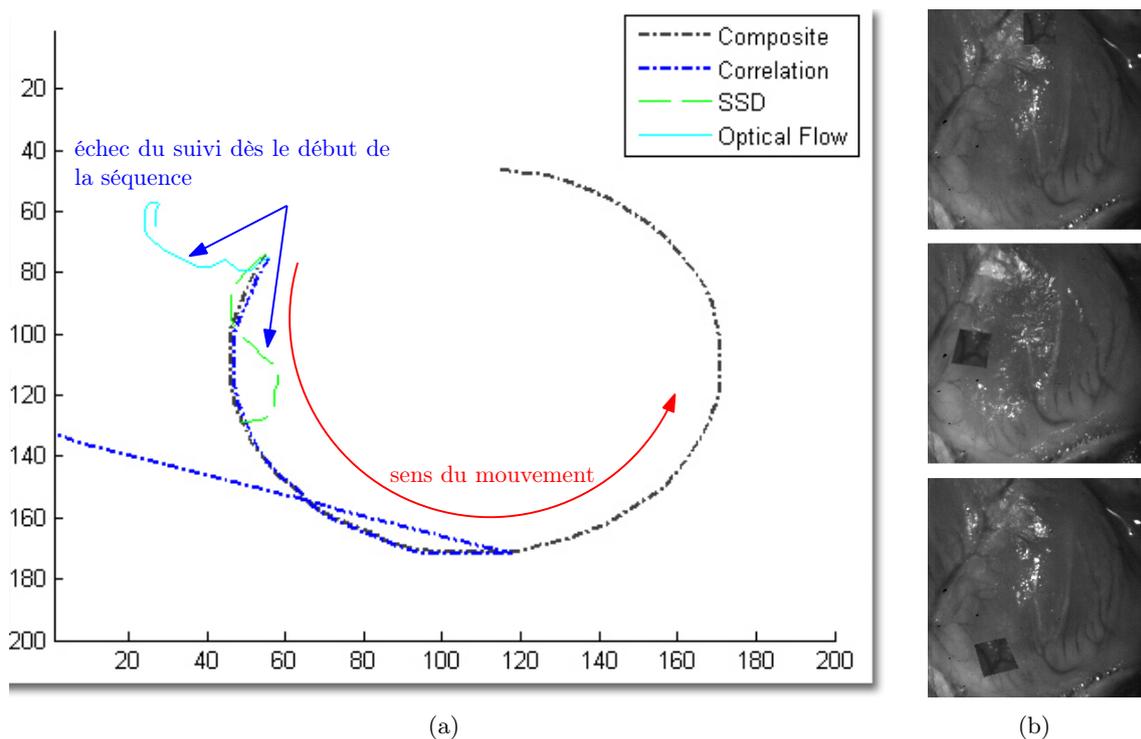


FIG. 3.9 – *Un autre exemple de suivi.* (a) trajectoires correspondant au mouvement, (b) extrait de la séquence correspondante. Dans cet exemple, on ajoute au mouvement du motif des déformations homographiques. Ce type de déformation fonctionnant parfaitement avec ESM (qui est conçu dans ce sens), nous n'avons représenté ici que les méthodes de suivi de point. Sur cet exemple particulier, on remarque facilement l'intérêt de l'approche par texture, qui améliore la robustesse du suivi, là où d'autres méthodes perdent la cible.

texture de cœur. Comme nous le constatons sur la figure 3.9, seules les approches ESM et Composite arrive à suivre le motif²⁸. Les autres critères ont tendance à “préférer” au motif suivi une autre partie de l’image et perdent rapidement la cible.

Les résultats sur les différentes séquence artificielles montrent bien l’avantage de l’approche Composite par rapport à des approches plus classiques. Cela est logique, car si la distance SSD (ou la corrélation de la même façon) est une bonne mesure de similarité, elle a tendance à diverger très vite dès que le motif change un peu. L’approche KLT souffre de son lien avec la SSD et également du fait que les images ne sont pas très contrastées.

Enfin, nous devons admettre que les approches par suivi d’homographie, comme la méthode ESM, sont très efficaces sur ce type d’images artificielles. Cependant, deux considérations permettent de tempérer notre jugement à leur propos : tout d’abord, on n’observe pas sur les images *in vivo* de grandes rotations ou changement d’échelle des motifs. De plus, certaines déformations observable sur ces images ne sont pas de nature homographique, et ESM a parfois tendance à diverger brusquement et à perdre sa cible au cours du suivi.

Ces aspects sont abordés au paragraphe suivant.

3.5.2 Séquences *in vivo*

3.5.2.1 Images

Pour affiner notre comparaison, nous avons évalué les différentes approches sur des séquences naturelles. La figure 3.10 illustre quelques unes de ces séquences. Celles-ci sont très variées en termes d’exposition, de résolution et de distance à l’organe, ce qui nous a permis d’évaluer de nombreux aspects de l’algorithme.

Il faut cependant nuancer nos résultats, et ce à travers plusieurs considérations :

- toutes nos images ont été acquises sur un panel de quatre cochons, soit quatre cœurs différents. Bien que nous ayons testé plusieurs éclairages, plusieurs optiques et des distances variables à l’organe, notre jeu d’images reste peu représentatif et il faudrait beaucoup plus d’échantillons pour tirer les conclusions globales sur la méthode ;
- tous les animaux considérés étaient en très bonne santé. Hors nous savons que, chez les personnes subissant des opérations de pontage, le cœur est souvent entouré de graisse. Aucun des animaux sujets de nos expérimentations ne présentait de telles particularités ;
- enfin nous n’avons pas de séquences issues de patients humains.

Pourtant, ces séquences nous permettent de comparer les différentes approches de façon plus réaliste, et soulignent certaines particularités des mouvements cardiaques.

²⁸l’approche ESM marche également très bien, mais nous voulions comparer sur cet essai les méthodes de suivi de point en priorité.

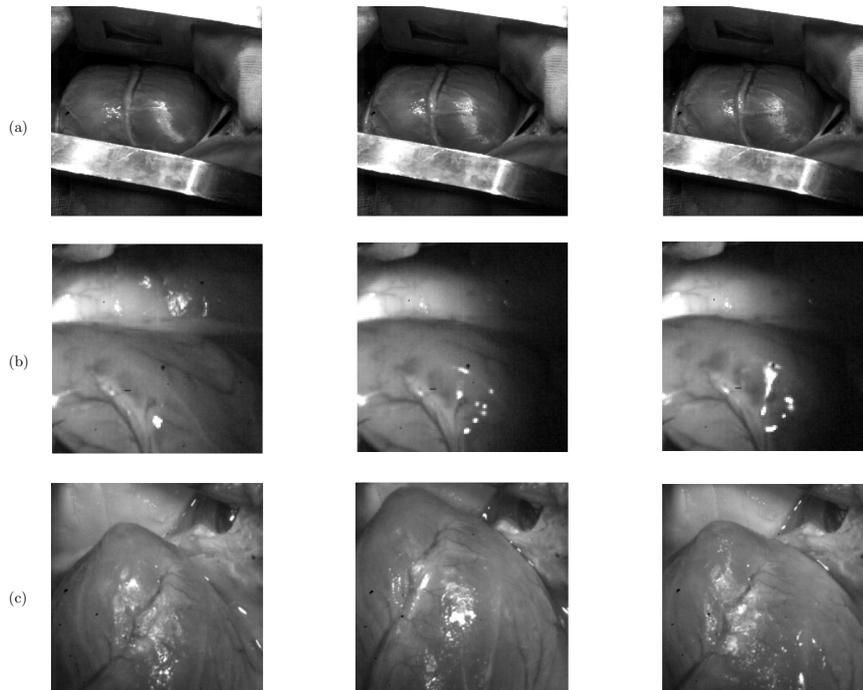


FIG. 3.10 – *Images extraites de séquences expérimentales prises à des distances, résolutions et avec des éclairages différents. (a) séquence 512×512 pixels, éclairage par scialitique chirurgical, distance : 70cm de l'organe, (b) séquence 128×128 pixels (à 250Hz), éclairage par LED, distance : 25cm de l'organe, (c) séquence 512×512 pixels, éclairage hallogène, distance : 10cm de l'organe.*

3.5.2.2 Résultats

Tout d'abord, les essais réalisés en tracking sur les séquences expérimentales montrent l'intérêt des coronaires dans le suivi de mouvement. Elles représentent en effet l'élément principal en terme de texture, face à la très forte homogénéité du reste de la surface du cœur. La figure 3.11 illustre une séquence de suivi avec notre méthode, en montrant la région d'intérêt considérée. On remarque que, dans cet exemple, le motif se déforme très peu et qu'il n'y a presque pas de composante par rotation dans le mouvement.

Même si l'on observe parfois des déformations plus importantes, la méthode de suivi Composite se comporte mieux que les méthodes classiques. Par exemple, sur la figure 3.12a on observe que les déformations induites par les mouvements cardiaques (ici une contraction du cœur) pénalisent rapidement un suivi de type SSD (en vert). L'approche Composite, par contre, est plus stable sur ce type de séquences.

Comme illustré sur la figure 3.12b, le suivi ESM a tendance à perdre la cible sur nos séquences expérimentales. La zone de recherche a alors tendance à "s'étirer" et ne correspond plus du tout à la zone initiale.

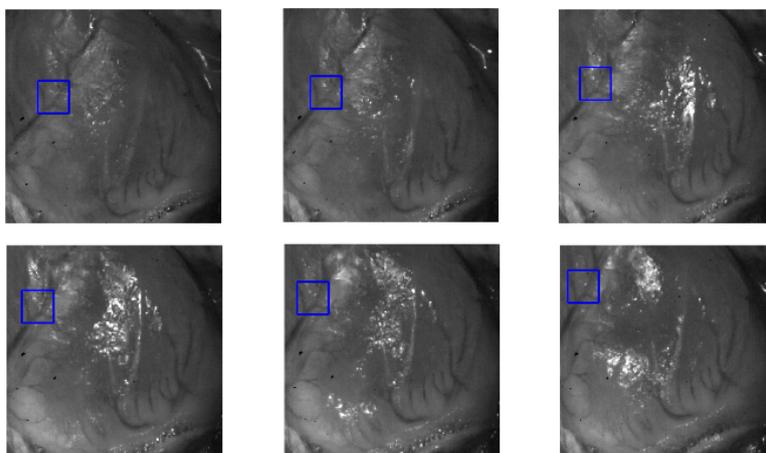


FIG. 3.11 – *Exemple de suivi Composite sur une séquence in vivo. Le carré bleu représente le motif suivi en entier.*

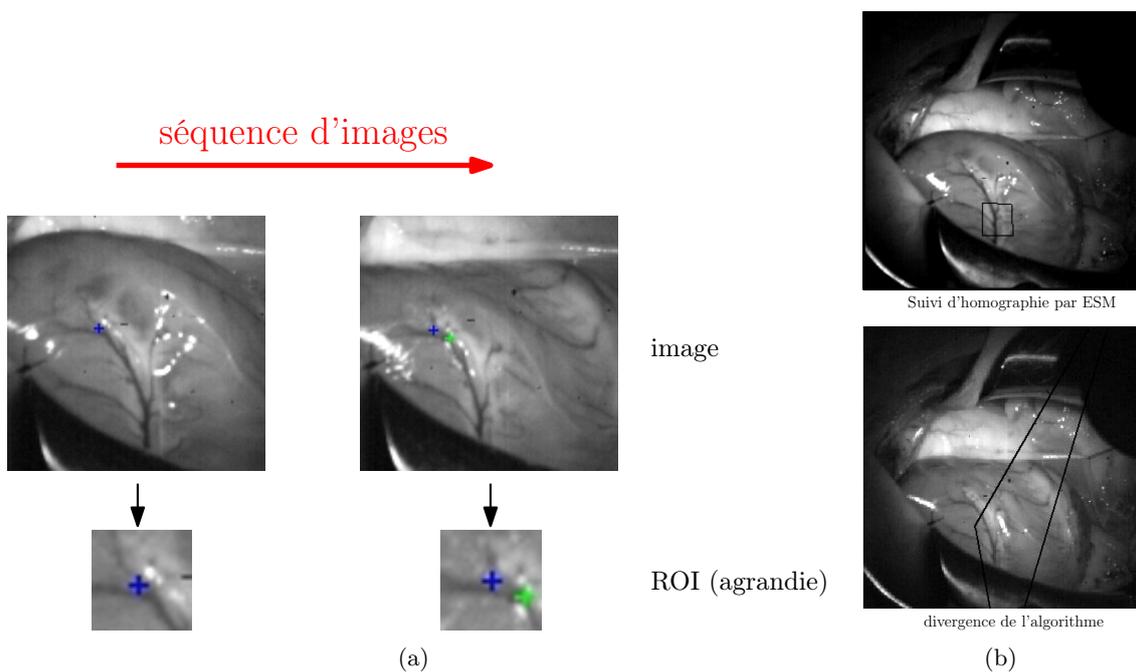


FIG. 3.12 – *Effet des déformations du motif sur le suivi. (a) comparaison du suivi SSD et du suivi Composite : l'approche par SSD, en vert, a tendance à dévier du motif original au cours de la déformation, contrairement au suivi Composite, en bleu. (b) l'algorithme ESM a tendance à perdre la cible et diverge alors rapidement.*

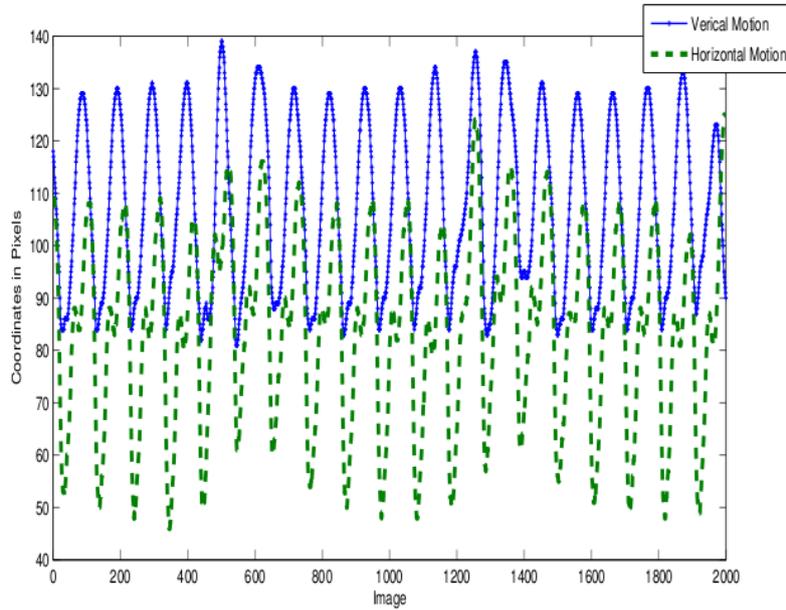


FIG. 3.13 – *Trajectoires obtenues en suivi dans l’image.* Les trajectoires verticales (trait continu) et horizontales (tirets) sont exprimées en fonction de l’indice de l’image dans la séquence.

A partir de ces mesures, il est possible de réaliser le suivi de mouvements dans l’image. La figure 3.13 montre les trajectoires obtenues avec notre algorithme, sur lesquelles on visualise facilement les cycles cardiaques.

Ces résultats sont comparables à ceux, de référence, obtenus par Nakamura²⁹, et illustrés sur la figure 3.14. Notre approche présente l’avantage de ne pas nécessiter la pose d’un marqueur sur l’organe, ce qui est précisément le but recherché. En contrepartie, le temps de calcul de 30ms par image pour un suivi de 6 points sur le cœur³⁰, est encore trop important pour une application à haute fréquence. La méthode reste cependant plus rapide qu’un suivi de type ESM, voire beaucoup plus rapide qu’un suivi de motif déformable comme [Gay-Bellile et al., 2006]. Un travail d’optimisation reste donc nécessaire, avec peut-être une adaptation aux nouveaux processeurs parallèles pour le calcul des attributs.

3.6 Conclusion

Dans ce chapitre, nous avons présenté l’état de l’art actuel des méthodes de suivi de mouvement dans les images, puis nous avons introduit une nouvelle méthode basée

²⁹figure 7 de l’article [Nakamura et al., 2001].

³⁰parfois moins lorsque la déformation de certaines régions d’intérêt est faible et que le suivi bascule en suivi SSD simple.

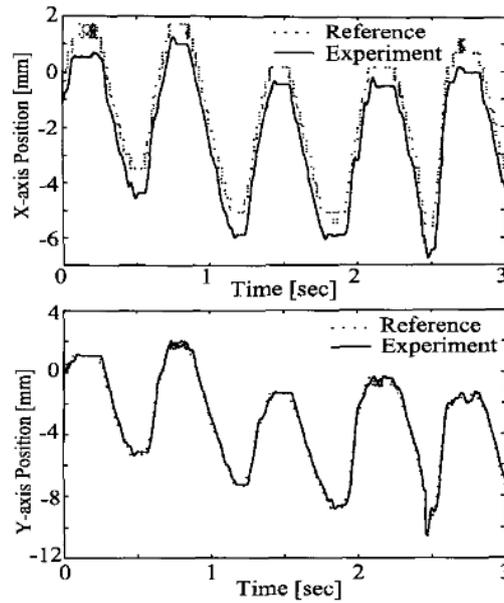


FIG. 3.14 – *Suivi de mouvements de cœur battant*, tiré de [Nakamura et al., 2001].

sur la texture. Notre approche est originale et se démarque positivement des approches existantes. Cependant, des améliorations sont encore à envisager pour obtenir un système totalement opérationnel, en particulier au niveau du temps de calcul qui doit être diminué. Enfin, vu le type de surface que nous suivons, les approches par suivi déformable [Gay-Bellile et al., 2006] pourraient à terme devenir une alternative aux méthodes actuelles. Dans cette perspective, les nouveaux types de processeurs [Sinha et al., 2006] pourraient apporter une réponse aux problèmes de performance inhérents à ces approches. Le chapitre suivant présente la problématique de l'évaluation des mouvements 3D du cœur, par l'utilisation de la stéréo-vision.

Chapitre 4

Évaluation du mouvement 3D

4.1 Introduction

Afin de pouvoir asservir le robot, il est nécessaire d'évaluer les mouvements de la surface du cœur dans l'espace cartésien. Pour ce faire, nous avons développé une solution à base de stéréo-vision. Cette solution permet d'effectuer la reconstruction d'une série de points à la surface du cœur, et ainsi d'évaluer le mouvement de l'organe.

Ce chapitre est divisé en deux parties. Tout d'abord le §4.2 présente les bases de la géométrie projective et présente les méthodes utilisées pour calibrer le système de caméras. Puis le §4.3 introduit la problématique de la reconstruction 3D et les méthodes que nous avons utilisées. Enfin le §4.4 présente nos résultats de reconstruction sur des images expérimentales de cœur battant.

4.2 Concepts fondamentaux

4.2.1 Géométries

La géométrie est une des branches des mathématiques les plus anciennes et les plus féconde. Durant les 2500 ans qui nous séparent de Thalès et Pythagore, elle a vu naître et s'éteindre nombre de théories sur des notions aussi fondamentales que les distances, les volumes ou le parallélisme. De nos jours, nombre de problèmes (et de solutions) naissent dans des intuitions géométriques¹. En vision par ordinateur, on utilise classiquement deux types de géométries² : la géométrie euclidienne et la géométrie projective.

La *géométrie euclidienne* est bien connue des physiciens, et particulièrement des mécaniciens, car elle est particulièrement bien adaptée à la description des systèmes physiques.

¹on parle ainsi d'arguments diagonaux, de projection d'espaces, de faisceaux. . .

²et accessoirement un certain nombre de géométries intermédiaires, présentées tableau 4.1.

type	ddl	matrice de transformation (coordonnées homogènes)	Invariants
<i>projective</i>	8	$\begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{pmatrix}$	intersection, colinéarité, ordre, birapport
<i>affine</i>	6	$\begin{pmatrix} a_{11} & a_{12} & t_x \\ a_{21} & a_{22} & t_y \\ 0 & 0 & 1 \end{pmatrix}$	parallélisme, rapport des aires et des segments parallèles, centres de gravité
<i>métrique</i>	4	$\begin{pmatrix} sr_{11} & sr_{12} & t_x \\ sr_{21} & sr_{22} & t_y \\ 0 & 0 & 1 \end{pmatrix}$	rapports de longueurs et d'aires
<i>euclidienne</i>	3	$\begin{pmatrix} r_{11} & r_{12} & t_x \\ r_{21} & r_{22} & t_y \\ 0 & 0 & 1 \end{pmatrix}$	longueurs, aires

TAB. 4.1 – *Hierarchie des transformations géométriques dans le plan.*

Elle repose sur trois notions fondamentales : la distance, l'aire (ou le volume) et la notion d'angle. L'objet de la géométrie euclidienne est l'étude des transformations qui laissent invariants ces paramètres : les isométries.

La *géométrie projective* quand à elle étudie une classe plus large de transformations, dites transformations projectives ou *homographies*. Cette classe est très libre, puisque la seule contrainte imposée aux transformations considérées est la conservation du birapport³

$$\frac{\overline{ca}/\overline{cb}}{\overline{da}/\overline{db}}, \text{ où } a,b,c,d \text{ sont 4 points} \quad (4.1)$$

Ainsi la géométrie projective peut être vue comme une forme “faible” de la géométrie euclidienne classique et les transformations euclidiennes peuvent être considérées comme des cas particuliers de transformations projectives. Le tableau 4.1, inspiré de [Hartley and Zisserman, 2004], résume les relations entre les différentes géométries utilisées en vision. En pratique, on utilisera la géométrie euclidienne 3D pour caractériser la position dans l'espace des différents repères : caméra, objet, robot. Au niveau des transformations entre les images, on considérera la géométrie dans l'*espace projectif* \mathcal{P}^2 .

³de manière équivalente on peut définir une transformation projective par le fait que 3 points alignés restent alignés. Cependant cette définition est moins utile en pratique.

4.2.2 Espace projectif, coordonnées homogènes et homographies

En géométrie euclidienne, on représente les points de l'espace par des vecteurs de \mathbb{R}^3 :

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3 \quad (4.2)$$

Cependant, cette représentation souffre de limitations dès que l'on tente de représenter certaines transformations élémentaires, telles que les translations. En effet il n'existe pas d'applications linéaires A telles que

$$\forall \mathbf{x} \in \mathbb{R}^3, A(\mathbf{x}) = \mathbf{x} + \mathbf{t} \quad (4.3)$$

pour $\mathbf{t} \in \mathbb{R}^3$ donné. Pour contourner ce problème, on peut choisir de représenter l'espace en *coordonnées homogènes*. Mathématiquement, cela revient à associer à un point de \mathbb{R}^3 espace euclidien un point de \mathcal{P}^3 espace projectif

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \longrightarrow \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} \propto \begin{pmatrix} k.x \\ k.y \\ k.z \\ k \end{pmatrix} \quad (4.4)$$

le symbole \propto signifiant “est proportionnel à” : en coordonnées homogènes, les points sont définis à un coefficient près. En fait les deux vecteurs de droite de l'équation (4.4), et même toute la classe de vecteurs définie par la relation d'équivalence “est proportionnel à”, représentent un seul et même point de l'*espace projectif* \mathcal{P}^3 . De fait les points de \mathcal{P}^3 ont 4 coordonnées mais seulement 3 degrés de liberté. Comme nous l'avons dit au §4.2.1, les transformations euclidiennes représentent une sous-classe des transformations projectives. Cette classe se paramétrise très facilement :

$$X' = \begin{pmatrix} \mathbf{R} & \mathbf{t} \\ 0 & 1 \end{pmatrix} \quad (4.5)$$

où \mathbf{R} est une matrice de rotation euclidienne (3×3) et \mathbf{t} un vecteur de translation.

Pour les relations 2D entre images, nous utilisons un modèle de transformation projective, c'est-à-dire une application inversible de \mathcal{P}^2 vers \mathcal{P}^2 qui conserve l'alignement de

3 points coplanaires. Dans ce cas, il existe une matrice régulière \mathbf{H} telle que⁴

$$\mathbf{x}^* \propto \mathbf{H}\mathbf{x}. \quad (4.6)$$

On a donc 8 degrés de liberté (9 composantes – 1 relation de proportionnalité) contre seulement 3 ddl pour les transformations euclidiennes du plan : 2 translations et une rotation.

4.2.3 Modèle de caméra

Dans le cadre de notre travail, nous avons considéré un modèle de caméra dit trou d'épingle (*pinhole* en anglais) ou modèle *sténopé* [Fau, Hartley and Zisserman, 2004]. On considère 4 repères principaux (figure 4.1) : le repère monde \mathcal{R}_w , le repère caméra \mathcal{R}_c placé au niveau du centre optique, le repère rétinien \mathcal{R}_r placé au niveau du plan image et le repère de l'image \mathcal{R}_i , en pixels. Soit un point M de l'espace et m ses coordonnées dans l'image. On peut découper la transformation en 3 étapes, qui vont permettre de passer par ces 4 repères. La première transformation est le passage des coordonnées de $M = (x, y, z)^t$ du repère monde dans le repère caméra. C'est une transformation rigide \mathbf{T}_1 qui se paramétrise, comme nous l'avons vu au §4.2.2, de la façon suivante

$$M_c \propto \begin{pmatrix} \mathbf{R} & \mathbf{t} \\ 0 & 1 \end{pmatrix} M = \begin{pmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \\ 0 & 0 & 0 & 1 \end{pmatrix} M. \quad (4.7)$$

Ces paramètres sont dits *paramètres extrinsèques*. L'étape suivante est la projection du point M_c sur l'image proprement dite. Cette transformation \mathbf{T}_2 peut se découper en deux étapes fondamentales : tout d'abord la projection de M_c sur le plan image, dans le repère rétinien

$$M_r \propto \begin{pmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} M_c, \quad (4.8)$$

puis on effectue un changement de repère \mathbf{T}_3 pour passer du repère rétinien aux coordonnées du point en pixels

$$m \propto \begin{pmatrix} k_u & r & u_0 \\ 0 & k_v & v_0 \\ 0 & 0 & 1 \end{pmatrix} M_r. \quad (4.9)$$

⁴en pratique, pour résoudre ce genre d'équations, on utilise un produit vectoriel caractérisant la colinéarité des vecteurs $\mathbf{x}^* \times \mathbf{x} = 0$.

Dans ces équations, f représente la distance focale de la caméra, k_u et k_v la taille des pixels et u_0 et v_0 donnent la position du repère en pixels par rapport au repère rétinien. Le paramètre r ou “skew” représente l’angle qui peut exister entre les axes \mathbf{u} et \mathbf{v} . Dans notre application nous avons considéré $r = 0$, ce qui est justifié par le fait que nous utilisons un capteur CMOS. On exprime généralement la matrice :

$$\begin{aligned}
K &= \begin{pmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} k_u & r & u_0 \\ 0 & k_v & v_0 \\ 0 & 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} k_u \cdot f & r & u_0 & 0 \\ 0 & k_v \cdot f & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \\
&= \begin{pmatrix} \alpha_u & r & u_0 & 0 \\ 0 & \alpha_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.
\end{aligned} \tag{4.10}$$

et donc

$$m = \mathbf{K} M_c \tag{4.11}$$

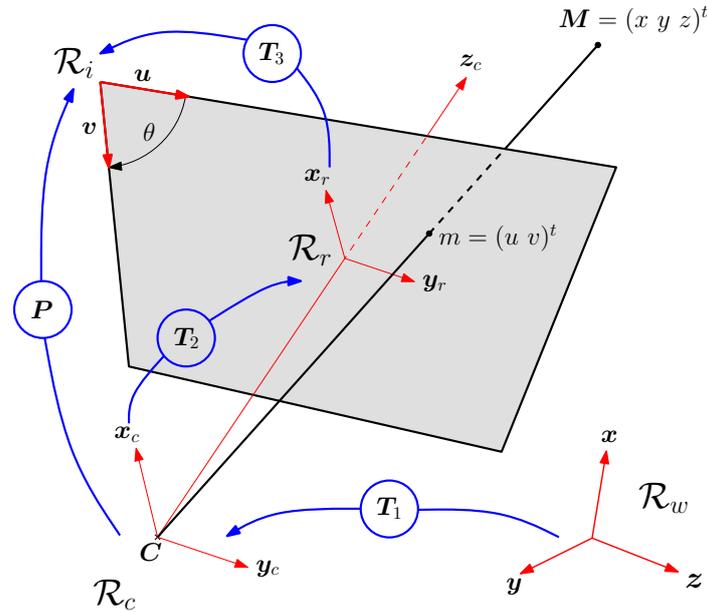
Les paramètres α_u , α_v , r , u_0 et v_0 représentent les *paramètres intrinsèques* de la caméra. On remarque au passage que l’on peut déterminer entièrement la transformation projective de la caméra sans déterminer la distance focale de celle-ci.

Enfin on peut exprimer la projection dans le repère monde (en utilisant les paramètres intrinsèques et extrinsèques) :

$$\begin{aligned}
P &= \begin{pmatrix} \alpha_u & r & u_0 & 0 \\ 0 & \alpha_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{pmatrix}
\end{aligned} \tag{4.12}$$

4.2.4 Calibrage

La calibration de la caméra est une opération permettant l’estimation des paramètres présentés au §4.2.3. Dans ce cas, on parle de *calibrage fort*. On rencontre également des

FIG. 4.1 – *Modèle de caméra et changements de repères associés.*

approches par *calibrage faible* sur les systèmes multi-caméras, c'est-à-dire que l'on calibre seulement la géométrie épipolaire de la paire stéréoscopique et on peut ainsi faire la reconstruction 3D à une homographie près.

Il existe de nombreuses méthodes pour calibrer une caméra. Certaines utilisent des objets de calibrage de dimension connues afin d'estimer les paramètres de la caméra. D'autres utilisent la rigidité de la scène pour réaliser le calibrage, on parle alors d'*auto-calibration*. Dans nos expériences, nous avons utilisé des mires planes et calibré nos caméras avec la méthode de Zhang [Zhang, 2000]. Cette méthode réalise le calibrage en se basant sur une séquence d'images d'une mire plane donc les dimensions sont connues précisément. Elle est très robuste et ne nécessite pas d'information sur la pose de l'objet de calibrage.

4.3 Reconstruction 3D

La projection d'un objet 3D dans une image 2D entraîne la perte d'une partie de l'information de position. La dimension perdue lors de la projection perspective crée une ambiguïté, illustrée sur la figure, qu'il faut lever soit en utilisant des contraintes sur l'objet reconstruit (§4.3.1), soit en utilisant plusieurs vues (§4.3.2).

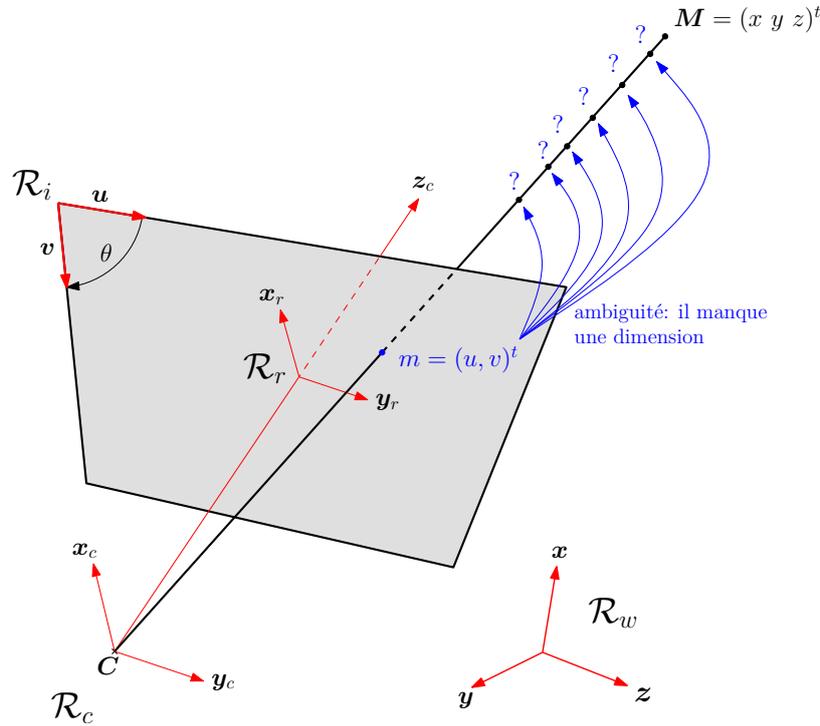


FIG. 4.2 – *Ambiguïté sur la profondeur de l'objet induite par la projection.*

4.3.1 Approches mono-caméra

4.3.1.1 Principales catégories d'approches

Les approches de reconstruction mono-caméra utilisent des contraintes géométriques sur l'objet reconstruit afin de lever l'ambiguïté présentée sur la figure 4.2. On trouve trois grandes familles de techniques utilisées pour la reconstruction mono-caméra :

- les méthodes de recalage sur un modèle CAO. Dans ce cas on utilise un algorithme itératif de type [DeMenthon and Davis, 1995] qui va reconstruire la *pose* de l'objet. Ces méthodes ont deux grands inconvénients : elles sont connues pour être sensibles aux erreurs de mesure et de modélisation, et elles ne fonctionnent que sur des objets rigides (il faut connaître précisément la géométrie de l'objet). Une application aux objets déformables serait envisageable si nous disposions d'un modèle fiable des mouvements du cœur, mais comme cela a été évoqué, nous ne sommes pas arrivés à dégager un tel modèle ;
- les méthodes dérivées de méthodes de calibrage, comme celle utilisée au LSIIT dans [Ginhoux et al., 2004, Gangloff et al., 2006] (inspirée de la méthode de Tsai [Tsai, 1986]), ou celle proposée par Mickaël Sauvée [Sauvée, 2006] (basée sur la méthode de Zhang [Zhang, 2000]). Les méthodes de calibrage permettant de retrouver la

matrice de projection à partir d'une mire particulière, on peut, une fois le système entièrement calibré, l'inverser numériquement pour retrouver la position de la mire à partir des vues. Ces approches présentent l'avantage de ne pas nécessiter un modèle complet de l'organe, vu que seule une petite primitive géométrique (typiquement un plan) est reconstruite. En contrepartie, ces approches nécessitent l'utilisation de marqueurs : c'est la contrainte supplémentaire ajoutée au système pour permettre sa résolution ;

- les *modèles d'apparence actifs*, qui sont utilisés généralement pour le suivi de visages. Ils sont très bien adaptés aux objets souples et permettent la génération de “modèles génériques” (basés sur l'analyse en composantes principales dans l'espace des images) pour permettre le recalage automatique. Cependant leur gros défaut est qu'ils ne permettent pas un suivi précis des mouvements, ce qui est une contrainte forte de notre application.

4.3.1.2 Solution étudiée

Parmi nos idées initiales pour ces travaux, l'une était le suivi d'un plan virtuel avec la méthode ESM, afin de réaliser une reconstruction par une méthode similaire à celle présentée dans [Sauvée, 2006] (basée sur la méthode de Zhang). Cette approche a posé plusieurs problèmes techniques :

- les difficultés de la méthode ESM à suivre les mouvements cardiaques ;
- l'erreur de modélisation inhérente à la méthode : on suppose que l'on suit un plan à la surface du cœur. Cette approximation est grossière suivant la région de l'organe considérée, et ce même sous une hypothèse de petits déplacements (pour les images haute fréquence). De plus, les résultats obtenus avec l'algorithme ESM sont d'autant meilleurs que la taille du patch est importante, ce qui fausse d'autant notre hypothèse ;
- le problème de la *taille* du patch suivi. En effet, la reconstruction par cette méthode implique de connaître la taille de l'objet suivi. Or nous n'avons pas de vérité terrain facilement accessible sur la géométrie de l'organe.

Le dernier point est le plus problématique. Nous avons proposé d'utiliser au début de la procédure deux caméras afin de calculer la taille du motif de référence. Puis il est possible d'enlever la seconde caméra afin d'effectuer le suivi.

Cependant cette approche paraît difficile à réaliser en pratique pour plusieurs raisons : tout d'abord il faut réaliser le calibrage de la paire stéréoscopique, tâche complexe en environnement chirurgical, qui n'est utile que pour la première reconstruction. De plus la précision obtenue n'est pas très bonne. Enfin, la méthode ESM a parfois tendance à perdre complètement sa cible lors du suivi. Cette dernière situation semble pouvoir être

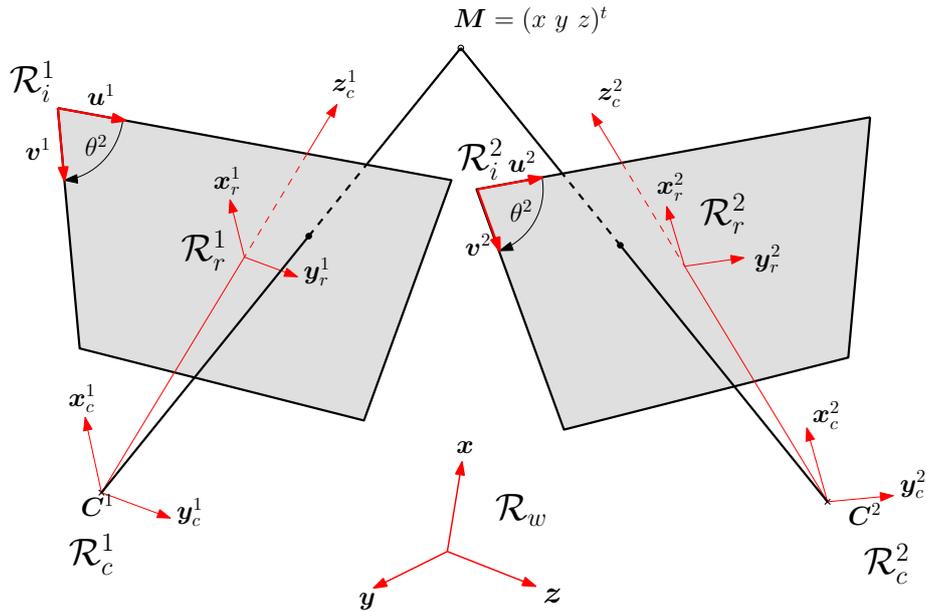


FIG. 4.3 – *Géométrie d'un capteur stéréoscopique.*

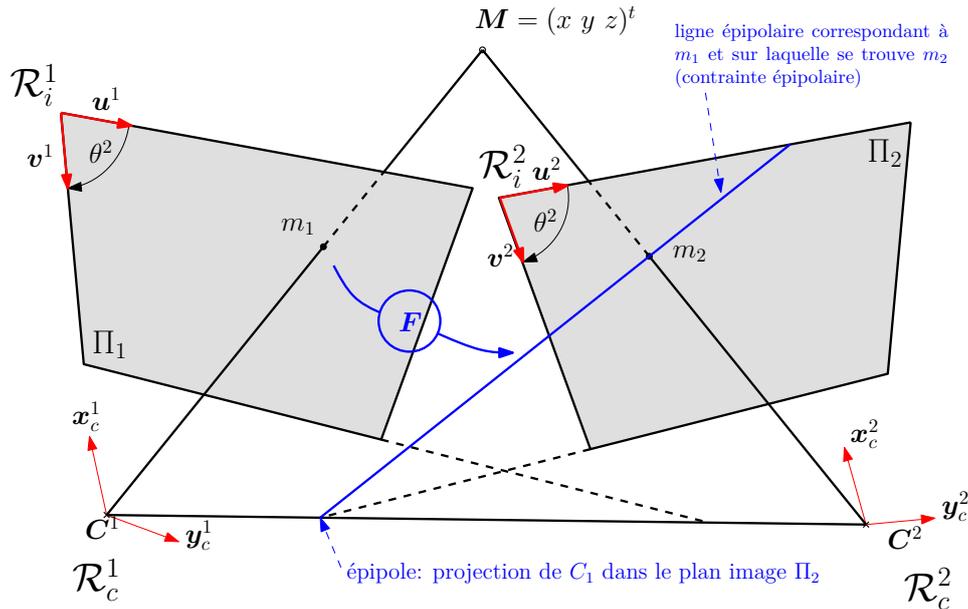
améliorée en utilisant plus d'informations : une zone de suivi plus grande, ou des images couleurs (comme dans [Bourger et al., 2007]). La première solution n'est pas applicable dans notre cas (nous cherchons une information aussi locale que possible). Le passage à la couleur constitue quant à lui une évolution possible de ce travail.

Après avoir étudié cette solution, nous avons décidé de nous concentrer sur l'approche par stéréo-vision⁵, introduite au paragraphe suivant.

4.3.2 Stéréo-vision

La *stéréo-vision* s'intéresse à la reconstruction de la position 3D d'un objet à partir de deux vues de cet objet. Il s'agit de l'approche la plus naturelle, d'un point de vue anthropomorphique, pour la perception visuelle d'environnements 3D. Elle permet entre autres de lever l'incertitude concernant la *profondeur*, et cela sans hypothèse sur la forme de l'objet (il est possible de raisonner à l'échelle du point). La figure 4.3 illustre comment la présence d'une seconde image permet de lever cette ambiguïté : la connaissance de la géométrie des caméras permet de retrouver M à l'intersection des rayons optiques.

⁵je remercie au passage Sylvain Petit, étudiant de Master 2, avec qui j'ai eu le plaisir de travailler pour mettre en place l'architecture de reconstruction stéréoscopique.

FIG. 4.4 – *Contrainte épipolaire et matrice fondamentale F.*

4.3.2.1 Géométrie des images stéréoscopiques

Dans une paire d'images stéréoscopiques, comme celle schématisée sur la Figure 4.4, on peut relier la position des points sur l'image de gauche à celle des points sur l'image de droite. C'est la *contrainte épipolaire*. Elle vient du fait que le point M , ses deux projections m_1 et m_2 , et les deux centres optiques C_1 et C_2 sont (par construction) coplanaires.

Cela entraîne une relation de dépendance linéaire entre les projections du point dans l'image de gauche et dans l'image de droite. On l'exprime par la matrice F , dite *matrice fondamentale*⁶ selon l'équation 4.13 de la contrainte épipolaire :

$$m_1^t F m_2 = 0. \quad (4.13)$$

Si, dans l'équation (4.13), on fixe le point m_1 , alors on obtient l'équation d'une droite dans la seconde image. Cette droite est la projection du rayon visuel caractérisant l'incertitude sur le point M (cf. figure 4.2). La matrice fondamentale est de rang⁷ 2 et a donc 7 degrés de liberté : 9 paramètres - 1 facteur de proportionnalité - 1 contrainte de rang 2. Ces contraintes sont très utiles en stéréo vision :

- elles facilitent l'appariement des points dans les deux images, car si l'on connaît un

⁶on trouve également dans la littérature la notion de *matrice essentielle* E . Il s'agit de la même relation exprimée dans le repère rétinien et non dans le plan image. On a donc $F = K^{-T} E K$.

⁷on force cette contrainte pendant la régression, en annulant la plus petite valeur singulière lors de la décomposition SVD.

- point on n'a plus qu'à chercher l'autre sur la droite épipolaire, cf. figure 4.4 ;
- réciproquement, elle permet de détecter des erreurs d'appariement, dans le cas de points ne vérifiant pas la contrainte épipolaire ;
- enfin, elle est liée au calibrage de la caméra : on peut évaluer F à partir des matrices de projection des deux caméras ; et on peut même se passer de calibrer les caméras et évaluer directement F lorsqu'on procède par *calibrage faible*⁸. Dans cette approche de la stéréo-vision, on cherche avant tout à calculer la matrice fondamentale, puis si nécessaire on remonte aux propriétés métriques en utilisant des indices visuels (droites parallèles, longueurs connues, etc.).

4.4 Expérimentations

4.4.1 Calibrage du système

4.4.1.1 Modèle de caméra

Nous avons effectué le calibrage fort du système stéréoscopique en utilisant une petite mire de calibrage de 5cm de côté placée à proximité du cœur. La figure 4.5 montre une séquence de calibrage réalisée au début d'une séquence in vivo. Ces séquences sont réalisées à chaque début d'expérimentation, puis la mire est retirée. Notre paire stéréoscopique étant fixe, les paramètres peuvent être considérés comme invariants.

Nos paramètres intrinsèques s'expriment donc :

$$K = \begin{pmatrix} \alpha_u & 0 & u_0 & 0 \\ 0 & \alpha_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad (4.14)$$

le paramètre de skew r est volontairement considéré comme nul, car nous utilisons directement l'image issue d'un capteur CMOS.

Nous évaluons également la distorsion radiale due à l'optique selon un modèle d'ordre quatre⁹ :

$$\hat{x}_n = x_o + x_n[k_1r^2 + k_2r^4], \quad (4.15)$$

$$\hat{y}_n = y_o + y_n[k_1r^2 + k_2r^4] \quad (4.16)$$

⁸notons que nous travaillons toujours en *calibrage fort*.

⁹au départ nous avons utilisé la *Camera Calibration Toolbox* de J.-Y. Bourget qui permet d'estimer la distorsion à l'ordre 6, mais les essais ont montré qu'un modèle d'ordre 4 est suffisant en pratique. Notre implémentation en C est directement inspirée du code Matlab de cette toolbox.

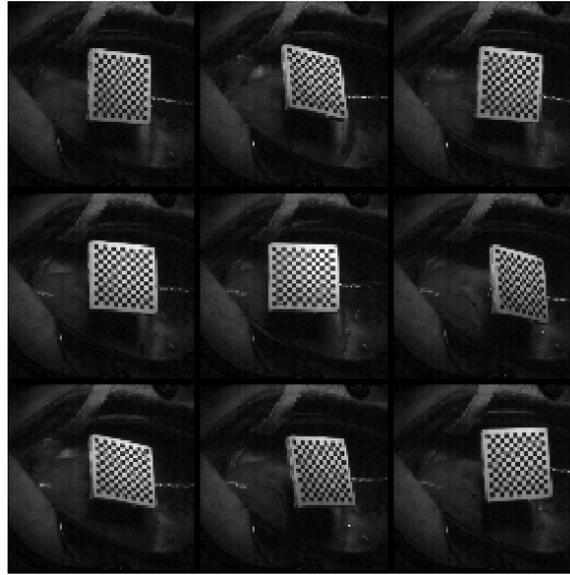


FIG. 4.5 – *Extraits d'une séquence de calibrage in vivo.* La mire de 5cm de côté spécialement réalisée pour cette expérimentation, est placée dans le champ du capteur en début d'expérimentation.

où (x_n, y_n) sont les coordonnées du points exprimées par rapport au centre (x_o, y_o) de l'image et r est la distance du point courant au centre de l'image, soit

$$r^2 = x_n^2 + y_n^2. \quad (4.17)$$

L'algorithme permet de déterminer les paramètres de distorsion k_1 et k_2 et ainsi de rectifier les images afin d'obtenir une meilleure précision. Pour cela, on procède en trois étapes :

1. calcul des paramètres intrinsèques (par la méthode de [Zhang, 2000]) sans la distorsion ;
2. projection des points de la mire (dont on la géométrie est connue) et calcul de k_1 et k_2 ;
3. nouvelle calibration par la méthode de Zhang en intégrant la correction de la distorsion dans le calcul.

Nous avons testé l'algorithme sur des mires de différentes tailles afin d'évaluer l'influence de la taille sur la précision du calibrage. Dans tous les cas, les valeurs de distorsion restent très faibles (autour de 0,2 pour $|k_1|$ et 0,1 pour $|k_2|$).

Cependant, comme nous le montrons au paragraphe suivant, la résolution de l'image influence la précision de la reconstruction.

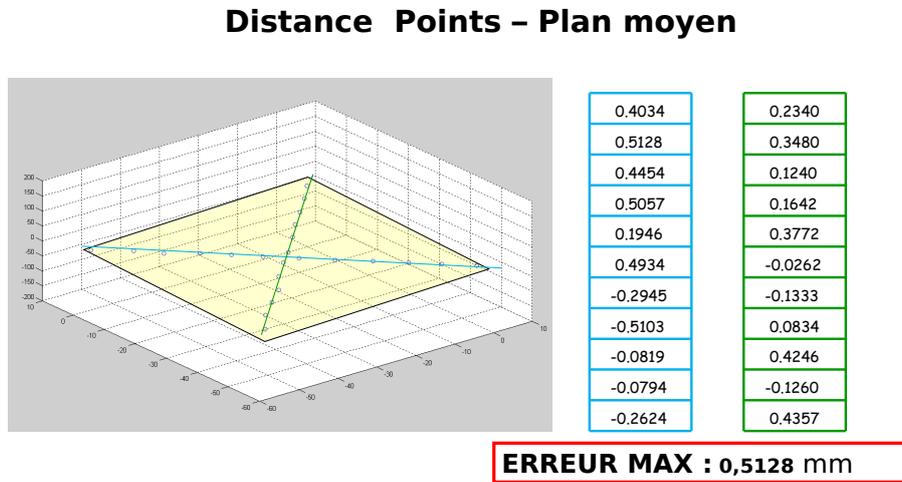


FIG. 4.6 – *Exemple de mesure de précision en utilisant la mire de calibrage. Pour des images 512×512 , sur cette série, on obtient une précision de reconstruction de l'ordre de 0,5mm.*

4.4.1.2 Mesure de la précision de la reconstruction

Une fois le système calibré, nous avons cherché à évaluer la précision de notre méthode. Comme nous ne disposons pas de vérité terrain telle que les dimensions exactes du cœur, nous avons utilisé la mire de calibrage, de géométrie connue. Notre approche se base sur un test de la planéité de la mire, comme illustré¹⁰ sur la figure 4.6. Comme nous le voyons, sur des images de résolution 512×512 nous obtenons une erreur inférieure au millimètre, et ce malgré la petite taille de la mire utilisée.

Cependant, lorsque la résolution devient plus faible (256×256 pixels et moins) l'erreur de reconstruction augmente rapidement et devient supérieure au millimètre, ce qui handicape clairement l'efficacité de la méthode. Le tableau 4.4.2.2 illustre des résultats issus du même test.

Vu l'effet de la résolution sur la fréquence d'acquisition (§3.4.1.1), il y a un réel compromis précision - fréquence d'acquisition à trouver pour ce type d'application.

¹⁰cette mesure a été effectuée sur une acquisition réalisée lors d'expérimentations du LRP, dans le cadre du travail de thèse de Marie-Aude Vitrani [Vitrani, 2006].

résolution	128 × 128	256 × 256	512 × 512	1024 × 1024
erreur maximale (plan moyen)	2,3mm	1,3mm	0,5mm	0,35mm

TAB. 4.2 – *Évolution de la précision selon la résolution des images sur le test de planéité.*

4.4.1.3 Calcul de la matrice fondamentale

Le calcul de la matrice fondamentale est primordial pour l'appariement des points correspondants entre les vues. Cette matrice peut se déduire des matrices de calibrage, mais il est conseillé¹¹ de la calculer indépendamment afin d'améliorer la précision de l'appariement [Hartley and Zisserman, 2004].

Pour réaliser ce calcul, nous utilisons les points de la mire (figure 4.5) extraits lors de la phase de calibrage. Nous disposons ainsi de 121 points (généralement un peu moins suivant la position de la mire) pour réaliser le calcul de cette matrice. Dans [Hartley and Zisserman, 2004], plusieurs méthodes de calcul de la matrice fondamentale sont présentées, dont la méthode des 8 points, qui évalue F selon la méthode des moindres carrés ($n \geq 8$) :

$$\begin{bmatrix} x_1x'_1 & y_1x'_1 & x'_1 & x_1y'_1 & y_1y'_1 & y'_1 & x_1 & y_1 & 1 \\ x_2x'_2 & y_2x'_2 & x'_2 & x_2y'_2 & y_2y'_2 & y'_2 & x_2 & y_2 & 1 \\ \vdots & \vdots \\ x_nx'_n & y_nx'_n & x'_n & x_ny'_n & y_ny'_n & y'_n & x_n & y_n & 1 \end{bmatrix} \begin{bmatrix} f_{11} \\ f_{12} \\ f_{13} \\ f_{21} \\ f_{22} \\ f_{23} \\ f_{31} \\ f_{32} \\ f_{33} \end{bmatrix} = Af = 0. \quad (4.18)$$

La matrice A est de rang au plus 8 et f est son noyau (d'après (4.18)). L'utilisation de plus de huit points assure généralement que A est de rang maximal. Comme suggéré dans [Hartley and Zisserman, 2004], nous normalisons notre nuage de points afin d'obtenir un résultat plus fiable.

Afin de pouvoir faire une résolution "classique" par moindre carrés, nous rendons le système non homogène en posant $f_{33} = 1$ (le cas $f_{33} = 0$ constituant alors une singularité). Il en découle :

¹¹voir [Hartley and Zisserman, 2004], notre référence pour les méthodes de stéréo vision.

$$(A_n \mathbf{1}) \begin{pmatrix} f_n \\ 1 \end{pmatrix} = 0 \quad (4.19)$$

avec A_n et f_n la matrice A et le vecteur f de l'équation (4.18) amputés respectivement de leur dernière ligne et dernière colonne. Le système à résoudre s'écrit finalement comme un système non homogène :

$$A_n f_n + \mathbf{1} = 0. \quad (4.20)$$

On en déduit alors f puis F par moindres carrés. Le problème est que cette matrice est généralement de rang plein, hors la matrice fondamentale doit être de rang 2. Pour renforcer cette contrainte, nous utilisons la méthode dite de Longuet-Higgins qui consiste à annuler la plus petite valeur singulière :

$$F \stackrel{\text{SVD}}{=} U \begin{pmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \end{pmatrix} V^t \quad (4.21)$$

avec $\sigma_1 \geq \sigma_2 \geq \sigma_3$ devient alors :

$$F' = U \begin{pmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & 0 \end{pmatrix} V^t \quad (4.22)$$

Cette méthode donne de très bons résultats sur nos images (les appariements, discutés au paragraphe (4.4.2), sont très bons avec la matrice fondamentale obtenue par cette méthode). Nous l'avons comparée à une méthode numérique d'estimation dite "robuste" : la méthode *RANSAC*. C'est une méthode probabiliste qui recherche un consensus :

1. l'algorithme sélectionne au hasard 8 points et estime la matrice fondamentale ;
2. puis il calcule, pour chaque point, la distance du point dans la seconde image à la droite épipolaire correspondante. Si cette distance est supérieure à un seuil (par exemple 1mm) le point est rejeté. Les autres points forment le *consensus* S_i
3. on itère le processus un nombre de fois défini à l'avance. Puis on choisit l'ensemble S_i qui a le plus grand nombre d'éléments, et on estime la matrice fondamentale avec ces points.

Cette méthode permet une estimation fiable en présence d'erreurs de mesure (jusqu'à 50% de mesures) et garantit ainsi une certaine sécurité de l'estimation. La figure 4.7 illustre un exemple de points sélectionnés par le consensus.

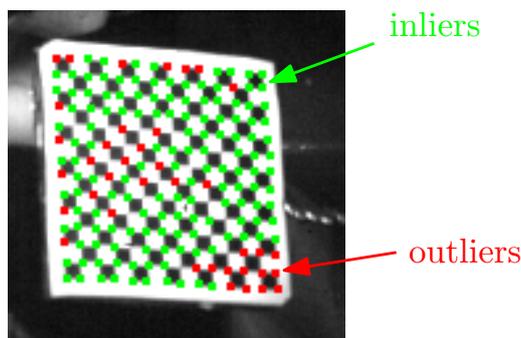


FIG. 4.7 – *Points considérés par RANSAC comme acceptables pour l'estimation, image 128×128 , tolérance 1,5mm. Les points en verts sont conservés pour l'estimation (inliers). Les points en rouge sont rejetés par RANSAC (outliers). La plupart des rejets sont dus à des échecs du détecteur de coins.*

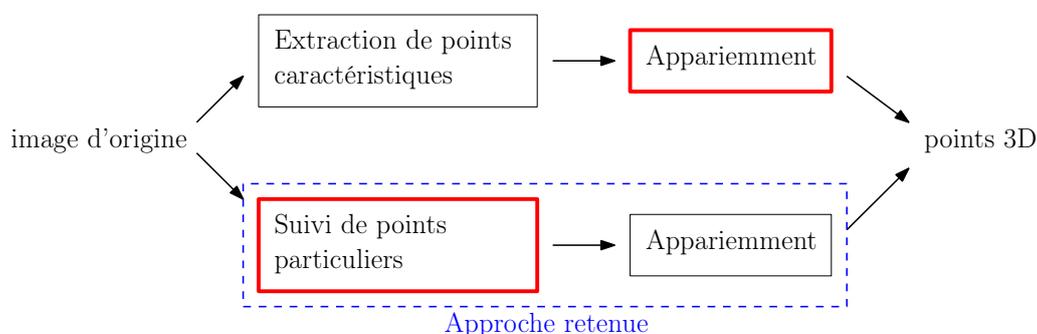


FIG. 4.8 – *Les deux approches pour la reconstruction stéréoscopique. Les étapes en sur-brillance sont celles qui sont les plus demandeuses en puissance de calcul.*

Cependant, nous n'avons pas relevé de différence notable entre les résultats fournis par cette méthode et le calcul par moindres carrés de la méthode des 8 points. Nous avons donc conservé notre méthode d'estimation, plus rapide, pour les expérimentation suivantes.

4.4.2 Appariement de Points

4.4.2.1 Différentes approches

Dans le but d'apparier les points entre les deux images, nous utilisons la contrainte épipolaire afin de guider la recherche du point dans l'image 2 à partir d'un point connu dans l'image 1. Comme évoqué précédemment, il y a deux approches pour ce type de problème, illustrées sur la figure 4.8 :

- une première consiste à extraire des points d'intérêt par un traitement bas niveau, comme les détecteurs de Harris ou KLT, puis à utiliser un classifieur pour réaliser l'appariement ;

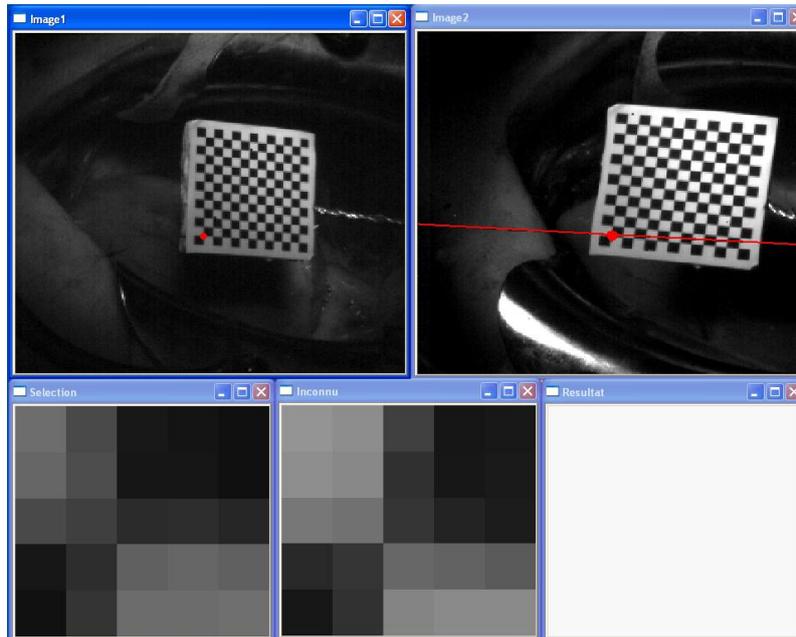


FIG. 4.9 – *Calcul des correspondances en utilisant la contrainte épipolaire. Exemple d'appariements de points de la mire sur l'implémentation OpenCV. Nous traçons dans l'image de droite la ligne épipolaire correspondant au point sélectionné dans l'image de gauche, et représentons en bas les pixels des régions d'intérêt considérées.*

- une seconde approche consiste à suivre une série de points particuliers le long d'une séquence et à les chercher dans l'autre séquence.

4.4.2.2 Notre approche

Au vu du travail présenté au chapitre 3 et comme nous cherchons à suivre les mouvements de l'organe, nous avons choisi la seconde approche. Nous utilisons donc notre algorithme de suivi dans l'image de gauche, puis nous utilisons la contrainte épipolaire introduite au §4.3.2 (figure 4.4) pour guider la recherche du point correspondant dans l'image de droite. L'utilisation de cette contrainte est illustrée sur la figure 4.9.

L'appariement des points est fait en utilisant la distance inter-texture. Nous avons également utilisé la précision sub-pixellique décrite au §3.4.3 afin d'améliorer la reconstruction, mais cela reste insuffisant pour les faibles résolutions (les résultats du tableau prennent en compte cette amélioration).

4.4.3 Reconstruction

Une fois les paramètres de la paire stéréoscopique connus et les points appariés, nous pouvons effectuer une reconstruction 3D de ces points dans le repère caméra. Là encore,

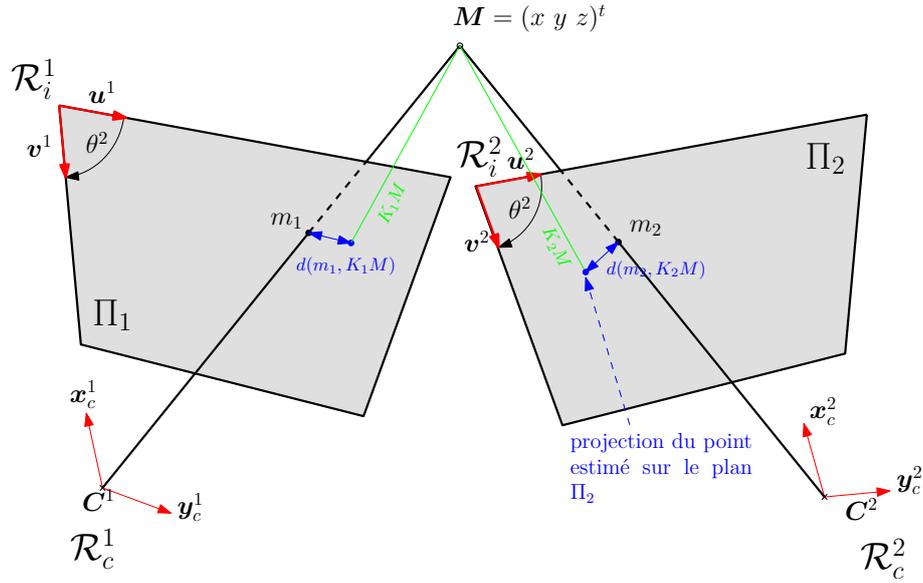


FIG. 4.10 – **Erreur de re-projection.** L'erreur de re-projection est la moyenne géométrique des distances entre les points relevés dans les images et la re-projection du point estimé dans ces images.

nous avons le choix entre plusieurs approches :

- une approche linéaire : dans laquelle on exprime, dans le repère de l'objet, les rayons visuels issus des deux caméras¹² :

$$AX = B \quad (4.23)$$

avec¹³

$$A = \begin{pmatrix} p_{11} - p_{31}u & p_{12} - p_{32}u & p_{13} - p_{33}u \\ p_{21} - p_{31}v & p_{22} - p_{32}v & p_{23} - p_{33}v \\ p'_{11} - p'_{31}u' & p'_{12} - p'_{32}u' & p'_{13} - p'_{33}u' \\ p'_{21} - p'_{31}v' & p'_{22} - p'_{32}v' & p'_{23} - p'_{33}v' \end{pmatrix}, B = \begin{pmatrix} p_{14} - u \\ p_{24} - v \\ p'_{14} - u' \\ p'_{24} - v' \end{pmatrix}.$$

Puis on résout ce système par un calcul de pseudo-inverse ;

- une approche non-linéaire, qui consiste à minimiser l'erreur de re-projection illustrée figure 4.10 par une méthode numérique, comme la méthode de Levenberg-Marquardt¹⁴ :

¹²on trouve initialement une équation de la forme $AX = 0$, que l'on transforme en équation inhomogène comme pour la méthode des huit points en posant $p_{44} = 1$.

¹³les p_{ij} et les p'_{ij} représentent les éléments des matrices globales de projection des deux caméras, (u, v) et (u', v') les coordonnées des points dans les images.

¹⁴la fonction `cvLevenbergMarquardtOptimization` d'OpenCV est très efficace pour réaliser cette estimation.

$$r = \sqrt{d^2(m_1, K_1M) + d^2(m_2, K_2M)}$$

Nous avons comparé les deux méthodes dont les résultats sont très proches pour des images ayant une résolution supérieure ou égale à 512×521 pixels. Cependant, pour des images de résolutions inférieures, la méthode non-linéaire est beaucoup plus intéressante. Ainsi, sur l'exemple de la figure 4.6, on obtient une erreur maximale de 0,51mm par la méthode non-linéaire, alors que l'erreur avec l'autre méthode est de 1,12mm. Cette erreur élevée peut s'expliquer par la mauvaise calibration des matrices lors de la résolution par la méthode linéaire.

Nous avons donc décidé d'utiliser la méthode non-linéaire pour la reconstruction des scènes 3D. La figure 4.11 illustre un cas de reconstruction de la surface proche du cœur (les vues droite et gauche sont représentées, ainsi que les points de reconstruction au centre).

Au delà de la phase de calibration, les calculs effectués restent limités. Ainsi, l'ensemble suivi - reconstruction fonctionne à une fréquence de 18 images par seconde (en moyenne 52ms pour la séquence illustrée figure 4.11), ce qui reste pour l'instant trop faible pour l'application considérée.

4.5 Conclusion

Dans ce chapitre, nous avons présenté une approche pour la mesure des mouvements 3D de la surface du cœur, basée sur la stéréo vision. Cette approche ne demande pas d'hypothèse préalable sur la nature des mouvements ou la géométrie de l'organe, et se montre efficace sur nos séquences expérimentales. Cependant, de gros efforts d'optimisation restent à faire afin d'atteindre la fréquence souhaitées de 100Hz

D'autres améliorations sont bien sûr envisageables, par exemple en intégrant un autre type de suivi tel qu'un suivi déformable [Malis, 2007, Gay-Bellile et al., 2006]. Ces méthodes prometteuses devraient permettre à terme de reconstruire directement les surfaces souples, avec une précision et une stabilité supérieures aux méthodes classiques. Une thèse est actuellement en cours au Lirmm pour évaluer et développer ces méthodes sur notre plateforme.

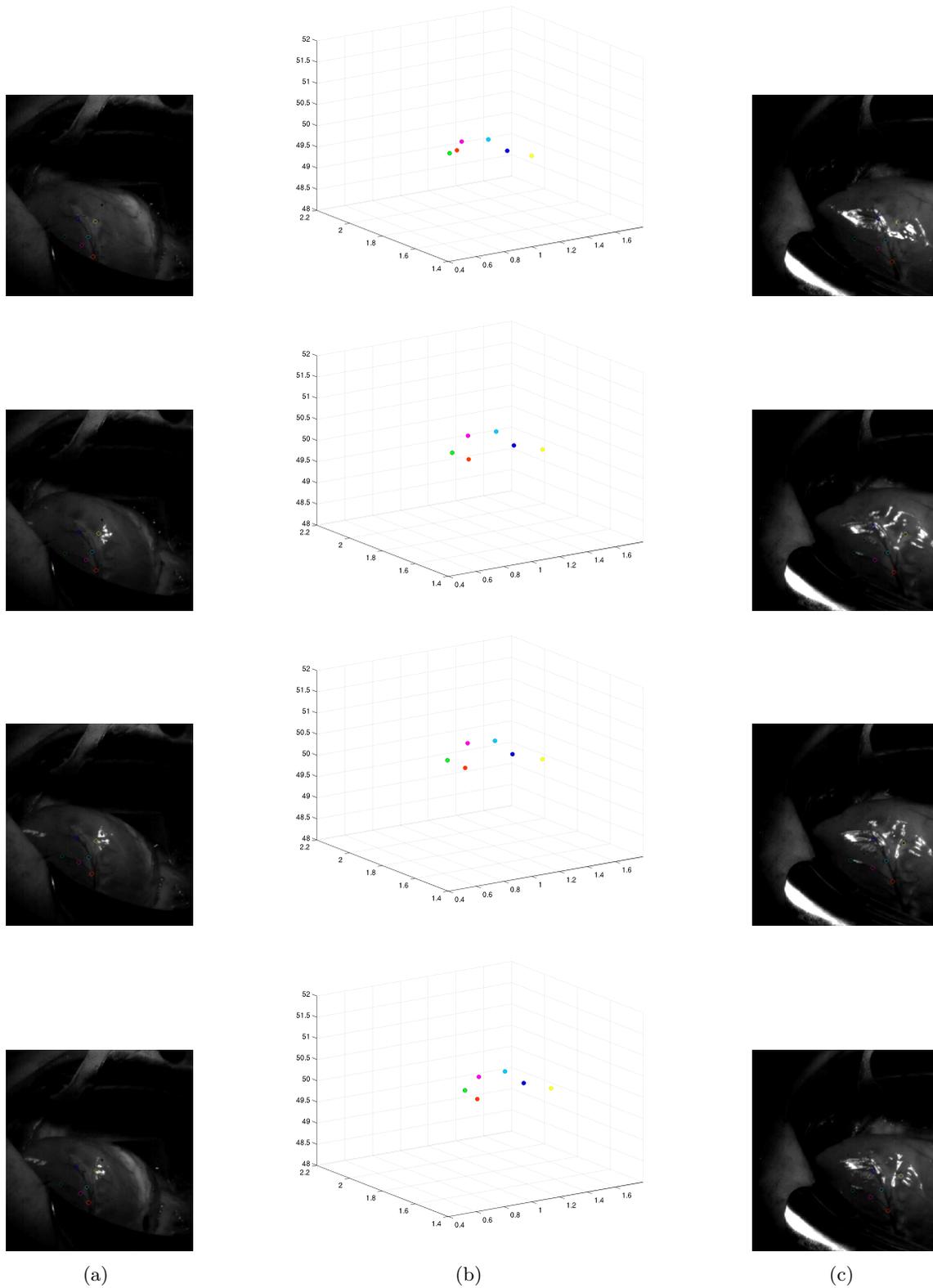


FIG. 4.11 – *Exemple de reconstruction stéréoscopique. (a) images de la caméra de gauche, (b) position 3D reconstruites, (c) images de la caméra de droite.*

Conclusion générale

Nous avons présenté dans ce manuscrit les travaux réalisés dans le contexte de l'analyse de mouvement de cœur battant pour la chirurgie cardiaque mini-invasive robotisée. Notre approche du suivi est basée sur l'analyse de texture, qui permet de pallier les limitations des approches plus classiques en termes de robustesse du suivi vis-à-vis des déformations. Ce suivi est alors intégré dans un système stéréoscopique afin de proposer une reconstruction 3D de la position de l'organe, sans nécessiter la pose de marqueurs artificiels à sa surface.

Comme nous l'avons vu, le problème du suivi comme celui de la reconstruction 3D peuvent se poser sous la forme de problème d'optimisation. Dans ce cadre, deux difficultés sont à considérer : tout d'abord le choix d'une fonction de coût qui permette de prendre en compte l'évolution des différents paramètres au cours du suivi, puis le choix d'une méthode de résolution suffisamment rapide pour une implémentation temps réel. La première difficulté a été résolue par le choix de paramètres de texture adaptés et par le développement d'une métrique utilisant cette information. C'est une approche originale qui présente, à notre avis, un fort potentiel pour ce type d'application (suivi d'organes ou d'objets déformables). La seconde difficulté est surmontée grâce à un système d'acquisition à haute fréquence, qui permet de ne considérer que des déplacements limités dans les images et de réduire les zones de recherche de l'algorithme. L'évaluation des résultats sur des séquences d'images générées artificiellement, ainsi que sur des acquisitions in-vivo, illustre l'intérêt de notre méthode par rapport à des méthodes classiques de suivi, et même vis-à-vis de méthodes plus récentes telles que celle proposée par [Malis, 2004].

La reconstruction 3D, suite logique de ce travail dans le cadre de l'asservissement du mouvement d'un robot sur celui du cœur battant permet une nouvelle fois d'utiliser l'analyse de texture pour seconder une tâche de vision. Là encore, l'analyse de texture améliore les résultats obtenus avec des méthodes plus classiques.

Cette thèse a permis de réaliser un certain nombre de développements en vision pour la plateforme cœur battant du LIRMM. Il s'agit d'un projet de grande ampleur, qui intégrera les aspects commande, téléopération et vision. Notre contribution est donc une "brique" de ce projet, essentielle dans le sens où le problème de compensation de mouvement ne pourra être résolu sans une solution efficace en termes de capteur. De ce point de vue,

l'approche par vision est idéale, car elle offre une solution sans contact et disponible sur toutes les plateformes existantes par le biais des systèmes endoscopiques.

Il reste cependant beaucoup à faire avant que ce type de méthode ne soit appliqué sur des patients. Parmi les perspectives actuelles de ce travail, on peut citer :

- l'amélioration de la robustesse du suivi, qui, si elle est meilleure que beaucoup de méthodes, n'est pas encore suffisante. Plusieurs solutions sont envisageables, comme le développement d'un modèle des mouvements de la surface du cœur (envisagé, mais pas encore développé), l'utilisation de filtres plus puissants comme les filtres particuliers pour améliorer le suivi, ou encore l'utilisation de plus de points pour le suivi ;
- l'augmentation de la fréquence d'acquisition et de traitement, par des mises à jour matérielles et des optimisations logicielles. L'arrivée massive des processeurs parallèles ouvre dans ce sens de nombreuses perspectives ;
- plus généralement, le suivi de surface déformable concentre actuellement de nombreux efforts de recherche, et de nouvelles méthodes apparaissent régulièrement dans la littérature. Cet engouement, dû tout autant au défi scientifique qu'à l'évolution naturelle de la communauté vers ces approches, maintenant que les solutions pour le suivi d'objets rigide ont atteint une certaine maturité, devrait amener d'importants développements dans les années à venir. Il est donc primordial de suivre l'émergence de nouvelles méthodes [Gay-Bellile et al., 2007, Bosch et al., 2007], sinon pour évoluer vers celles-ci, du moins pour évaluer la pertinence de notre approche dans le temps.

Enfin, au-delà des aspects vision, l'implémentation de lois de commandes, utilisant le logiciel développé dans cette thèse, sur la plate-forme robotique du LIRMM, est indispensable afin d'évaluer la viabilité de l'approche. Deux thèses sont actuellement en cours au LIRMM (une sur la commande robuste et une sur les méthodes de suivi déformable) afin de faire évoluer la plate-forme de robotique médicale vers un système complet, téléopéré et intégrant la compensation de mouvements par asservissement visuel. Nous espérons que les bases construites durant cette thèse aideront à atteindre ces objectifs.

Bibliographie

- Site virtual medical world. URL <http://www.hoise.com/vmw/00/articles/vmw/LV-VM-02-00-5.html>. 2000.
- IEEE Transactions On Robotics and Automation : Special Issue on Medical Robotics*, volume 19. Oct. 2003.
- P. Abolmaesumi, S. Salcudean, W. Zhu, M. Sirouspour, and S. DiMaio. Image-guided control of a robot for medical ultrasound. *IEEE Transactions on Robotics and Automation*, 18(1) :11–23, Feb. 2002.
- E. H. Adelson, C. H. Anderson, J. R. Bergen, P. J. Burt, and J. M. Ogden. Pyramid methods in image processing. Technical Report 6, Nov. 1984.
- S. Aiono, J. Gilbert, B. Soin, P. Finlay, and A. Gordon. Controlled trial of the introduction fo a robotic camera assistant (EndoAssist) for laparoscopic cholesytectomy. In *Proceedings of the Annual Scientific Meeting, Society for Minimally Invasive Therapy*, Boston, USA, Sept. 1999.
- B. Alacam, B. Yazici, and N. Bilgutay. Breast cancer detection based on ultrasound b-scan texture analysis and patient age information. In *Proceedings of IEEE 29th Annual Bioengineering Conference* , pages 98–99, Mar. 2003.
- M. Allaf, S. Jhackman, P. Shulam, J. Cadeddu, B. Lee, R. Moore, and L. Kavoussi. Laparoscopic Visual Field : Voice versus Foot Pedal Interface for Control of the AESOP Robot. *Surgical Endoscopy*, 12(12) :1415–1418, Dec. 1998.
- A. Asano, M. Miyagawa, and M. Fujio. Texture modelling by optimal gray scale structuring elements using morphological pattern spectrum. In *Proceedings of the 15th International Conference on Pattern Recognition*, volume 3, pages 475–478, Sept. 2000.
- N. Ayache. RR3364 - l’analyse automatique des images médicales état de l’art et perspectives. Technical report, INRIA Sophia Antipolis, Feb. 1999.

- S. Baker and I. Matthews. Lucas-Kanade 20 Years On : A Unifying Framework : Part 1, technical report CMU-RI-TR-02-16. Technical report, Robotics Institute, Carnegie Mellon University, July 2002.
- D. Benboudjema. *Champs de Markov triplets et segmentation bayésienne non supervisée d'images*. PhD thesis, Institut National des Télécommunications, Feb. 2005.
- S. Benhimane and E. Malis. Homography-based 2D Visual Tracking and Servoing. *International Journal of Robotic Research (Special Issue on Vision and Robotics joint with the International Journal of Computer Vision)*, 26(7) :661–676, July 2007.
- A. Blake and M. Isard. Condensation - CONDitional DENsity propagaTION for visual tracking. *International Journal of Computer Vision*, 29(1) :5–28, Aug. 1998.
- L. Bonilha, E. Kobayashi, G. Castellano, G. Coelho, E. Tinois, F. Cendes, and L. M. Li. Texture Analysis of Hippocampal Sclerosis. *Epilepsia*, 44(12) :1546–1550, 2003.
- A. Bosch, A. Zisserman, and X. Munoz. Image classification using random forests and ferns. In *Proceedings of the 11th International Conference on Computer Vision (ICCV'07)*, Rio de Janeiro, Brazil, Oct. 2007.
- F. Bourger, C. Doignon, P. Zanne, and M. de Mathelin. Model-free Vision-based Robot Control for Minimally Invasive Surgery using ESM Tracking and Pixels Color Selection. In *Proceedings of the IEEE International Conference On Robotics and Automation (ICRA'07)*, pages 3579–3584, Roma, Italy, Apr. 2007.
- M. Cavusoglu and F. Tendrick. A laparoscopic telesurgical workstation. *IEEE Transactions on Robotics and Automation*, 15(4) :728–739, Aug. 1999.
- M. Cavusoglu, W. William, F. Tendrick, and S. Sastry. Robotics for telesurgery : second generation Berkeley/UCSF laparoscopic telesurgery workstation and looking towards the future applications. In *Proceedings of the Allerton Conference on Communication, Control and Computing*, Monticello, USA, 2001.
- S. Chen, J. Keller, and R. Crownover. On the Calculation of Fractal Features from Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(10) :1087–1090, Oct. 1993.
- D. Cianflone, MD, D. R. MacLean, MD, T. A. PEarson, MD, S. Stachenko, MD, and V. B. Harris. Déclaration de Milan : Pour un positionnement de la technologie au service de la santé cardiovasculaire mondiale. 5e Conférence internationale sur la santé cardiovasculaire, June 2004. Enter text here.

- D. A. Clausi and Y. Zhao. Rapid determination of co-occurrence texture features. In *Proceedings of the IEEE Geoscience and Remote Sensing Symposium (IGARSS'01)*, volume 4, pages 1880–1882, Seoul, Korea, July 2001.
- J. P. Cocquerez, S. Philipp, P. Bolon, J. M. Chassery, D. Demigny, C. Graffigne, A. Montanvert, R. Zéboudj, and J. Zérubia. *Analyse d'Images : Filtrage et segmentation*. Masson, 1995.
- A. Comport, M. E., and C. F. A real-time tracker for markerless augmented reality. In *Proceedings of the ACM/IEEE International Symposium on Mixed and Augmented Reality for Minimally Invasive Therapy (ISMAR'03)*, pages 36–45, Tokyo, Japan, Oct. 2003.
- E. Corbillon and A.-I. Pouillé. La chirurgie assistée par ordinateur - rapport d'étape. Technical report, Agence Nationale d'Accréditation et d'évaluation en Santé (ANAES), May 2002.
- A. Cornuéjols, L. Miclet, and Y. Kodratoff. *Apprentissage artificiel : concepts et algorithmes*. Eyrolles, 2003.
- E. Cuevas, D. Zaldivar, and R. Rojas. Kalman filter for vision tracking, technical report b 05-12. Technical report, Freie Universität Berlin, Institut für Informatik, Aug. 2005.
- L. Cuvillon. *Compensation du battement cardiaque en Chirurgie Robotisée*. PhD thesis, Université Louis Pasteur - Strasbourg I, June 2006.
- B. Davies. A review of robotics in surgery. In *Proceedings of the Institution of Mechanical Engineers*, volume 214, pages 129–140, 2000.
- C. Delgorge, F. Courrèges, L. A. Bassit, C. Novales, C. Rosenberger, N. Smith-Guerin, C. Brù, R. Gilabert, M. Vannoni, G. Poisson, and P. Vieyeres. A tele-operated mobile ultrasound scanner using a light weight robot. *IEEE Transactions on Information Technology in Biomedicine*, 9(1) :50–58, Mar. 2005.
- D. DeMenthon and L. S. Davis. Model-Based Object Pose in 25 Lines of Code. *International Journal of Computer Vision*, 15(1-2) :123–141, June 1995.
- M. Dhome, M. Richetin, J.-T. Lapresté, and Rives. G. Determination of the attitude of 3D objects from a single perspective view. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(12) :1265–1278, Dec. 1989.
- M. Do and M. Vetterli. Rotation invariant texture characterization and retrieval using steerable wavelet-domain hidden markov models. *IEEE Transactions on Multimedia*, 4(12) :517–527, Dec. 2002.

- E. Dombre, G. Duchemin, P. Poignet, and F. Pierrot. Dermarob. A safe robot for reconstructive surgery. *IEEE Transactions on Robotics and Automation, Special Issue on Medical Robotics*, 19(5) :876–884, Oct. 2003.
- E. Dombre, M. Michelin, F. Pierrot, P. Poignet, P. Bidaux, G. Morel, T. Ortmaier, D. Sallé, N. Zemiti, P. Gravez, M. Karouia, and N. Bonnet. Marge project : design, modelling, and control of assistive devices for minimally invasive surgery. In *Proceedings of Medical Image Computing and Computer-Assisted Interventions (MICCAI'04)*, pages 1–8, S^t Malo, France, Sept. 2004.
- G. Duchemin, P. Poignet, E. Dombre, and F. Pierrot. The challenge of designing and manufacturing of actuated medical robots for safe human interaction. *IEEE Robotics and Automation Magazine, Special Issue on Robot Dependability*, 11(2) :46–55, June 2004.
- P. Dulyakarn, Y. Rangsanseri, and Thitimajshima. Comparison of Two Features for Multispectral Imagery Analysis. In *Proceeding of Asian Conference of Remote Sensing (ACRS'2000)*, Taipei, Taiwan, Dec. 2000.
- M. E., L. M., D. F., and S. P. 3D reconstruction of complex structures with bundle adjustment : an incremental approach. In *Proceedings of the 2006 IEEE International Conference on Robotics and Automation*, pages 3055–3061, Orlando, Florida, Sept. 2007a.
- R. E., L. M., D. F., and C. T. Monocular Vision for Mobile Robot Localization and Autonomous Navigation. *International Journal of Computer Vision*, 74(3) :237–260, Sept. 2007b.
- J. Gangloff, R. Ginhoux, M. de Mathelin, L. Soler, and J. Marescaux. Model predictive control for compensation of cyclic organ motions in teleoperated laparoscopic surgery. *IEEE Transactions on Control Systems Technology*, 14(2) :235–246, Mar. 2006.
- V. Gay-Bellile, M. Perriollat, A. Bartoli, and P. Sayd. Image Registration by Combining Thin-Plate Splines With a 3D Morphable Model. In *Proceedings of the International Conference on Image Processing (ICIP'06)*, Atlanta, GA, USA, Oct. 2006.
- V. Gay-Bellile, A. Bartoli, and P. Sayd. Gestion des auto-occultations pour l'augmentation d'une surface déformable sur une vidéo. In *Actes du colloque COmpression et REpresentation des Signaux Audiovisuels (CORESA '07)*, Montpellier, France, Nov. 2007.
- R. Ginhoux. *Compensation des mouvements physiologiques en chirurgie robotisée par commande prédictive*. PhD thesis, Université Louis Pasteur - Strasbourg I, 2003.

- R. Ginhoux, J. A. Gangloff, M. F. de Mathelin, L. Soler, M. M. A. Sanchez, and J. Marescaux. Beating heart tracking in robotic surgery using 500 hz visual servoing, model predictive control and an adaptive observer. In IEEE, editor, *Proceedings of International Conference on Robotics and Automation (ICRA '04)*, volume 1, pages 274–279, New Orleans, LA, USA, Apr. 2004.
- M. Gröger, W. Sepp, T. Ortmaier, and G. Hirzinger. Reconstruction of image structure in presence of specular reflections. In B. Radig and S. Florczyk, editors, *Proceedings of the 23rd DAGM Symposium on Pattern Recognition*, pages 53–60, Munich, Germany, Sept. 2001.
- G. S. Guthart and J. K. Salisbury. the intuitive telesurgery system : overview and applications. In *Proceedings of IEEE International Conference on Robotics and Automation (ICRA '00)*, volume 1, pages 618–621, San Francisco, USA, Apr. 2000.
- R. Haralick. Statistical and Structural Approches to Textures. *Proceedings of IEEE*, 5 (67) :786–804, Apr. 1979.
- S. J. Harris, Q. Mei, F. Arambula-Cosio, R. D. Hibberd, S. Nathan, Wickham, and B. L. Davies. Design of a robotic prostatectomy system. In *Proceedings of Medimec'95 Second International Workshop of Mechatronics in Medicine and Surgery*,, pages 7–14, Sept. 1995.
- S. J. Harris, F. Arambula-Cosio, Q. Mei, R. D. Hibberd, B. L. Davies, J. E. A. Wickham, M. S. Nathan, and B. Kundu. The Probot - an active robot for prostate resection. *Proceedings of the Institution of Mechanical Engineers, Part H : Journal of Engineering in Medicine*, 211(H4) :317–326, Sept. 1997a.
- S. J. Harris, W. J. Lin, R. D. Hibberd, J. Cobb, R. Middleton, and B. L. Davies. Experiences with robotic systems for knee surgery. In Springer Verlag, editor, *Proceedings of the First Joint Conference on Computer Vision, Virtual Reality and Robotics in Medicine (CVRMed) and Medial Robotics and Computer-Assisted Surgery (MRCAS)*, volume 1205, pages 757–766, Mar. 1997b.
- R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2nd edition, 2004.
- D. J. Heeger and J. R. Bergen. Pyramid-Based texture analysis/synthesis. In Robert Cook, editor, *Proceedings of the 22th Conference and Exhibition on Computer Graphics and Interactive Techniques (SIGGRAPH'95)*, pages 229–238, Los Angeles, CA, USA, Aug. 1995. ACM SIGGRAPH, Addison Wesley.

- S. Herlidou. *Caractérisation tissulaire en Imagerie par Résonance Magnétique Nucléaire par l'analyse de texture : étude du tissu musculaire et de tumeurs intracrâniennes*. PhD thesis, Université de Rennes 1, Oct. 1999.
- J. Hong, T. Dohi, M. Hashizume, K. Konisji, and N. Hata. An ultrasound-driven needle insertion robot for percutaneous cholecystostomy. *Physics in Medicine and Biology*, 19 (3) :441–455, Feb. 2004.
- T.-I. Hsu, A. D. Calway, and R. Wilson. Texture Analysis Using the Multiresolution Fourier Transform. In *Proceedings of the 8th Scandinavian Conference on Image Analysis*, pages 823–830. IAPR, 1993.
- S. Hutchinson, G. Hager, and I. Corke. A Tutorial on Visual Servo Control. *IEEE Transactions on Robotics and Automation*, 12(5) :651–670, Oct. 1996.
- K. Ikuta, S. Daifu, T. Hasegawa, and H. Higashikawa. Hyper-finger for remote minimally invasive surgery in deep area. In *Proceedings of Medical Image Computing and Computer-Assisted Intervention (MICCAI'02)*, pages 173–181, Tokyo, Japan, 2002.
- K. Jafari-Khouzani, M.-R. Siadat, H. Soltanian-Zadeh, and K. Elisevich. Texture analysis of hippocampus for epilepsy. *Proceedings of the SPIE*, 5031 :279–288, May 2003.
- D. Jaffray, D. Herk, J. Lebesque, and A. Martinez. Image guided radiotherapy of the prostate. In *Proceedings of Medical Image Computing and Computer Intervention (MICCAI'01)*, pages 1075–1080, Utrecht, Holland, 2001.
- B. Jähne. *Digital Image Processing*. Springer-Verlag, 2002.
- R. Jennane, R. Harba, G. Lemineur, S. Bretteil, A. Estrade, and C. L. Benhamou. Estimation of the 3d self-similarity parameter of trabecular bone from its 2d projection. *Medical Image Analysis*, 11 :91–98, Feb. 2007.
- S.-S. Jingaposao and C. Anni. Fractal based texture analysis for retrieval of image data. In *Proceedings of Fifth Asia-Pacific Conference on Communications and Fourth Optoelectronics and Communications Conference*, volume 2, pages 845–848, 1999.
- B. Julesz. Experiments in the Visual Perception of Textures. *Scientific American*, 232 : 34–43, 1975.
- S. Julier and J. Uhlmann. A new extension of the Kalman filter to nonlinear systems. In *International Symposium on Aerospace/Defense Sensing, Simulation and Controls*, Orlando, FL, 1997.

- R. E. Kalman. A new approach to linear filtering and prediction problems. *Transactions of the AMSE, Journal of Basic Engineering*, 82(Series D) :35–45, 1960.
- T. Kienzle, S. Stulberg, M. Peshkin, A. Quaid, J. Lea, A. Goswami, and W. Chi-Haur. Total knee replacement. *IEEE Engineering in Medicine and Biology Magazine*, 14(3) : 301–306, May 1995.
- A. Krupa. *Commande par vision d'un robot de chirurgie laparoscopique*. PhD thesis, Institut National Polytechnique de Lorraine, July 2003.
- R. Kumar, P. Berkelman, G. P., A. B. P. Jensen, L. Whitcomb, and R. Taylor. Preliminary experiments in cooperative human/robot force control for robot assisted microsurgical manipulation. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICuufacturingRA'00)*, pages 610–617, San Fransisco, USA, 2000.
- Y. Kwoh, J. Hou, E. Jonckheere, and S. Hayati. A robot with improved absolute positionning accuracy for CT guided stereostatic brain surgery. *IEEE Transactions on Biomedical Engineering*, 32(2) :153–160, Feb. 1988.
- J. Lacroix. Processus stationnaires et prévision, cours de l'université pierre et marie curie, 2004.
- A. Lange. Simultaneous Statistical Calibration of the GPS signal delay measurements with related meteorological data. *Physics and Chemistry of the Earth, Part A : Solid Earth and Geodesy*, 26(6-8) :471–473, Jan. 2001.
- M. Lemma, A. Mangini, A. Redaelli, and F. Acocella. Do cardiac stabilizers really stabilize? Experimental quantitative analysis of mechanical stabilization. *Interactive CardioVascular and Thoracic Surgery*, 4 :222–226, Mar. 2005.
- L. Lepistö, I. Kunttu, and A. VISA. Retrieval of non-homogeneous textures based on directionality. In *Proceedings of the 4th International Workshop on Image Analysis for Multimedia Interactive Services, (WIAMIS'2003)*, Mar. 2003.
- J. Levy Vehel. RR1707 - texture analysis using fractal probability functions. Technical report, INRIA Rocquencourt, 1992.
- J. Levy Vehel, P. Mignot, and J. Berroir. RR1706 - texture and multifractals : new tools for image analysis. Technical report, INRIA Rocquencourt, 1992.
- A. Loh. *The recovery of 3-D structure using visual texture patterns*. PhD thesis, University of Western Australia, Feb. 2006.

- A. Lorette. *Analyse de texture par méthodes markoviennes et par morphologie mathématique : application à l'analyse des zones urbaines sur des images satellitales*. PhD thesis, Université de Nice Sophia Antipolis, 1999.
- D. G. Lowe. Object recognition from local scale-invariant features. In *Proceedings of International Conference on Computer Vision (ICCV'99)*, pages 1150–1157, Corfu, Greece, Sept. 1999.
- B. D. Lucas and T. Kanade. An Iterative Image Registration Technique with an Application to Stereo Vision. *Proceedings of the 7th International Joint Conference on Artificial Intelligence (IJCAI '81)*, pages 674–679, Apr. 1981.
- J. Mackay and G. A. Mensah. Atlas of Heart Disease and Stroke. Technical report, World Health Organization, 2005. URL http://www.who.int/cardiovascular_diseases/resources/atlas/en/.
- P. Maillet, B. Nahum, L. Blondel, P. Poignet, and E. Dombre. Brigit, a robotized tool guide for orthopedic surgery. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA'05)*, pages 211–216, Barcelone, Espagne, Apr. 2005.
- E. Malis. An efficient unified approach to direct image registration of rigid and deformable surfaces. Technical Report 6089, INRIA, Jan. 2007.
- E. Malis. Improving vision-based control using efficient second-order minimization techniques. In *Proceedings of IEEE International Conference on Robotics and Automation (ICRA'04)*, volume 2, New Orleans, LA, USA, Apr. 2004.
- S. Mallat. *A wavelet tour of signal processing*. Éditions de l'École Polytechnique, 1999.
- J. Marescaux, J. Leroy, and M. Gagner. “opération lindbergh” une première mondiale en télé-chirurgie : le geste chirurgical a traversé l'atlantique! Technical report, IRCAD, France Télécom, Computer Motion, Sept. 2001.
- A. Materka, P. Cichy, and J. Tuliszkiwicz. Texture Analysis of X-Ray Images for Detection of Changes in Bone Mass and Structure. In *Proceedings of Infotech Oulu Workshop on Texture Analysis in Machine Vision*, pages 29–34, Oulu, Finland, jun 1999.
- B. Maurin, L. Barbe, B. Bayle, P. Zanne, J. Gangloff, M. de Mathelin, A. Gangi, L. Soler, and F. Forgione. In vivo study of forces during needle insertions. In *Proceedings of Medical Robotics, Navigation and Visualisation Scientific Workshop (MRNV'04)*, Remagen, Germany, Mar. 2004.

- T. Mengko and J. Tjandra Pramudito. Implementation of Gabor filter to texture analysis of radiographs in the assessment of osteoporosis. In *Proceedings of Asia-Pacific Conference on Circuits and Systems (APCCAS'02)*, volume 2, pages 251–254, Aug. 2002.
- M. E. Munich, P. Pirjanian, E. Di Bernardo, L. Goncalves, N. Karlsson, and D. Lowe. SIFT-ing through features with ViPR. *IEEE Robotics and Automation Magazine*, 13(3) :72–77, Sept. 2006.
- Y. Nakamura, K. Kishi, and H. Kawakami. Heartbeat Synchronization for Robotic Cardiac Surgery. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA'01)*, volume 2, pages 2014–2019, Seoul, Korea, May 2001.
- A. Nasri and M. El Hitmy. Algorithme Génétique et Critère de la Trace pour l'Optimisation du Vecteur Attribut : Application à la Classification Supervisée des Images de Textures. In *Proceedings of the 15th International Conference on Vision Interface (VI'2002)*, pages 407–412, Calgary, Canada, Apr. 2002.
- R. Negenborn. Robot Localization and Kalman Filters : On finding your position in a noisy world. Master's thesis, Utrecht University, Sept. 2003.
- A. Noce, J. Triboulet, P. Poignet, and E. Dombre. Features Selection for Visual Servoing of the Beating Heart. In *Proceedings of the IEEE/RAS-EMBS BIOROB'06*, number 183, Pisa, Italy, Feb. 2006.
- E. Oja and K. Valkealahit. Reduced Multidimensional Histograms in Color Texture Description. In *Proceedings of the International Conference on Pattern Recognition (ICPR'98)*, volume 2, pages 1057–1061, Brisban, Australia, Aug. 1998.
- T. J. Ortmaier. *Motion Compensation in Minimally Invasive Robotic Surgery*. PhD thesis, Technischen Universitat Munchen, 1999.
- G. Peyré. *Géométrie multi-échelles pour les images et les textures*. PhD thesis, École Polytechnique, 2005.
- W. Pieczynski and A.-N. Tebbache. Pairwise Markov random fields and its application in textured images segmentation. In *Proceedings of the 4th IEEE Southwest Symposium on Image Analysis and Interpretation*, pages 106–110, 2000.
- F. Pierrot, E. Dombre, E. Dégoulange, L. Urbain, P. Caron, S. Boudet, J. Gariépy, and J. Megnien. Hippocrate : A Safe Robot Arm for Medical Applications with Force Feedback. *Medical Image Analysis*, 3(3) :285–300, 1999a.

- F. Pierrot, E. Dombre, L. Dégoulange, L. Urbain, P. Caron, S. Boudet, J. Gariepy, and J. Megnien. Hippocrate : a safe robot arm for medical applications with force feedback. *Medical Image Analysis*, 3(3) :285–300, 1999b.
- R. Polet and J. Donnez. Gynecologic laparoscopic surgery with a palm-controlled laparoscope holder. *The Journal of the American Association of Gynecologic Laparoscopists*, 11(1) :73–78, 2004.
- J. Portilla and E. P. Simoncelli. A Parametric Texture Model based on Joint Statistics of Complex Wavelet Coefficients. *International Journal of Computer Vision*, oct 2000.
- W. K. Pratt. *Digital Image Processing : PIKS Inside, Third Edition*, pages 509–550. John Wiley & Sons, 2001. ISBN 0471374075.
- M. Pressigout and E. Marchand. Hybrid tracking algorithms for planar and non-planar structures subject to illumination changes. In *ACM/IEEE International Symposium on Mixed and Augmented Reality, (ISMAR'06)*, pages 52–55, Santa Barbara, CA, October 2006.
- G. Rellier. *Analyse de textures dans l'espace hyperspectral par des méthodes probabilistes*. PhD thesis, ESSI, Université de Nice Sophia Antipolis, Nov. 2002.
- H. Rininsland. ARTEMIS. A telemanipulator for Cardiac Surgery. *European Journal of Cardio-Thoracic Surgery*, 16 :106–111, 1999.
- C. Rivière, W. Ang, and P. Khosla. Toward active tremor cancelling in handheld microsurgical instruments. *IEEE Transactions on Robotics and Automation, special issue on Medical Robotics*, 19(5) :793–800, 2003.
- D. Salle. Modélisation et simulation d'un système de chirurgie robotisée pour l'analyse du retour haptique. In *8th International Symposium on Experimental Robotics (ISER)*, jun 2005.
- M. Sauvée. *Contribution à l'aide aux gestes pour la chirurgie cardiaque à coeur battant. Guidage échographique par asservissement prédictif non linéaire*. PhD thesis, Université Montpellier II, 2006.
- M. Sauvée, P. Renaud, P. Poignet, J. Triboulet, E. Dombre, M. Karouia, N. Bonnet, and A. Noce. Mitral valve leaflet motion tracking in ultrasound images. In *Surgetica : Computer Assisted Medical Interventions : Tools and Applications (Surgetica'05)*, Chambéry, France, jan 2005.

- P. Scheunders, S. Livens, G. Van de Wouwer, P. Vautrot, and D. Van Dyck. Wavelet-based Texture Analysis. *International Journal on Computer Science and Information Management, (IJCSIM)*, 1(2) :22–34, 1998. URL <http://wcc.ruca.ua.ac.be/~visielab/papers/ijcsim.ps.gz>.
- M. Shoham, M. Burman, E. Zehavi, L. Joskowicz, E. Batkin, and Y. Kunicer. Bone-mounted miniature robot for surgical procedures : concept and clinical applications. *IEEE Transactions on Robotics and Automation*, 19(5) :893–901, 2003.
- S. N. Sinha, J.-M. Frahm, M. Pollefeys, and Y. Genc. GPU-based Video Feature Tracking And Matching, Technical Report TR 06-012. Technical report, Department of Computer Science, UNC Chapel Hill, may 2006.
- D. Stoianovici, K. Cleary, A. Patriciu, D. Mazilu, A. Stanimir, W. Craciunoiu, V. Watson, and L. Kavoussi. Acubot : A robot for radiological interventions. *IEEE Transactions on Robotics and Automation*, 19(5) :927–930, 2003.
- E. Taillant, J.-C. Ailla-Vilchis, I. Bricault, and P. Cinquin. Implementation and Performances of the LPR light puncture robot. In *Proceedings of Surgetica, Computer Aided Medical Interventions : Tools and Applications*, pages 165–172, 2005.
- R. H. Taylor, H. A. Paul, D. D. Mittelstad, B. Glassman, M. B., and W. Bargar. Robotic total hip replacement surgery in dogs. In *IEEE Engineering in Medicine and Biology Society, 11th Annual International Conference*, 1989.
- M. Tendera. How much does Europe invest in the treatment of cardiovascular diseases? *European Heart Journal*, 27(13) :1521–1522, June 2006.
- H. Tran and E. Marchand. Real-time keypoints matching : application to visual servoing. In *Proceedings of IEEE International Conference on robotics and automation (ICRA'07)*, pages 3787–3792, Rome, Italy, Apr. 2007.
- R. Y. Tsai. An Efficient and Accurate Camera Calibration Technique for 3D Machine Vision. In *Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition (CVPR'86)*, pages 364–374, Miami Beach, FL, June 1986.
- F. Van Meer. *Conception et Réalisation d'une Instrumentation Terminale Intégrée en Chirurgie Mini Invasive Robotisée*. PhD thesis, LAAS, jan 2005.
- A. Vilchis, P. Cinquin, J. Troccaz, A. Guerraz, B. Hennion, F. Pellissier, P. Thorel, F. Courrèges, A. Goudron, G. Poisson, P. Vieyeres, P. Caron, O. Mérieux, L. Urbain, C. Daimo, S. Lavallée, P. Arbeille, M. Althuser, J.-M. Ayoubi, B. Tondu, and

- S. Ippolito. A system for robotic tele-echography. In *Proceedings of Medical Image Computing and Computer Assisted Intervention (MICCAI'01)*, pages 326–334, Utrecht, Pays Bas, 2001.
- M.-A. Vitrani. *Asservissement visuel à partir d'images échographiques. Application à la chirurgie intracardiaque*. PhD thesis, Université Pierre et Marie Curie, Dec. 2006.
- C.-H. Wei, C.-T. Li, and R. Wilson. *Database Modeling for Industrial Data Management : Emerging Technologies and Applications*, chapter A Content-based Approach to Medical Image Database Retrieval, pages 258–291. Idea Group Publishing, 2006.
- G. Winkler. *Image Analysis, Random Fields, and Dynamic Monte Carlo Methods : A Mathematical Introduction*. Application of Mathematics. Springer, 1991.
- World Health Organisation. Avoiding Heart Attacks and Strokes : Don't be a victim – Protect yourself. Technical report, WHO Library, 2005a. URL http://www.who.int/cardiovascular_diseases/resources/avoid_heart_attack_report/en/index.html.
- World Health Organisation. Preventing Chronic Diseases : A Vital Investment. Technical report, WHO Library, 2005b. URL http://www.who.int/chp/chronic_disease_report/full_report.pdf.
- Z. Zhang. A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11) :1330–1334, 2000.

Chapitre 5

Annexe : taxinomie des attributs texturaux

Introduction

Cette annexe présente les différents attributs considérés dans notre étude, donne leur expression et présente leurs principales caractéristiques. Au delà de la dimension “inventaire”, cette section présente la logique et les aspects calculatoires de ces méthodes.

Au cours des expérimentations menées afin de déterminer les meilleurs attributs (§2.3), nous avons comparés tous les attributs suivants entre eux, et même plusieurs versions de ces attributs :

- pour les matrices de co-occurrence, nous avons évalué les quatre directions principales ($\theta = 0^\circ, 45^\circ, 90^\circ, -45^\circ$), des pas allant de $d = 1$ à $d = 3$, et un nombre de niveau de gris considérés allant de 32 à 256 ;
- pour les matrices de longueur des plages, nous avons également évalué quatres directions ($\theta = 0^\circ, 45^\circ, 90^\circ, -45^\circ$) ;
- pour la méthode des différences de niveaux de gris, nous avons également considéré des déplacements suivant les vecteurs $(0, 1)$, $(1, 0)$, $(-1, 0)$, $(0, -1)$, $(0, 2)$, $(2, 0)$, $(-2, 0)$ et $(0, -2)$.

Nous avons également, dans le cas des matrices de cooccurrence et de niveaux de gris, calculé des matrices “isotropes” en faisant la moyennes des matrices obtenues selon deux ou quatre directions. Cette approche offre généralement de très bons résultats de classification, même si elle consomme plus de ressources.

Attributs texturaux du premier ordre

Les attributs statistiques du premier ordre donnent des relations statistiques caractérisant l'histogramme des niveaux de gris de la région d'intérêt considérée. Cet histogramme représente pour chaque valeur possible de luminance la proportion de pixels de l'image correspondants. Il permet d'étudier la répartition de l'information dans l'image, et même d'y apporter des traitements tels que : l'augmentation des nuances par égalisation de l'histogramme ou étirement d'histogramme ou des opérations de seuillage ou de segmentation simples.

De par leur méthode de calcul, ces attributs ne prennent pas en compte la position relative des pixels, et ne peuvent donc extraire qu'une information globale sur l'image. Cependant, ils fournissent des informations parfois très utiles (le PSNR par exemple, qui renseigne sur la quantité d'information contenue dans l'image). Ces attributs sont :

Moyenne : cet attribut représente la "luminosité" de l'image :

$$\mu = \frac{1}{N} \sum_{i=0}^{N-1} x_i \quad (5.1)$$

avec N le nombre total de pixels sur l'image et x_i l'intensité lumineuse du $i^{\text{ème}}$ pixel de l'image.

Variance : la variance (ou l'écart-type) de l'histogramme renseigne sur le niveau de contraste de l'image.

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2 \quad (5.2)$$

Rapport Signal sur Bruit (PSNR) : c'est le rapport entre la moyenne et l'écart type :

$$SNR = \frac{\mu}{\sigma} \quad (5.3)$$

Il donne une bonne idée de la dispersion du spectre de l'image et donne une estimation sur le niveau de bruit dans l'image.

Symétrie (Skewness) : la symétrie donne une information sur la déviation de la distribution par rapport à la valeur centrale de l'histogramme :

$$\gamma_1 = \sum_{i=0}^{N-1} \frac{(x_i - \mu)^3}{N\sigma^3}. \quad (5.4)$$

Un skewness positif indique une image globalement plus claire que la moyenne, un skewness négatif une image globalement plus foncée.

Aplatissement (Kurtosis) : l’aplatissement donne une information sur la répartition des niveaux de gris au sommet de l’histogramme :

$$\gamma_2 = \sum_{i=0}^{N-1} \frac{(x_i - \mu)^4}{N\sigma^4}. \quad (5.5)$$

Un kurtosis élevé implique un histogramme plus “pointu”, et donc une dominance de certaines valeurs des niveaux de gris.

Les matrices de cooccurrence

Présentation

Les matrices de cooccurrence (souvent notées GLCM pour “Gray Level Cooccurrence Matrix”) permettent de calculer des attributs texturaux d’ordre 2, c’est-à-dire qui prennent en compte des relations sur des couples de pixels. Introduite par Julez puis développée par Haralick dès 1973, cette approche est particulièrement féconde en applications et se trouve toujours en bonne place dans les ouvrages spécialisés Pratt [2001], Jähne [2002], Cocquerez et al. [1995]. Une matrice de cooccurrence représente l’histogramme des couples de pixels selon une direction θ et une distance d données. Par construction, la matrice de cooccurrence a une taille $N_h \times N_h$, N_h étant le nombre de valeurs que prennent les niveaux de gris. Pour les images que nous considérons, il faudra donc calculer des matrices 256×256 car nous travaillons avec des images en 256 niveaux de gris. De plus il faut recalculer la matrice de cooccurrence pour chaque couple (θ, d) , ce qui augmente généralement beaucoup les calculs. Cependant il est envisageable de réduire le nombre de calculs nécessaires par une série de transformations sur la matrice de cooccurrence :

Normalisation

La normalisation de la matrice de cooccurrence :

$$C_n(\theta, d) = p(\theta, d) = \frac{C(\theta, d)}{N_p}$$

permet de représenter l’information sous forme probabiliste : chaque valeur représente alors la probabilité de rencontrer un certain couple de pixels dans l’image.

Symétrisation

Elle est utilisée pour simplifier les calculs, sur la base de la relation $C(-\theta, d) = C^T(\theta, d)$:

$$C_s = \frac{1}{2}(C(\theta, d) + C(-\theta, d)) = \frac{1}{2}(C(\theta, d) + C^T(\theta, d)).$$

Dans notre application nous calculons directement la matrice symétrique.

Réduction du nombre de niveaux de gris

Pour réduire le nombre de calculs engendrés par les matrices de cooccurrence, il est possible de sous échantillonner l'image pour obtenir un nombre de niveaux de gris plus restreint. Cette approche réduit énormément les calculs, au prix d'une perte de l'information contenue dans l'image.

Attributs texturaux

Tous les attributs présentés ici sont calculés sur la matrice de cooccurrence *normalisée*¹.

Moyenne de la GLCM :

$$\mu_2 = \sum_{i,j} i \cdot p(i, j) \quad (5.6)$$

Variance de la GLCM :

$$\sum_{i,j} (i - \mu)^2 p(i, j) \quad (5.7)$$

Énergie : cet attribut caractérise l'homogénéité de l'image : plus il y a de régions homogènes sur l'image, plus l'énergie sera grande. Pour une image constante, l'énergie vaut 1.

$$f_1 = \sum_{i,j} p(i, j)^2 \quad (5.8)$$

Contraste : il permet également de caractériser l'homogénéité de l'image : plus il y a de régions homogènes et plus le contraste sera faible

$$f_2 = \sum_{i,j} (i - j)^2 p(i, j) \quad (5.9)$$

¹nous calculons toujours les matrices de cooccurrence selon plusieurs directions simultanément afin d'inclure un comportement isotrope. Dans les formules énoncées, i et j désignent les indices de ligne et de colonne et $p(i, j)$ la valeur de la matrice de cooccurrence normalisée associée.

Entropie : l'entropie donne une information sur la complexité de l'image. On l'utilise pour différencier les textures "lisses" des textures "rugueuses" :

$$f_9 = - \sum_{i,j} p(i,j) \ln(p(i,j)) \quad (5.10)$$

Corrélation : cet attribut représente la corrélation croisée entre les séries des lignes et colonnes de la matrice de cooccurrence :

$$\rho = \sum_{i,j} \frac{(i - \mu_i)(j - \mu_j)}{\sigma_i \sigma_j} p(i,j) \quad (5.11)$$

avec μ_i et μ_j les moyennes des lignes et des colonnes respectivement, σ_i et σ_j étant les variances des lignes et la colonnes. Ces paramètres se calculent comme suit :

$$\left\{ \begin{array}{ll} p'(i) = \sum_{j=0}^N p(i,j) & p''(j) = \sum_{i=0}^N p(i,j) \\ \mu_i = \sum_{i=1}^N i p'(i) & \mu_j = \sum_{j=1}^N j p''(j) \\ \text{puis :} & \\ \sigma_i = \sqrt{\sum_{i=0}^N (i - \mu_i)^2 p'(i)} & \sigma_j = \sqrt{\sum_{j=0}^N (j - \mu_j)^2 p''(j)} \end{array} \right.$$

Les définitions (et les valeurs) de μ_i , μ_j , σ_i et σ_j restent les mêmes pour tous les attributs qui les utilisent.

Homogénéité ou **moment différentiel inverse**. Cet attribut se comporte de façon inverse au contraste : plus il y a de régions homogènes et plus l'homogénéité est élevée

$$H = \sum_{i,j} \frac{1}{1 + (i - j)^2} p(i,j) \quad (5.12)$$

Moment diagonal : cet attribut renseigne sur la présence de régions "inhomogènes" dans l'image :

$$m = \sum_{i,j} \sqrt{\frac{1}{2} |i - j|} p(i,j) \quad (5.13)$$

Cluster shade et **cluster prominence** : le cluster shade (5.14) ou *groupe de nuance* en français, et le cluster prominence (5.15) ou *groupe de proéminence* sont des moments d'ordre 3 et 4 sur la matrice de cooccurrence. Il permettent de pousser l'étude de la répartition des niveaux de gris autour de la moyenne :

$$c_s = \sum_{i,j} (i + j - 2\mu)^3 p(i,j) \quad (5.14)$$

$$c_p = \sum_{i,j} (i + j - 2\mu)^4 p(i,j) \quad (5.15)$$

Autres attributs : Parmi les 13 attributs texturaux² introduits par Haralick dans son article fondateur il reste à considérer :

$$f2 = \sum_{k=0}^{Ng-1} k^2 p_{x-y}(k) \quad (5.16)$$

$$f6 = \sum_{i=2}^{2Ng} i p_{x+y}(i) \quad (5.17)$$

$$f7 = \sum_{i=2}^{2Ng} (i - f_8)^2 p_{x+y}(i) \quad (5.18)$$

$$f8 = - \sum_{i,j} p_{x+y}(i) \log(p_{x+y}(i)) \quad (5.19)$$

$$f11 = \sum_{i=0}^{Ng-1} p_{x-y}(i) \log(p_{x-y}(i)) \quad (5.20)$$

$$f12 = \frac{f_9 - HXY1}{\max(HXY1, HXY2)} \quad (5.21)$$

$$f13 = \sqrt{1 - e^{-2(HXY2 - f_9)}} \quad (5.22)$$

avec les notations :

$$\left\{ \begin{array}{l} p_x(i) = \sum_{j=1}^{Ng} p(i, j) \\ p_y(j) = \sum_{i=1}^{Ng} p(i, j) \\ p_{x+y}(k) = \sum_{i=1}^{Ng} \sum_{j=1}^{Ng} p(i, j) \text{ avec } i + j = k, k = 2, 3, \dots, Ng \\ p_{x-y}(k) = \sum_{i=1}^{Ng} \sum_{j=1}^{Ng} p(i, j) \text{ avec } |i - j| = k, k = 0, 1, \dots, Ng - 1 \\ HXY1 = - \sum_{i=1}^{Ng} \sum_{j=1}^{Ng} p(i, j) \log(p_x(i)p_y(j)) \\ HXY2 = - \sum_{i=1}^{Ng} \sum_{j=1}^{Ng} p_x(i)p_y(j) \log(p_x(i)p_y(j)) \end{array} \right.$$

Ces attributs sont moins utilisés dans la littérature que ceux précédemment introduits, cependant nous avons choisi des les inclure dans notre étude afin de comparer le jeu d'attributs le plus riche possible.

Méthode des différences de niveaux de gris

Cette méthode, décrite dans Herlidou [1999], consiste à calculer des statistiques sur une image constituée de l'image initiale moins l'image décalée. En effet, devant la complexité de certains attributs de la matrice de cooccurrence, la méthode des différences de niveaux de gris permet d'évaluer certaines relations "inter-pixels" (on a donc des attributs d'ordre 2) sur l'image. La construction de la matrice de différences est simple : on part d'une

²nous avons omis de la liste ceux déjà cités.

image digitalisée I , puis on lui soustrait pixel par pixel une image translatée d'un vecteur $d = (\Delta_i, \Delta_j)$. C'est ce qui est exprimé de manière plus formelle dans l'équation 5.23 :

$$d = (\Delta_i, \Delta_j)$$

$$g_d = |I(i, j) - I(i + \Delta_i, j + \Delta_j)| \quad (5.23)$$

$$\text{puis : } p_d(k) = P(g_d(i, j) = k) \quad (5.24)$$

On calcule alors les probabilités $p_d(i)$ associées aux valeurs possibles de g_d . Ainsi on calcule l'histogramme de l'image issue de la différence, sur lequel on calculera nos attributs. Sur cette base on peut construire 5 attributs texturaux :

Moyenne

La moyenne se calcule selon l'équation 5.25, de la même façon que pour la matrice de cooccurrence.

$$\mu = \sum_{i=1}^{N_g} i \cdot p_d(i) \quad (5.25)$$

Contraste

Le contraste se calcule comme moment d'ordre 2 sur l'image, comme indiqué par l'équation 5.26. Il permet comme son nom l'indique déterminer l'importance du contraste de l'image.

$$CONT = \sum_{i=1}^{N_g} i^2 \cdot p_d(i) \quad (5.26)$$

Moment angulaire d'ordre 2 ou uniformité

Ce critère, donné par la formule 5.27, est élevé si les variations de la texture sont concentrées autour de certaines valeurs.

$$ASM = \sum_{i=1}^{N_g} [p_d(i)]^2 \quad (5.27)$$

Entropie

Comme dans le cas de l'uniformité, ce paramètre donné par l'équation 5.28 informe sur l'homogénéité de la répartition des valeurs de l'image différence.

$$\rho = - \sum_{i=1}^{N_g} p_d(i) \log(p_d(i)) \quad (5.28)$$

Moment différentiel inverse

Ce paramètre, dont la formule est présentée par l'équation 5.29, est le cousin du paramètre d'homogénéité défini pour la matrice de cooccurrence. Il favorise les faibles valeurs de i , qui représentent de faibles variations de la texture sur l'image (c'est l'approche contraire du contraste défini plus haut).

$$\rho = - \sum_{i=1}^{N_g} \frac{p_d(i)}{i^2 + 1} \quad (5.29)$$

Méthode des gradients

Cette méthode Herlidou [1999] consiste en une étude des gradients de l'image calculés sur des fenêtres 3×3 ou 5×5 . Les Équations 5.30 et 5.31 présentent le calculs de ces gradients.

$$gd_1(i, j) = \sqrt{[g(i+1, j) - g(i-1, j)]^2 + [g(i, j+1) - g(i, j-1)]^2} \quad (5.30)$$

$$gd_2(i, j) = \sqrt{[g(i+2, j) - g(i-2, j)]^2 + [g(i, j+2) - g(i, j-2)]^2} \quad (5.31)$$

avec $g(i, j)$ la valeur de luminance du pixel à la position (i, j) .

On peut alors calculer deux attributs sur ces gradients :

Moyenne des gradients :

$$\mu = \frac{1}{N} \sum_{i,j} gd(i, j) \quad (5.32)$$

Variance des gradients :

$$\sigma = \frac{1}{N} \sum_{i,j} (gd(i, j) - \mu)^2 \quad (5.33)$$

Matrice des longueurs de plages des niveaux de gris

Cette méthode consiste à compter le nombre de plages de même luminance dans une image. La méthode des longueurs de plages consiste à compter le nombre de plages dans une direction donnée et à utiliser des statistiques sur ces plages pour construire des attributs. Ces plages sont déterminées suivant une direction³ θ . À partir de cette matrice, on détermine une série d'attributs texturaux⁴ :

³notation introduite au paragraphe 2.2.1.3.

⁴Dans les formules énoncées, i et j désignent les indices de ligne et de colonne et $r(i, j)$ la valeur de la matrice de longueur des plages associée. N_r et N_p désignent respectivement le nombre de plages et le nombre de pixels dans l'image.

Poids des plages courtes :

$$SRE = \frac{1}{Nr} \sum_{i,j} \frac{r(i,j)}{j^2} \quad (5.34)$$

Poids des plages longues :

$$LRE = \frac{1}{Nr} \sum_{i,j} j^2 \cdot r(i,j) \quad (5.35)$$

Distribution des niveaux de gris :

$$GDIST = \frac{1}{Nr} \sum_i [\sum_j r(i,j)]^2 \quad (5.36)$$

Distribution des longueurs de plages :

$$RLDIST = \frac{1}{Nr} \sum_j [\sum_i r(i,j)]^2 \quad (5.37)$$

Pourcentage de plages :

$$RPC = Nr/Np \quad (5.38)$$

Poids des longueurs de plages de faible niveau de gris :

$$LGRE = \frac{1}{Nr} \sum_{i,j} \frac{r(i,j)}{i^2} \quad (5.39)$$

Poids des longueurs de plages de fort niveau de gris :

$$HGRE = \frac{1}{Nr} \sum_{i,j} i^2 \cdot r(i,j) \quad (5.40)$$

Poids des petites plages de faible niveau de gris :

$$LGSRE = \frac{1}{Nr} \sum_{i,j} \frac{r(i,j)}{i^2 \cdot j^2} \quad (5.41)$$

Poids des petites plages de fort niveau de gris :

$$HGSRE = \frac{1}{Nr} \sum_{i,j} \frac{i^2 \cdot r(i,j)}{j^2} \quad (5.42)$$

Poids des longues plages de faible niveau de gris :

$$LGHRE = \frac{1}{Nr} \sum_{i,j} \frac{j^2 \cdot r(i,j)}{i^2} \quad (5.43)$$

Poids des longues plages de fort niveau de gris :

$$LRE = \frac{1}{Nr} \sum_{i,j} i^2 j^2 r(i,j) \quad (5.44)$$

Attributs basés sur la transformée de Fourier

La transformée de Fourier permet de capturer facilement les régularités du signal, et par sa capacité à diagonaliser les opérateurs linéaires [Mallat, 1999], présente un cadre naturel pour la décomposition et la synthèse des signaux. Du point de vue de l'analyse de texture, elle peut être utilisée pour décomposer la texture ou pour caractériser la textures par des attributs.

Dans notre étude nous avons utilisé une FFT bidimensionnelle et les attributs issus de [Dulyakarn et al., 2000] sur le spectre de puissance. En notant F la transformée de Fourier :

$$F(U, V) = \sum_{x=0}^{N_1-1} \sum_{y=0}^{N_2-1} e^{-i2\pi V \frac{y}{N_2}} e^{-i2\pi \frac{x}{N_1}} f(x, y) \quad (5.45)$$

nos attributs sont définis ainsi :

Maximum d'amplitude :

$$\max\{|F(i, j)|, (i, j) \neq (0, 0)\} \quad (5.46)$$

Amplitude moyenne :

$$\frac{1}{N} \sum_{i,j} |F(i, j)| \quad (5.47)$$

Energie :

$$\sum_{i,j} |F(i, j)|^2 \quad (5.48)$$

variance du spectre d'amplitude :

$$\frac{1}{N} \sum_{i,j} [|F(i, j)| - A_m]^2 Energy \quad (5.49)$$