

Numéro d'ordre : 66-2004

Année 2004

N° attribué par la bibliothèque

/ \_ / \_ / \_ / \_ / \_ / \_ / \_ / \_ / \_ /

# THÈSE

*présentée*

devant l'UNIVERSITÉ CLAUDE BERNARD - LYON 1

*pour l'obtention*

## du DIPLOME DE DOCTORAT

*arrêté du 25 avril 2002*

Discipline : Ingénierie Médicale et Biologique

présentée et soutenue publiquement le 1<sup>er</sup> juin 2004

par

**Nicoleta M. PAUNA**

Maître es Sciences Physiques

Titre :

**Evaluation des méthodes de mise en correspondance en  
imagerie multimodale IRM/TEP thoracique et  
cardiaque**

### Jury

---

M. Marc JANIER	Co-Directeur
M. Cozar ONUC	Co-Directeur
M. Patrick CLARYSSE	Co-Encadrant
M. Bernard GIBAUD	Rapporteur
M. Radu CIUPA	Rapporteur
Mme. Adelina GEORGESCU	Examineur
Mme. Isabelle MAGNIN	Présidente

Centre de Recherche et d'Applications en Traitement de l'Image et du Signal

CREATIS - CNRS UMR 5515 - INSERM U630



# Avant-propos

Ce travail de thèse s'est déroulé au Centre de Recherche et d'Applications en Traitement de l'Image et du Signal (CREATIS), Unité UMR CNRS 5515 et INSERM U630. Cette thèse a fait l'objet d'une co-tutelle entre l'Université Claude Bernard Lyon 1 et l'Université Babes Bolyai, Cluj Napoca, Roumanie.

Tout d'abord, je tiens à exprimer ma reconnaissance à Madame Isabelle Magnin, DR INSERM et Monsieur Didier Revel, PUPH, co-directeurs de CREATIS, qui m'ont accueilli au sein du laboratoire en me permettant ainsi de travailler dans un environnement scientifique de haut niveau et d'avoir une collaboration très enrichissante avec des médecins et des ingénieurs du domaine de l'imagerie médicale.

Je remercie mes co-directeurs de thèse, les professeurs Marc Janier et Onuc Cozar qui m'ont orientée vers un sujet de recherche à la fois riche et stimulant. Merci Marc de m'avoir fait confiance en acceptant la co-tutelle !

Je suis extrêmement reconnaissante à Monsieur Bernard Gibaud, CR INSERM et au Professeur Radu Ciupa d'avoir accepté d'être rapporteurs de ma thèse. Leurs remarques constructives m'ont été d'une aide précieuse. Je remercie également le professeur Adelina Georgescu et la présidente de mon jury, Madame Isabelle Magnin.

J'exprime toute ma gratitude à Patrick Clarysse, CR CNRS, qui a co-encadré cette thèse et qui m'a dirigée et conseillée avec beaucoup de patience tout au long de cette période. Je le remercie de m'avoir généreusement fait part de sa grande expérience en imagerie médicale. Ce manuscrit de thèse sous sa forme actuelle doit beaucoup à ses relectures minutieuses et remarques judicieuses. J'ajoute aussi son aide décisif pour constituer les nombreux dossiers de demande de financement pendant la thèse !

Un grand merci à Nicolas Costes, ingénieur au CERMEP, pour le temps et l'énergie qu'il a consacré à ce projet et de m'avoir fait part de ses compétences en imagerie TEP et dans les problématiques de recalage/validation, ainsi que pour ses qualités humaines. Ses encouragements réguliers, sa simplicité et sa rigueur m'ont été précieuses. Qu'il soit également remercié pour la relecture attentive de ce manuscrit.

Le travail de thèse m'a fait découvrir le travail en équipe, avec un échange constant avec d'autres chercheurs. Les discussions menées avec Quoc Cuong Pham, Thomas Grenier (CREATIS) et Timo Mäkelä du Laboratory of Biomedical Engineering de l'Université Technologique d'Helsinki m'ont fait beaucoup avancer dans ma recherche.

Merci à Johan Montagnat et Luc Bracoud à qui je dois mes débuts en programmation en langage C ! Je remercie Dr. Pierre Crosille pour les données IRM, et Nicolas Costes et Franck Lavenne pour celles en TEP.

Je ne voudrais surtout pas oublier mes amis avec qui j'ai partagé les joies et les soucis : Marcela, Laurence et Fabrice. Et puis les amis de CREATIS : Luc, Leonardo, Julien, Jérémie, Sabin, Rana, Marlène, Alejandro, JPR. Je remercie Elisabeth et Marie Ange pour leurs conseils pour mes dossiers de demande de bourses et les démarches de post-doctorat !

Mille mercis à "ma deuxième famille en France", Geneviève et Bernard Longère, pour leur soutien et gentillesse !

Je tiens à remercier les membres des deux Clubs Soroptimistes de Lyon, et, en particulier, Janine Sophie Guimiot et Evelyne Blum (présidentes des Clubs Soroptimistes lors de mon arrivée à Lyon) qui m'ont apporté l'aide financier des leurs clubs et qui ont toujours été à mon écoute pendant cette période. Grâce à eux, j'ai rencontré Stéphanie, Françoise et Martine !

Je remercie les initiateurs de la convention de collaboration entre les universités des deux pays, tout particulièrement à messieurs O. Cozar et T. Tauth.

Mes autres remerciements sincères sont destinés aux divers organismes qui m'ont apporté leur soutien financier : l'Université Babes Bolyai de Roumanie, la Région Rhône-Alpes, l'Union Soroptimiste d'Europe, la Fondation pour la Recherche Médicale.

*Pour leurs efforts, amour et confiance en moi, je dédie cette thèse à ma mère et à mon père, à ma soeur Serenella (et au petit Raul!), et à ma grand-mère.*

*J'aimerais remercier mes parents de tout mon coeur de m'avoir soutenue pour continuer mes études en France, ce qui m'a permis de découvrir de nouveaux horizons ! Merci maman de m'avoir guidée vers le domaine de l'imagerie médicale. L'aboutissement de cette thèse doit beaucoup à tes encouragements, à ton enthousiasme et à ta tendresse, qui m'ont donné la force d'avancer même dans les moments les plus difficiles !*



# Table des matières

<b>Introduction</b>	<b>17</b>
<b>1 Contexte</b>	<b>19</b>
1.1 Fusion d'informations multimodales pour l'étude de la viabilité myocardique	19
1.1.1 Contexte médical . . . . .	19
1.1.2 Techniques d'exploration du coeur . . . . .	23
1.1.3 Mise en correspondance d'images dans le contexte cardiaque . . . .	24
1.2 L'importance de l'évaluation . . . . .	33
<b>2 Évaluation des méthodes de traitement d'images médicales</b>	<b>35</b>
2.1 Cadre conceptuel de l'évaluation . . . . .	36
2.2 Procédures d'évaluation . . . . .	38
2.2.1 Standardisation des procédures d'évaluation . . . . .	38
2.2.2 Données d'évaluation . . . . .	41
2.2.3 Métriques d'évaluation . . . . .	44
2.3 Évaluation des méthodes de recalage . . . . .	46
2.3.1 Critères de qualité du recalage . . . . .	46
2.3.2 Évaluation de la précision des méthodes de recalage . . . . .	47
2.3.3 Stratégies d'évaluation des méthodes de recalage d'images cardiaques et thoraciques . . . . .	49
<b>3 Stratégie d'évaluation de la précision des méthodes de mise en corres- pondance en imagerie thoracique et cardiaque TEP et IRM</b>	<b>57</b>
3.1 Présentation de la stratégie . . . . .	58
3.2 Construction de la référence . . . . .	58
3.2.1 Données IRM de référence . . . . .	59
3.2.2 Modèle numérique 3D (anatomique et métabolique) . . . . .	60
3.2.3 Simulation des données TEP . . . . .	60
3.3 Évaluation des performances des méthodes de mise en correspondance . . .	66
3.3.1 Transformations de référence . . . . .	66
3.3.2 Métriques d'évaluation de la précision . . . . .	66
3.4 Hypothèse d'évaluation . . . . .	68
3.4.1 Niveau d'évaluation . . . . .	68

3.4.2	Objectif de l'évaluation . . . . .	68
3.4.3	Test d'hypothèses . . . . .	68
<b>4</b>	<b>Evaluation comparative de méthodes de mise en correspondance</b>	<b>70</b>
4.1	Présentation des méthodes de mise en correspondance évaluées . . . . .	70
4.1.1	Les méthodes de recalage iconiques . . . . .	70
4.1.2	Méthode de recalage basée sur des surfaces . . . . .	74
4.2	Évaluation de la précision des méthodes . . . . .	75
4.2.1	Mise en correspondance d'images de thorax . . . . .	75
4.2.2	Mise en correspondance d'images de coeur . . . . .	80
4.3	Discussion . . . . .	87
4.4	Limites et perspectives . . . . .	89
	<b>Conclusion</b>	<b>91</b>
	<b>Annexes</b>	<b>93</b>
<b>A</b>	<b>Modalités d'acquisition d'images médicales</b>	<b>93</b>
A.1	Tomographie par Émission de Positons . . . . .	93
A.1.1	Émission de positons . . . . .	93
A.1.2	Annihilation de positons . . . . .	94
A.1.3	Détection en coïncidence . . . . .	94
A.1.4	Projections . . . . .	95
A.1.5	Facteurs de dégradation des images . . . . .	96
A.2	Imagerie par Résonance Magnétique . . . . .	97
A.2.1	Principe physique de la résonance magnétique nucléaire . . . . .	98
A.3	Les ultrasons . . . . .	101
A.4	Les rayons X . . . . .	102
<b>B</b>	<b>Mesures de Similarité</b>	<b>105</b>
<b>C</b>	<b>Matrices des transformations rigides 3D</b>	<b>109</b>
C.1	Angles d'Euler . . . . .	109
C.2	Représentation des quaternions . . . . .	110
<b>D</b>	<b>Test d'hypothèses</b>	<b>112</b>
D.1	Test d'hypothèses unilatéral à gauche pour une moyenne . . . . .	112
	<b>Bibliographie</b>	<b>114</b>



# Table des figures

1.1	Anatomie du cœur . . . . .	20
1.2	Représentation des ventricules gauche et droit du cœur. . . . .	20
1.3	Electrocardiogramme et cycle cardiaque. . . . .	22
1.4	Modalités impliquées dans l'estimation de la viabilité. . . . .	24
1.5	Types de transformations géométriques d'après [Maintz 98] . . . . .	27
2.1	Performance d'un algorithme en fonction de sa complexité théorique (d'après [Bowyer 98]) . . . . .	37
2.2	Le modèle générique d'évaluation GEM (d'après [Buvat 99]) . . . . .	39
2.3	Étude de consistance (d'après [Jenkinson 01]) . . . . .	40
3.1	Stratégie d'évaluation : vue générale . . . . .	58
3.2	Les données de référence IRM/TEP : (a) image par RM réelle, (b) image par RM sans les bras, (c) image par RM segmentée et étiquetée, (d) image simulée de TEP de transmission, (e) image simulée de TEP d'émission, (f) image simulée du cœur en émission . . . . .	59
3.3	Image de coeur TEP d'émission (a) réelle, (b) simulée . . . . .	63
3.4	Image 3D de thorax TEP de transmission simulée . . . . .	63
3.5	Image de thorax TEP de transmission (a) réelle, (b) simulée . . . . .	64
3.6	(a) Modèle anatomique pathologique du fantôme numérique, (b) structure des régions ischémiques du ventricule gauche . . . . .	65
3.7	Simulation d'images TEP de captation glucidique (a) normale et (b) dans un cas ischémique . . . . .	65
3.8	Boîte à moustaches pour la représentation de la distribution des erreurs . .	67
4.1	Principe de la méthode de recalage rigide HMR . . . . .	71
4.2	Principe de la méthode de recalage SBR basée sur les surfaces (issu de [Mäkelä 03a]) . . . . .	76
4.3	Recalage d'images transverses par RM et TEP de transmission de thorax : (a) image par RM, (b) image TEP transformée ; visualisation en damier des volumes avant le recalage (c) et après le recalage avec les méthodes IMP (d) et HMR+CC (e) . . . . .	78

4.4	Erreurs de recalage sur les paramètres de la transformation pour les images par RM et TEP de transmission de thorax avec la méthode SBR et les optimisations de Powell (a), grille (b) et Simplexe (c), avec la méthode HMR et les mesures de similarité IM (d), RC (e) et CC (f) et avec la méthode IMP (g). (h) erreurs $d_{RMS}$ pour chaque méthode. Les erreurs en translation sont en voxels (1 voxel=1.95 mm) et les erreurs en rotation sont en degrés. La taille du voxel en TEP ( $V_{TEP}$ ) et la résolution physique de la caméra TEP sont respectivement égales à 3.52 mm et 4.1 mm. . . . .	79
4.5	Image du coeur sain en vue petit-axe avec la région d'intérêt contenant le ventricule gauche (a) coupe IRM, (b) coupe TEP simulée correspondante .	81
4.6	Recalage d'images par RM et TEP d'émission de coeur : (a) image par RM, (b) image TEP transformée, visualisation en damier des volumes avant le recalage (c) et après le recalage avec les méthodes HMR+IM (d) et HMR+CC (e) . . . . .	82
4.7	Erreurs de recalage pour les images de coeur sain en utilisant la méthode de recalage HMR avec les mesures de similarité CC (a), RC (b), IM (c) et la méthode IMP (d). (h) erreurs $d_{RMS}$ pour chaque méthode. Les erreurs en translation sont en voxels (1 voxel=1 mm) et les erreurs en rotation sont en degrés. La taille du voxel en TEP ( $V_{TEP}$ ) et la résolution physique de la caméra TEP sont respectivement égales à 2.43 mm et 4.1 mm. . . . .	83
4.8	Erreurs de recalage pour les images de coeur pathologique en utilisant la méthode de recalage HMR avec les mesures de similarité CC (a), RC (b), IM (c) et la méthode IMP (d). (h) erreurs $d_{RMS}$ pour chaque méthode. Les erreurs en translation sont en voxels (1 voxel=1 mm) et les erreurs en rotation sont en degrés. La taille du voxel en TEP ( $V_{TEP}$ ) et la résolution physique de la caméra TEP sont respectivement égales à 2.43 mm et 4.1 mm.	84
A.1	(a) annihilation électron/positon, (b) détection en coïncidence des photons	95
A.2	Coïncidences fortuite et diffusée . . . . .	95
A.3	Principe du codage spatial dans une direction (a) application d'un champ magnétique uniforme, (b) application d'un gradient de champ magnétique (issu de [Kastler 97]) . . . . .	100
A.4	Trois images par résonance magnétique d'un même niveau de coupe pondérées en (a) T1, (b) T2, (c) densité protonique . . . . .	101
A.5	Système d'acquisition d'images ultrasonores . . . . .	102
A.6	Principe d'acquisition en tomodensitométrie . . . . .	103
A.7	Tomographie à rayons X : parcours de l'espace en rotation simple et avec une acquisition spiralée par scanneur hélicoïdal . . . . .	104
B.1	Notations pour un histogramme joint d'après [Sarrut 00] . . . . .	106

# Bibliographie personnelle

T. Mäkelä, P. Clarysse, O. Sipilä, N. Pauna, Q. C. Pham, T. Katila, I. E. Magnin, 2002, A review of cardiac image registration methods, IEEE Transactions on Medical Imaging, 21(9), pp 1011-1021.

M. Pollari, J. Lötjönen, T. Mäkelä, N. Pauna, A. Reilhac, P. Clarysse, Evaluation of cardiac PET-MRI registration methods using a numerical breathing phantom, IEEE International Symposium on Biomedical Imaging, 2004, pp 1447-1450.

N. Pauna, P. Croisille, A. Reilhac, T. J. Mäkelä, M. Janier, O. Cozar, P. Clarysse, A strategy to quantitatively evaluate MRI/PET cardiac rigid registration methods using a Monte Carlo simulator, in Lecture Notes in Computer Science 2674 : Second International Workshop on Functional Imaging and Modeling of the Heart, FIMH 2003, I.E. Magnin, J. Montagnat, P. Clarysse, J. Nenonen, T. Katila (Eds), pp 194-204.

T. J. Mäkelä, M. Pollari, N. Pauna, A. Reilhac, P. Clarysse, I. Magnin, T. Katila, Evaluation and comparaison of surface and intensity based registration methods for thorax and cardiac MR and PET images, in Lecture Notes in Computer Science 2674 : Second International Workshop on Functional Imaging and Modeling of the Heart, FIMH 2003, I.E. Magnin, J. Montagnat, P. Clarysse, J. Nenonen, T. Katila (Eds), pp 224-233.



---

# NOTATIONS UTILISÉES

## Techniques d'imagerie médicale

IRM	Imagerie par Résonance Magnétique
TEP	Tomographie par Emission de Positons
TR	Image TEP de Transmission
EM	Image TEP d'Emission
TEMP	Tomographie par Emission Mono-photonique
CT	Tomodensitométrie par rayons X
US	Ultrasons
MCG	Magnétocardiographie
ECG	Electrocardiographie

## Divers

VD	Ventricule Droit
VG	Ventricule Gauche
ECG	Electrocardiogramme
FDG	Fluorodésoxyglucose

## Fonctions de coût

IM	Information Mutuelle
IMN	Information Mutuelle Normalisée
SDA	Somme des Différences Absolues
SDC	Somme des Différences au Carré
MDC	Moyenne des Différences au Carré
CSS	Changement Stochastique de Signe
MV	Maximum de vraisemblance
CC	Coefficient de Corrélation
RC	Rapport de Corrélation
VRI	Variance du Rapport des Intensités

## Mesures statistiques

SDV	déviatiion standard
moy	moyenne
eqm	erreur quadratique moyenne



# Resumé

Ce travail porte sur l'évaluation de la précision des méthodes de mise en correspondance en imagerie multimodale IRM/TEP thoracique et cardiaque.

En cardiologie, les résultats les plus récents relatifs aux pathologies ischémiques montrent que la viabilité des tissus myocardiques peut être estimée à partir de l'analyse conjointe de la perfusion, du métabolisme oxydatif et de la fonction contractile, d'où l'intérêt d'une imagerie multimodale et intégrative. L'Imagerie par Résonance Magnétique (IRM) est la modalité d'imagerie non-invasive de choix pour accéder à l'anatomie cardiaque. L'information fonctionnelle liée au métabolisme du coeur peut être évaluée en utilisant la Tomographie d'Emission de Positons (TEP) de captation glucidique au fluorodesoxyglucose marqué. La mise en correspondance (ou recalage) des territoires analogues imagés dans différentes modalités est une étape préalable de la fusion des données anatomiques et fonctionnelles. Dans ce contexte, nous nous posons le problème de l'exploitation de programmes de recalage automatique d'images.

L'utilisation en routine clinique d'outils de traitement de l'information et des images n'est envisageable qu'après une évaluation rigoureuse de leurs performances. Une stratégie d'évaluation requiert une *référence absolue* (*gold standard* en anglais), qui représente la solution idéale. Elle peut être définie comme un jeu de données test parfaitement décrit associé à des métriques d'erreur qui permettront de quantifier de manière absolue les performances d'une méthode. En imagerie médicale, une telle référence n'existe généralement pas.

Dans cette thèse, nous proposons (chapitre 3) une stratégie pour l'évaluation rigoureuse de la précision des méthodes de mise en correspondance rigide d'images TEP et par RM thoraciques et cardiaques. Elle repose sur la simulation réaliste d'images TEP à partir d'images par RM réelles. Pour cela, nous construisons un modèle théorique de captation glucidique caractéristique du contexte clinique étudié (sujet sain et ischémique) à partir de mesures sur des examens réels. Les principaux processus physiques d'une acquisition TEP sont modélisés dans un simulateur de type Monte Carlo. Notre approche fournit ainsi des images TEP simulées qui sont en parfaite correspondance avec les images par RM réelles.

---

La stratégie est mise en oeuvre pour l'évaluation de trois méthodes de recalage rigide IRM/TEP (chapitre 4) ; deux d'entre elles sont des approches iconiques basées sur les intensités des images, tandis que la troisième est une approche surfacique. Une des méthodes iconiques a été développée dans le cadre de cette thèse. Un ensemble de transformations géométriques rigides de référence aléatoirement générées est appliqué sur l'image TEP afin d'obtenir une série de couples d'images IRM/TEP non-recalées. Les transformations estimées sur ces couples d'images peuvent être alors comparées aux transformations appliquées, et l'écart quantifié à l'aide de critères d'erreur. Des tests statistiques ont été utilisés afin de répondre à la question posée par l'objectif de validation dans le contexte clinique de l'étude de la viabilité myocardique. Dans le cas de l'imagerie thoracique, les méthodes iconiques ont conduit à des résultats de recalage plus précis que la méthode de recalage de surfaces. Dans le cas du recalage d'images de coeur, il est nécessaire de sélectionner une région d'intérêt et de pré-orienter les images afin d'obtenir un recalage satisfaisant avec la mesure d'information mutuelle et le rapport de corrélation.

Notre travail de recherche souligne la difficulté de concevoir une stratégie rigoureuse d'évaluation et le besoin de définir une terminologie et une méthodologie communes pour l'évaluation. La standardisation de protocoles d'évaluation et leur adoption par la communauté scientifique faciliteraient la comparaison des processus de traitements d'images médicales (e.g. recalage) ainsi que leur utilisation en clinique et leur éventuelle valorisation industrielle .



---

# Introduction

L’essor des techniques d’imagerie médicale permet d’obtenir des images du corps humain de plus en plus riches en information. Les modalités<sup>1</sup> d’imagerie médicale se sont multipliées depuis 30 ans sous la poussée d’innovations technologiques majeures. Les images acquises offrent de plus en plus souvent une vue de l’anatomie 3D du patient et non plus seulement une projection, comme en radiologie conventionnelle, ou une coupe 2D, comme en échographie. Leur résolution spatiale augmente également : la taille du voxel descend couramment en dessous du millimètre, voire du demi-millimètre<sup>2</sup>. La nature des informations peut être de nature physiologique ou anatomique.

La multiplication des techniques d’exploration a pour conséquence l’augmentation drastique de la quantité d’images à analyser. Face à cette croissance, l’intérêt de développer des outils de traitement des informations et des images facilitant l’analyse par les médecins s’impose comme une évidence. Les traitements d’images visent généralement à améliorer la qualité des images et extraire des paramètres quantitatifs relatifs aux formes des structures anatomiques, à leur mouvement, etc... La mise en correspondance d’images (ou recalage) est un exemple particulièrement important d’un tel traitement dont le développement a été principalement suscité en imagerie cérébrale anatomique et fonctionnelle.

En imagerie cardiaque, l’étude de la viabilité myocardique constitue un exemple d’analyse conjointe d’informations issues de différentes modalités. Dans ce contexte, la combinaison de données anatomiques et fonctionnelles (perfusion myocardique, fonction contractile ...) implique de trouver les moyens de mettre en correspondance des territoires analogues imagés dans différentes modalités. En routine clinique, cette mise en correspondance est généralement réalisée mentalement par les médecins.

Des méthodes de recalage sont développées dans le but d’offrir une meilleure précision, une reproductibilité de la procédure et un gain de temps. Plusieurs méthodes sont aujourd’hui disponibles et les échanges entre différents utilisateurs ou ingénieurs visent à améliorer leurs performances. Pour une utilisation en clinique, une étape de validation est incontournable afin que le médecin sache la précision qu’il peut espérer, les risques d’échec. La difficulté de cette tâche, souvent négligée, réside dans le fait que la *vérité terrain* (ou *gold standard* en anglais) est souvent inaccessible en imagerie médicale. Il est donc très difficile d’évaluer objectivement les performances d’une méthode et de la comparer à d’autres.

Notre sujet de thèse s’inscrit dans le cadre des travaux menés sur l’étude de la viabilité myocardique par l’équipe “*Imagerie fonctionnelle de l’ischémie myocardique*” du laboratoire CREATIS. Le problème que nous proposons de traiter dans cette thèse est celui de l’évaluation des méthodes de recalage en imagerie multimodale IRM/TEP cardiaque.

Nous présentons une stratégie pour réaliser l’évaluation de la précision des méthodes

---

<sup>1</sup>Une modalité d’imagerie est un type d’imagerie reposant sur la détection d’un signal dépendant d’une propriété physique de la matière : opacité aux rayons X, résonance magnétique nucléaire, opacité aux ultrasons, radioactivité.

<sup>2</sup>Certains systèmes d’imagerie spécifiques comme ceux dédiés à l’étude des petits animaux ont des résolutions spatiales du 10<sup>ième</sup> de *mm*

rigides de mise en correspondance. Elle repose sur la simulation réaliste d'images TEP à partir d'images par RM. Pour cela, nous construisons un modèle théorique de captation glucidique caractéristique du contexte clinique étudié (sujet sain et ischémique) à partir de mesures sur des examens réels. Les processus physiques d'une acquisition TEP sont modélisés dans un simulateur de type Monte Carlo. La stratégie est mise en oeuvre pour l'évaluation de trois méthodes de recalage rigide d'images TEP/IRM thoraciques et cardiaques. L'une d'entre elles a été développée dans le cadre de cette thèse.

Ce mémoire est organisé de la manière suivante :

- Le premier chapitre introduit le contexte médical et présente l'enjeu de l'utilisation conjointe de données obtenues en imagerie cardiaque anatomique et fonctionnelle. Différentes techniques d'exploration du coeur sont également présentées, suivi d'une étude bibliographique portant sur la mise en correspondance dans le contexte cardiaque et sur l'importance de l'évaluation des méthodes de traitement d'images médicales.
- Le deuxième chapitre est essentiellement bibliographique, présentant successivement le cadre conceptuel d'une évaluation des méthodes de traitement d'images, les caractéristiques de la procédure d'évaluation et une classification des études portant plus spécifiquement sur les méthodes de recalage en imagerie cardiaque.
- Le troisième chapitre décrit la stratégie que nous avons mise en oeuvre pour l'évaluation de la précision des méthodes de mise en correspondance en imagerie cardiaque TEP et IRM, la construction de la référence et l'identification des étapes d'une procédure d'évaluation dans le cas particulier du recalage d'images multimodales IRM/TEP.
- Le quatrième chapitre présente une évaluation comparative de trois méthodes de mise en correspondance multimodale IRM/TEP d'images de thorax et de coeur. Dans cette partie, nous décrivons la méthode de recalage que nous avons développée, ainsi que les deux autres méthodes existantes évaluées.
- Une conclusion synthétise les résultats obtenus et présente les perspectives de ce travail.

# Chapitre 1

## Contexte

### 1.1 Fusion d'informations multimodales pour l'étude de la viabilité myocardique

#### 1.1.1 Contexte médical

Première cause de mortalité dans les pays occidentaux, les maladies cardio-vasculaires constituent un enjeu de santé publique de première importance. D'après les estimations de l'Organisation Mondiale pour la Santé <sup>1</sup>, chaque année, 17 millions de personnes décèdent d'une maladie cardio-vasculaire dans le monde, et plus de 1.5 million à l'échelle européenne. En France, ces maladies sont à l'origine de près de 180 000 décès et 120 000 personnes sont victimes d'un infarctus du myocarde chaque année. Hormis l'hérédité, le sexe et l'âge, les principaux facteurs de risque cardio-vasculaire sont : le tabagisme, l'hypertension artérielle, le cholestérol, la sédentarité et le surpoids. Face à un tel fléau, les armes sont multiples, de la prévention aux dernières innovations thérapeutiques, en passant par une meilleure connaissance des gestes d'urgence.

Le système cardio-vasculaire a un fonctionnement d'une grande complexité, brièvement décrit en section 1.1.1. Les progrès récents de l'imagerie médicale multimodale et multidimensionnelle permettent une compréhension de plus en plus précise de l'anatomie et de la fonction des organes. Les techniques d'imagerie médicale permettent d'avancer dans la compréhension du fonctionnement du cœur normal et pathologique. Notre contribution dans cette thèse porte sur le développement d'une stratégie rigoureuse d'évaluation de méthodes de fusion d'informations issues d'images multimodales dans le but d'aider le clinicien à évaluer la viabilité myocardique dans les pathologies ischémiques (section 1.1.1).

#### Éléments d'anatomie et de physiologie du coeur

Le cœur est un organe constitué presque en totalité de tissus musculaires, que l'on appelle le *myocarde*. Ce myocarde s'insère dans un sac séreux, le *péricarde*. Les parois

---

<sup>1</sup>[http ://www.who.int/](http://www.who.int/)

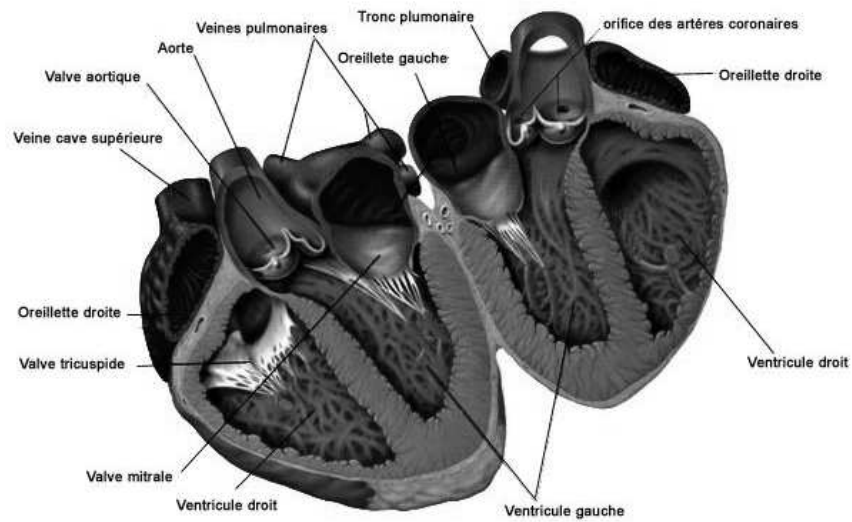


FIG. 1.1: Anatomie du cœur (Source : <http://www.medecine-et-sante.com/>)

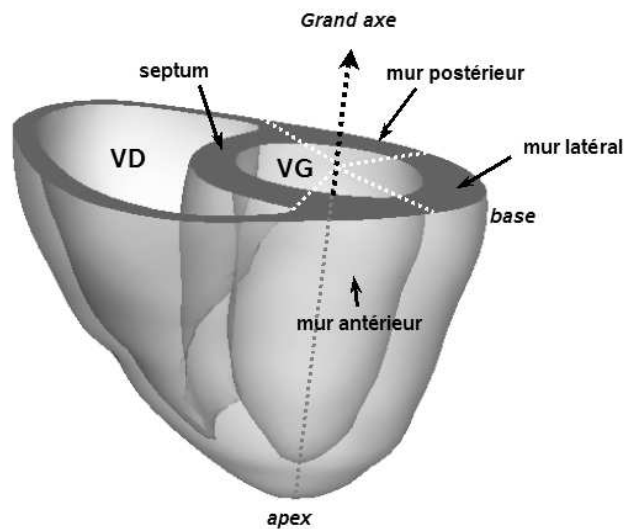


FIG. 1.2: Représentation schématique des ventricules gauche et droit du cœur.

interne et externe du myocarde sont respectivement nommées *endocarde* et *épicarde*. La pointe du coeur est appelée apex, et la partie haute est appelée base (figure 1.2). Les dimensions du coeur dépendent de l'individu, de son état d'entraînement et correspondent au moins à la taille du poing fermé de son porteur. Chez un individu moyen, il mesure environ 13 centimètres de long et 8 cm de large, et pèse moins de 500 grammes.

Le coeur est une pompe aspirante et refoulante qui se trouve au point de croisement de deux circulations sanguines : la *grande circulation* assurée par le coeur gauche (oreillette et ventricule gauches) et desservant tous les organes, et la *petite circulation* (circulation pulmonaire) assurée par le coeur droit (oreillette et ventricule droits) qui assure les échanges gazeux. Tous les vaisseaux qui expulsent le sang du coeur sont des artères et tous ceux qui l'acheminent au coeur sont des veines. Entre les artères et les veines se situent les capillaires (dans la grande comme dans la petite circulation). Les ventricules gauche (VG) et droit (VD) sont séparés par une cloison interventriculaire appelée le septum. Les parois du VG sont divisées en segments myocardiques identifiés en fonction de leur localisation antérieure, latérale, postérieure et septale (figure 1.1). Les parties droite et gauche fonctionnent en parallèle selon le schéma suivant :

- la partie droite reçoit des veines caves, qui s'abouchent dans l'oreillette droite, un sang non hématosé et l'envoie depuis le VD dans les poumons par l'artère pulmonaire,
- la partie gauche du coeur reçoit par les veines pulmonaires, qui s'abouchent dans l'oreillette gauche, le sang oxygéné provenant des poumons, et le propulse dans la circulation générale par l'aorte.

Le ventricule gauche est considéré comme l'élément principal de la pompe cardiaque. Il assure à lui seul 80% de la fonction cardiaque. Le VG peut être grossièrement représenté par une coque épaisse de forme semi-ellipsoïdique. On définit classiquement une direction longitudinale, le grand-axe du VG, qui constitue un système de repérage simple, allant de la pointe (l'apex) jusqu'à la base du ventricule. Une coupe dans un plan orthogonal au grand-axe est communément appelée vue en petit-axe. Le VD possède une paroi beaucoup plus fine et sa géométrie est plus complexe (figure 1.2). La surface périphérique du coeur est parcourue par un réseau vasculaire de coronaires qui assure l'irrigation sanguine du muscle.

Le myocarde est constitué d'un ensemble de cellules musculaires autonomes et contractiles. Il éjecte le sang dans l'organisme en se contractant de manière périodique. Au repos, le cycle cardiaque dure environ 800 ms (75 battements par minute) et se décompose en deux phases : la *diastole* (environ 60% du cycle cardiaque) et la *systole* (40% du cycle cardiaque). Pendant la *diastole*, les ventricules se remplissent du sang provenant des oreillettes. Pendant la *systole*, les deux ventricules se contractent et se vident dans leurs artères respectives à raison de 5 à 6 litres par minute. La *systole* correspond à une augmentation de la pression dans la cavité gauche suivie d'une contraction du VG, jusqu'à l'éjection du sang à travers l'aorte. La *télédiastole* est l'instant qui précède le début de la contraction ; la *télésystole* indique la fin de la contraction et le début du relâchement du muscle. Ces différentes phases sont repérées par rapport à l'Electrocardiogramme (ECG)

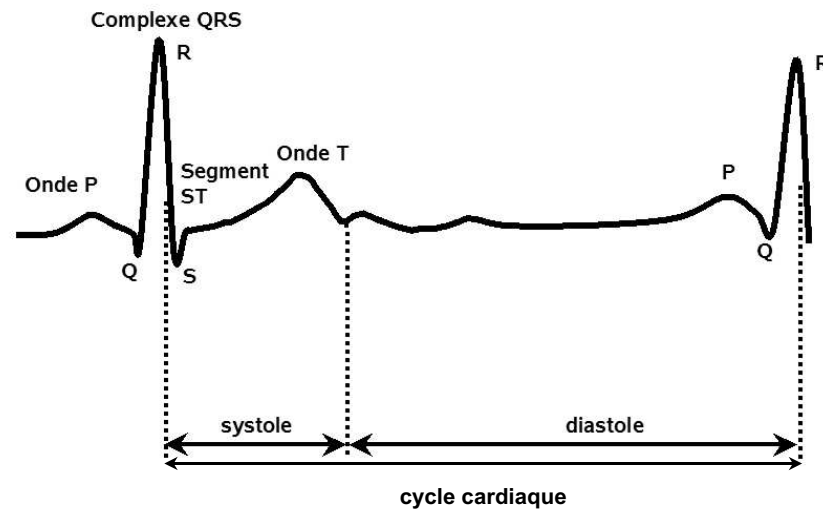


FIG. 1.3: Electrocardiogramme et cycle cardiaque.

représentant l'évolution de l'activité électrique globale au niveau du cœur (figure 1.3). Le fonctionnement du cœur à l'échelle cellulaire et moléculaire met en jeu simultanément des processus chimiques basés essentiellement sur l'action des ions calcium, électriques, liés à la propagation d'une onde de dépolarisation à la surface du cœur, mécaniques qui vont engendrer la contraction effective des cellules myocardiques. Pour un état de ces connaissances, on pourra se référer à l'ouvrage collectif de Pinet *et al.* [Pinet 02].

## L'ischémie myocardique et ses conséquences

### Ischémie myocardique

Le cœur apporte le sang oxygéné à tout l'organisme, mais doit également assurer sa propre perfusion grâce aux artères coronaires. L'adaptation à une demande accrue au cours de l'effort ne peut se faire que par l'augmentation du débit sanguin.

Pour expulser le sang dans tout le corps, le myocarde a besoin d'énergie, qu'il trouve dans des substrats spécifiques présents dans le sang. La réduction de l'apport sanguin qui peut se produire lors de l'obstruction d'une artère coronaire (athérosclérose) entraîne une suite de conséquences connue sous le nom de *cascade ischémique*. En simplifiant, le déficit de perfusion implique une réduction du métabolisme qui, si elle persiste, aboutit à une réduction de la capacité contractile du cœur. Le sang est ainsi moins efficacement distribué à l'organisme. La pathologie associée à ce mécanisme porte le nom d'*ischémie myocardique*. A son stade ultime, l'*ischémie myocardique* conduit à l'infarctus du myocarde.

### La viabilité myocardique

L'*ischémie myocardique* résulte d'un déséquilibre entre perfusion, métabolisme oxydatif et fonction contractile, provenant d'une insuffisance de la perfusion dans un territoire cardiaque donné. Une ischémie prolongée peut entraîner la nécrose des tissus et déclencher

un *infarctus*. L'état de la région myocardique atteinte est cependant réversible selon les délais de re-perfusion et l'action de la circulation collatérale, d'où la notion de *viabilité des tissus*. Une définition de la viabilité est la suivante : *un tissu sujet à une perfusion insuffisante est dit viable s'il retrouve sa fonction après re-perfusion* [Janier 98].

L'estimation de la viabilité d'un tissu est une étape essentielle pour guider le choix d'une procédure de revascularisation et la thèse de A. Mazzadi illustre la complexité des facteurs qui interviennent dans l'ischémie myocardique [Mazzadi 03].

### 1.1.2 Techniques d'exploration du coeur

Les résultats les plus récents relatifs aux pathologies ischémiques montrent que la viabilité des tissus myocardiques peut être estimée à partir de l'analyse conjointe de la perfusion, du métabolisme et de la fonction contractile, d'où l'intérêt d'une imagerie multimodale et intégrative [Behloul 01]. La complexité des phénomènes biologiques est telle qu'elle ne peut pas être explicitée par une seule technique d'exploration. La nature complémentaire des techniques d'exploration existantes aide le clinicien, dans un contexte pathologique, à affiner son diagnostic, à préparer et même à réaliser les procédures thérapeutiques ou, dans un contexte de recherche, à mieux comprendre le fonctionnement du coeur normal et pathologique.

Pour accéder à la morphologie cardiaque, on fait appel à l'Imagerie par Résonance Magnétique (IRM) [Gallippi 01], aux rayons X [Dey 99] ou l'imagerie ultrasonore (US) [Gallippi 01].

L'information fonctionnelle liée au métabolisme peut être évaluée en utilisant la Tomographie par Émission de Positons (TEP ou Positron Emission Tomography -PET- en anglais) et celle liée à la perfusion en utilisant la Tomographie par Émission Monophotonique (TEMP ou Single Photon Emission Computed Tomography -SPECT- en anglais), l'IRM, l'imagerie ultrasonore, ou la TEP. La TEP de captation glucidique au fluorodésoxyglucose marqué (FDG18) est la technique de référence pour l'étude de la viabilité myocardique [Hartiala 95].

L'échographie, le ciné IRM et l'IRM de marquage tissulaire [Zerhouni 88], fournissent des informations sur la fonction contractile. On sait aussi que l'ischémie peut altérer l'activation électrique du cœur. En conséquence, le diagnostic de l'ischémie peut aussi être envisagé à partir de l'enregistrement de différences de potentiels électriques en électrocardiographie (ECG) ou la mesure d'un champ magnétique en magnétocardiographie (MCG) [Siltanen 88]. En MCG, on mesure sur le torse du patient, des champs magnétiques faibles émis par le cœur et engendrés par des courants électriques parcourant les cellules actives du muscle cardiaque. Cette méthode, de plus en plus utilisée, est totalement non invasive et se révèle être très efficace pour l'étude de l'activité bioélectrique du cœur. Ainsi, l'analyse de ces acquisitions permet la localisation des régions du muscle cardiaque où l'activité électrique est anormale [Nenonen 01].

Les principales techniques d'imagerie impliquées dans l'étude des pathologies cardiaques ischémiques sont présentées dans la figure 1.4. Un résumé de plusieurs techniques

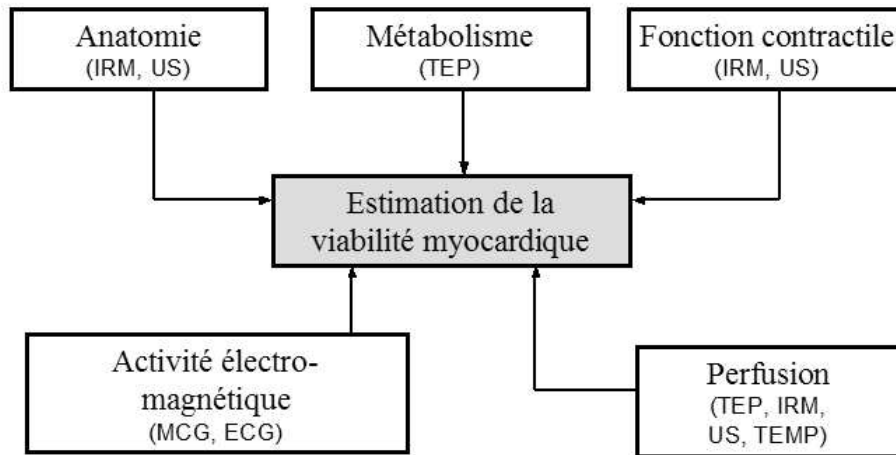


FIG. 1.4: Modalités d'imagerie impliquées dans l'estimation de la viabilité myocardique.

d'acquisition d'images médicales est proposé en Annexe A. Le lecteur peut également se reporter à l'état de l'art très complet de [Acharya ].

### 1.1.3 Mise en correspondance d'images dans le contexte cardiaque

Dans le cadre de l'étude de la viabilité myocardique, la fusion de données anatomiques et fonctionnelles implique, en préalable, de trouver les moyens de mettre en correspondance des territoires analogues imagés dans différentes modalités. Avant le développement d'outils de recalage assisté par ordinateur, la fusion était réalisée par une simple mise en correspondance mentale par les médecins, s'appuyant sur des structures communes et facilement identifiables.

Parmi les besoins les plus courants du praticien se trouvent une variété de tâches de *comparaison* : comparer des images d'un même patient acquises à des instants différents ou selon des modalités différentes (images multimodales), comparer des images de patients différents, ou encore comparer une image avec un atlas anatomique ou fonctionnel. Ces tâches de comparaison relèvent toutes d'une même problématique du traitement d'images : *le recalage* ou la *mise en correspondance d'images*.

Le recalage d'images, en anglais *image registration*, est la procédure ayant pour but de trouver une transformation entre deux jeux de données pour compenser des variations spatiales ou d'intensité dues, par exemple, à l'utilisation de systèmes d'acquisition différents ou à des différences de conditions ou de géométrie d'acquisition. L'objectif est de situer les images dans un même repère géométrique.

Le recalage d'images cardiaques est un problème difficile. Nous en énonçons les principales raisons inhérentes à l'organe étudié :

- le coeur est en perpétuel mouvement, sans compter le mouvement respiratoire des structures thoraciques, d'où la difficulté accrue du recalage d'images cardiaques par rapport aux images de cerveau, par exemple, qui est un organe quasi-statique.



- l'aspect lisse du coeur ne permet l'identification que d'un nombre très limité de structures anatomiques : jonctions interventriculaires, apex, muscles papillaires, etc...
- les images acquises présentent des caractéristiques très différentes en termes de résolutions spatiale et temporelle, et surtout de contraste (deux images fonctionnelles de métabolisme et de perfusion ne présenteront pas le même contraste).
- les techniques imagerie ne sont pas parfaites : elles peuvent générer des distorsions géométriques. Ce sont des déformations virtuelles dans la mesure où elles ne caractérisent pas une évolution physique de l'objet. Le simple fait qu'une image numérique soit échantillonnée à une certaine résolution spatiale implique une incertitude sur la localisation de l'objet représenté ; c'est en soi une source de distorsion géométrique. D'autres types de distorsions, parfois beaucoup plus sévères, sont dues à l'inexactitude des modèles qui sous-tendent les procédés d'imagerie (hypothèse de célérité constante en imagerie ultrasonore, hypothèse de linéarité du champ magnétique en IRM, etc...).

### Classification des techniques de recalage

De nombreuses méthodes de recalage ont été proposées dans la littérature. Plusieurs articles de synthèse, spécifiques au traitement d'images médicales en général, constituent d'excellentes introductions techniques à la problématique de recalage [Brown 92, Maintz 98, Audette 00, van den Elsen 93, Zitova 03]. Mäkelä *et al.* proposent plus particulièrement une synthèse des méthodes de recalage en imagerie cardiaque [Mäkelä 02].

Les procédures de recalage sont généralement décomposées selon quatre aspects : l'espace des transformations, l'espace des paramètres, la fonction de coût et la méthode d'optimisation [Barillot 99, Sarrut 00].

Soit  $I_1$  et  $I_2$  les images à recaler, le processus général de mise en correspondance peut être résumé formellement de la manière suivante. Il s'agit de trouver une transformation  $T$  qui maximise une fonction de coût  $S$  :

$$T^* = \underset{T \in C_T}{argmax} S(f_1(I_1), f_2(T(I_2))), \quad (1.1)$$

- $T^*$  est la transformation optimale recherchée,
- $f_1$  et  $f_2$  sont les fonctions extrayant des caractéristiques homologues des images,
- $S$  une fonction ou critère de similarité, définie sur l'espace des caractéristiques extraites par  $f_1$  et  $f_2$ ,
- $argmax$  est la procédure d'optimisation de la mesure de similarité, recherchant le maximum dans l'espace  $C_T$  défini par les paramètres de  $T$ .

Le recalage entre deux jeux de données consiste donc en l'optimisation du critère de similarité mesuré sur les informations homologues extraites des données.

Il existe quatre contextes de recalage qu'il est fondamental d'identifier lors de la mise en place d'une procédure de recalage. Ils permettent de choisir au mieux le type de transformation géométrique ou en intensité recherchée, ou encore d'identifier la nature de l'information homologue entre les données à recaler [Jannin 01] :

- *recalage intra-sujet et intra-modalité* : ce contexte concerne la fusion entre deux jeux de données provenant d'une même modalité chez un même sujet. Trois principaux objectifs cliniques sont associés à ce contexte : (1) Le suivi dans le temps de modifications dans l'anatomie du patient permet, par exemple, de suivre la croissance ou d'étudier l'évolution d'une zone lésionnelle ou les effets d'un traitement thérapeutique. En chirurgie ou en radiothérapie, ce contexte de fusion permet de comparer des images pré et post-traitement. (2) Ce contexte peut aussi être utilisé pour comparer différents états du sujet (stimulation/repos). (3) L'imagerie par soustraction est généralement basée sur le calcul d'une différence entre deux jeux de données image correspondant à deux acquisitions différentes réalisées avec et sans traceur ou agent de contraste. Le recalage intra-sujet et intra-modalité permet aussi de corriger les éventuels mouvements du sujet pendant une acquisition, notamment dans les séries temporelles où un léger déplacement pendant l'acquisition peut influencer sur l'analyse ultérieure. Pour ces applications, la transformation géométrique reliant les données à recaler est souvent considérée comme rigide.
- *recalage intra-sujet et inter-modalités* : cette catégorie permet de profiter de la nature complémentaire des différentes modalités d'imagerie. La mise en correspondance entre des données IRM et TEP permet conjointement la localisation anatomique des différentes structures et de leur métabolisme, ce qui fournit une corrélation anatomo-fonctionnelle, utile notamment pour guider le choix de la thérapie. Le choix du type de la transformation géométrique (rigide ou élastique) est complètement dépendant de l'application clinique. La nature de l'information homologue n'est pas évidente surtout lorsqu'une modalité représente l'anatomie et l'autre une fonction physiologique (e.g. métabolisme en TEMP).
- *recalage patient et système de traitement* : ce contexte concerne la mise en correspondance entre le patient et les outils thérapeutiques (systèmes robotiques, systèmes d'imagerie interventionnelle, outils chirurgicaux, etc). Les applications de ce type se rencontrent couramment en neurochirurgie ou en radiologie interventionnelle. Les systèmes de thérapie guidés par l'image, les systèmes de neuronavigation, sont de plus en plus utilisés en routine clinique.
- *recalage inter-sujets et intra- ou inter-modalités* : ce contexte intéresse la fusion entre des données mono ou multimodalités concernant des sujets différents. Les applications principales de cette catégorie sont liées à l'étude de la variabilité anatomique ou fonctionnelle au sein d'une population (témoins ou patients). Ce type de recalage permet de construire des modèles ou des gabarits ("template" en anglais) correspondant à chaque modalité d'imagerie, de comparer les données d'un sujet à un atlas ou à un gabarit, de comparer des données concernant des sujets différents appartenant à une même population ou, enfin, de comparer deux populations différentes de sujets. Le recalage dans un référentiel commun, de données provenant de sujets différents, mais d'une même modalité et utilisant un même protocole d'acquisition, permet de mettre en évidence l'existence de différences ou de points communs via une analyse statistique voxel par voxel des données ainsi normalisées spatialement.

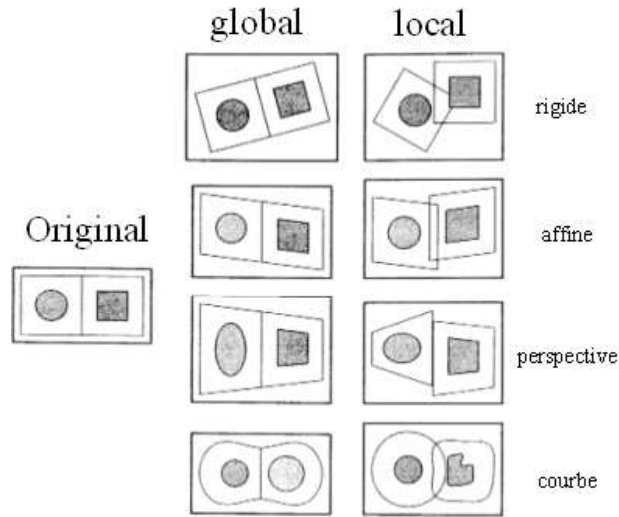


FIG. 1.5: Types de transformations géométriques d'après [Maintz 98]

Une telle démarche peut être utilisée pour des études statistiques sur la morphologie ou la pathologie, par comparaison avec une population de sujets sains. Elle est utilisée en imagerie cérébrale mais quasiment pas en imagerie cardiaque à cause de la variabilité anatomique inter-sujets trop importante et des difficultés liées au mouvement.

### Espace de recherche

L'espace de recherche est défini par la nature de la fonction de correspondance recherchée entre les structures homologues de deux images. Il s'agit généralement d'une transformation purement spatiale, mais il peut être intéressant d'inclure dans cet espace une transformation portant sur les intensités. Les transformations géométriques modélisent les variations géométriques que l'on cherche à compenser entre les données à recalculer. Les transformations peuvent être *globales*, c'est à dire appliquées à toute l'image, ou *locales*, l'image est alors découpée en régions possédant leur propre transformation. Leur nature peut être linéaire ou non-linéaire. La figure 1.5 illustre les principaux types de transformations [Maintz 98].

**Transformations rigides.** Les transformations rigides (rotation et translation, 6 paramètres en 3D et 3 paramètres en 2D) forment le plus petit espace des transformations possible pour un algorithme de recalage. L'hypothèse de rigidité convient au cas où l'on cherche à compenser la différence de positionnement d'un patient par rapport aux capteurs sans tenir compte d'éventuelles déformations des tissus imagés ou de distorsions géométriques créées par les procédés d'imagerie.

Les transformations rigides peuvent être définies sous forme matricielle par :

$$T(x) = Rx + t,$$

$t$  étant un vecteur colonne  $3 \times 1$  représentant la translation, et  $R$  étant une matrice orthogonale directe  $3 \times 3$  ( $RR^t = I_3$  et  $\det R = 1$ ) représentant une rotation.

Il existe plusieurs représentations possibles de la rotation [Horn 87] : le vecteur rotation, les angles d'Euler ou le quaternion unitaire (voir Annexe C).

**Similitudes.** Les similitudes sont des transformations à 4 degrés de liberté en 2D et 7 en 3D. Elle généralisent les transformations rigides par l'adjonction d'un facteur d'échelle global  $s$  :

$$T(x) = sRx + t, s > 0$$

Ce type de transformations ne change pas les relations géométriques entre les points (un triangle reste un triangle). Il conserve également les angles et les rapports de distance.

**Transformations affines.** De façon plus générale, les transformations affines sont définies par :

$$T(x) = Ax + t,$$

où  $A$  est une matrice  $3 \times 3$  quelconque. L'espace des transformations affines est de dimension 12 en 3D (6 en 2D). Le facteur d'échelle n'est plus forcément isotrope, puisqu'il y a un facteur par dimension de l'image. Les transformations de type torsion (*shear* et *skew* en anglais) sont également incluses.

**Transformations perspectives ou projectives.** Les transformations projectives conviennent typiquement au cas où des images 3D sont recalées avec des images 2D de type projection acquises avec une caméra (satellite, vidéo, images radiologiques ou portales). Cette classe de transformations permet de prendre en compte les effets de perspective liés à la formation de l'image 2D. Les transformations projectives préservent les lignes droites mais pas le parallélisme. Il y a 8 paramètres en 2D et 12 en 3D, mais le nombre de paramètres dépend de la connaissance a priori des paramètres des systèmes d'acquisition.

**Transformations courbes ou élastiques.** Les transformations élastiques sont des transformations non linéaires modélisant des variations spatiales complexes. Le nombre de paramètres est variable et généralement important. Ce type de transformation change les droites en courbes. Les transformations non-linéaires sont principalement utilisées dans deux types de situations [Jannin 01] :

- quand le jeu de données à recalculer ne correspond pas à une même configuration de la structure anatomique (dans le cas par exemple du recalage inter-sujets, ou de recalage vers un atlas ou vers un modèle)
- quand une structure anatomique commune existe entre les données mais qu'elle est distordue par une déformation non-rigide : correction de déformations liées à l'acquisition, recalage de données portant sur des structures anatomiques en mouvement (cœur, poumons, mouvements dus à la respiration, etc.) ou des structures ayant des états différents (dus, par exemple, à la comparaison d'une structure saine avec une structure pathologique, ou à la comparaison de données pré-opératoires

avec des données post-opératoires ou à la croissance de la structure pathologique ou anatomique)

### Espace des paramètres

L'espace des paramètres est défini par les informations homologues entre les données à recalcr. L'information homologue correspond à une information extraite des données à recalcr et sur laquelle est fondée la recherche de la transformation géométrique optimale au sens d'une métrique de similarité.

L'information homologue peut être classée par sa dimension : 0D (point), 1D (contour), 2D (surface), 3D (volume), nD (hypersurface) avec ou sans dimension temporelle supplémentaire [Barillot 99].

L'information homologue est aussi fréquemment qualifiée d'*extrinsèque* ou d'*intrinsèque*. Les méthodes de recalcr fondées sur une information homologue extrinsèque utilisent des objets artificiels liés au sujet, appelés *marqueurs externes*. Dans le cas multimodal, les marqueurs externes doivent être visibles dans l'une et l'autre des modalités à recalcr. Par conséquent, ces méthodes nécessitent un protocole particulier d'acquisition d'images. Les techniques invasives utilisent des marqueurs vissés directement dans les structures osseuses du patient ou des cadres stéréotaxiques utilisés en neurochirurgie. Les techniques non invasives utilisent, par exemple, des marqueurs collés sur la peau du patient ou des cadres qui empêchent le patient de bouger (ex. cadre en mousse). Plusieurs auteurs s'appuient sur des marqueurs externes pour guider le recalcr d'images multimodales du thorax et du coeur [Gilardi 98, Yu 95, Dey 99] et du cerveau [Maurer 93].

L'utilisation de marqueurs non invasifs est susceptible d'induire des erreurs de localisation dues au léger déplacement des marqueurs entre deux examens, et à la flexibilité du corps. Les marqueurs invasifs sont en général localisés avec plus de précision. Les méthodes de recalcr qui utilisent des marqueurs extrinsèques sont limitées aux organes statiques (imagerie du cerveau et orthopédique) et reposent sur une transformation rigide. En imagerie cardiaque, l'utilisation de la position de marqueurs sur la peau ne résout pas le problème de recalcr lié aux mouvements relatifs du coeur par rapport au corps du patient. Les transformations non rigides du coeur peuvent être parfois estimées à l'aide de marqueurs extrinsèques implantés sur la paroi du coeur dans les études menées sur les animaux [Duncan 94].

Les méthodes utilisant une information de nature intrinsèque utilisent différentes caractéristiques des images qui peuvent être :

- des *repères anatomiques* : des points anatomiques connus (par exemple commissures antérieure et postérieure dans le cas d'images cérébrales, ou les points de jonction interventriculaire pour les images cardiaques), des lignes ou des surfaces anatomiques remarquables (surface de la peau, surface des ventricules), ou enfin des volumes anatomiques (cerveau, structures osseuses, vaisseaux). Les points anatomiques sont choisis manuellement alors que les lignes, surfaces ou volumes nécessitent une procédure de segmentation manuelle ou automatisée. Les repères anatomiques servent

également à l'évaluation d'une méthode de recalage [Gilardi 98]. Dans [Faber 91], les surfaces cardiaques sont automatiquement détectées dans les images TEMP cardiaques synchronisées sur l'ECG et sont recalées de manière rigide avec les surfaces correspondantes extraites des images par RM.

La définition des repères anatomiques suppose que ces structures anatomiques sont détectables dans les deux images à recaler. Dans le cas des images cardiaques, ces invariants ne sont pas faciles à identifier. La difficulté de détecter des structures anatomiques peut encore être accentuée dans le cas pathologique [Gilardi 96].

- des *primitives géométriques* : qui consistent en des points, lignes, surfaces ou volumes qui ne correspondent pas directement aux structures anatomiques mais à des caractéristiques issues des images. Sinha *at al.* ont développé une approche de mise en correspondance d'images par RM et TEP-FDG par l'intermédiaire de surfaces déterminées interactivement [Sinha 95].
- les *intensités d'images* (approches dites *iconiques*). Ces approches s'appuient sur les intensités des images avec la définition d'un critère de similarité ou de dissimilarité basé sur des mesures statistiques [Fitzpatrick 00] ou sur les intensités pondérées par la localisation du pixel ou voxel (axes principaux [Maurer 93], moments d'inertie [Dey 99, Slomka 95]). Les méthodes de recalage basées sur des mesures statistiques de similarité occupent une place privilégiée parmi les méthodes de recalage de données médicales. C'est surtout le cas des approches entropiques, notamment de l'information mutuelle, qui ont montré leur efficacité et leur robustesse dans un grand nombre de situations, notamment les problèmes de recalage mono- ou multimodalités rigide ou non-linéaire. Les méthodes iconiques ne requièrent en général aucun pré-traitement et elles sont préférées quand il est difficile d'extraire des structures communes des jeux de données. Elles facilitent aussi l'automatisation des méthodes de recalage.

## Fonctions de coût et mesures de similarité

La fonction de coût correspond à un critère objectif, utilisé pour estimer la qualité du recalage entre deux structures homologues pour une transformation donnée. Elle est définie en fonction de l'information homologue et du type de la transformation à estimer. On peut identifier deux principales familles de fonctions de coût : celles basées sur des distances entre les primitives homologues dans l'espace des voxels des images et celles basées sur des mesures de similarité dans l'espace des intensités des images.

Lorsque l'information homologue est de nature extrinsèque aux données, la fonction de coût correspond souvent à une estimation de la distance euclidienne moyenne entre deux ensembles de primitives. Le problème de la mise en correspondance de points de type marqueurs externes ou références anatomiques est généralement formulé par une approche de type "minimisation aux moindres carrés". Quand les structures homologues sont plus complexes, d'autres critères de distance sont utilisés comme la méthode "head&hat" [Pelizzari 89, Sinha 95, Faber 91, Yu 95]. La *distance de Chamfrein* [Borgefors 88] inter-

vient dans certains algorithmes de recalage pour l'estimation de la distance moyenne entre des surfaces et/ou des nuages de points [Pallotta 95, Cai 99, Mäkelä 01, Gilardi 96]. L'algorithme du *Point le plus Proche Itéré* (*Iterative Closest Point* (ICP)) introduit par Besl et McKay [Besl 92] a été très largement utilisé, en particulier pour le recalage non-rigide d'images cardiaques TEMP dans [Declerck 97].

Quand le recalage repose sur une information homologue basée sur la distribution des intensités de l'image, les critères introduisent des mesures statistiques. Le fondement commun aux nombreuses mesures proposées dans la littérature est l'idée que les intensités de deux images manifestent une dépendance statistique d'autant plus forte que les images sont bien alignées. Le rôle de la mesure de similarité est précisément de donner une signification quantitative à cette notion de dépendance. On présente une classification indicative des mesures selon le type des hypothèses de dépendance en suivant *A. Roche* [Roche 01] :

- *Conservation de l'intensité :*

Les mesures de cette classe considèrent comme idéale la situation où les images sont identiques. Elles sont le plus souvent calculées à partir d'une image de différence. Les plus usitées sont la somme des différences au carré (SDC) et la somme des différences absolues (SDA) [Brown 92]. Le changement stochastique de signe (CSS) est un comptage des occurrences négatives et positives des différences d'intensités. Hoh *et al.* [Hoh 93] ont mis en oeuvre les critères SDA et SDC dans le cadre du recalage monomodal rigide d'images cardiaques acquises en TEP d'émission. Slomka *et al.* [Slomka 95] optent pour la SDA pour le recalage affine entre des images de coeur en TEMP d'émission et un modèle de référence. La moyenne des différences au carré (MDC) a été utilisée pour des problématiques de compensation de mouvements dus à la respiration [Bidaut 01].

- *Dépendance linéaire ou affine :*

Les intensités sont reliées par une fonction affine, c'est à dire :  $I \approx \alpha J + \beta$ . Le cas où le paramètre  $\beta$  est supposé nul correspond à une dépendance linéaire. Les mesures adaptées sont pour la plupart des variantes du Coefficient de Corrélation (CC) [Brown 92] ainsi que la Variance du Rapport des Intensités (VRI) [Woods 92]. CC est souvent un excellent choix pour un recalage monomodal d'images de cerveau [Fitzpatrick 00]. Turkington *et al.* [Turkington 97] ont choisi le coefficient de corrélation pour recalculer en translation des séries dynamiques de coeur acquises en TEP, par l'intermédiaire d'un modèle de séries dynamiques de référence. Le même critère, appliqué à des images cardiaques acquises en TEP d'émission, intervient dans [Bacharach 93].

- *Dépendance fonctionnelle :*

Pour des images de modalités différentes, l'hypothèse de dépendance affine est souvent fautive. On peut supposer l'existence d'une relation fonctionnelle plus générale (non-linéaire, non-monotone) entre les intensités. Dey *et al.* traitent du recalage multimodal rigide entre les images de thorax TEMP et la tomographie par rayons X [Dey 99], en minimisant un critère similaire au VRI [Woods 93]. Eberl *et al.* [Eberl 96] analysent les performances de différents critères de similarité dans le contexte du recalage monomodal d'images cardiaques en TEMP, et multimodal entre des images acquises en TEP et en

TEMP, et constatent que le meilleur critère est le SDA lorsque les images traitées sont de même modalité, et le VRI dans le cas multimodal. Dans [Roche 98], le Rapport de Corrélation (RC) est employé pour le recalage d'images multimodales de cerveau, et appliqué dans [Mäkelä 03b, Pauna 03] pour le recalage d'images multimodales de thorax par RM et TEP.

- *Dépendance statistique :*

L'hypothèse d'une dépendance fonctionnelle, même très générale, reste discutable. La théorie de l'information fournit une multitude de critères de nature statistique, parmi lesquels l'entropie conjointe, le coefficient de corrélation entropique et surtout l'information mutuelle (IM) [Viola 95]. L'information mutuelle est maintenant reconnue comme un critère performant et robuste pour la mise en correspondance d'images multimodales, comme l'attestent les recherches qui ont porté sur le recalage d'images cérébrales [Pluim 03]. Slomka *et al.* emploient l'IM pour aligner des images de thorax en TEP et en tomographie par rayons X [Slomka 00]. Une autre application est le recalage d'images par RM de l'abdomen [Carrillo 01]. Dans le domaine de l'imagerie cardiaque, Lötjönen *et al.* ont proposé une approche mettant en jeu le critère de l'IM, l'information du gradient des images, et des déformations élastiques [Lötjönen 01]. Cette méthode est appliquée à des images cardiaques par RM et TEP acquises sur le même patient.

La plupart des critères de similarité décrits ci-dessus sont estimés via des histogrammes joints (voir définition en Annexe B), nécessitant l'utilisation de méthodes d'interpolation comme le plus proche voisin, l'interpolation trilinéaire ou par volume partiel [Maes 97].

## Méthodes d'optimisation

Un autre aspect fondamental d'un algorithme de recalage est le choix de la méthode d'optimisation. La méthode d'optimisation définit la manière selon laquelle la mesure de similarité atteint son coût minimum. Elle permet le calcul de la transformation optimale au sens de la métrique de similarité mesurée sur les informations homologues. De nombreuses méthodes existent dans la littérature [Press 92]. Le projet "Vanderbilt" [West 97] a permis notamment d'évaluer la précision de plusieurs méthodes d'optimisation de l'information mutuelle pour le recalage rigide multimodal d'images cérébrales. Plus récemment, dans [Jenkinson 00, Jenkinson 01], la robustesse de différentes méthodes a été comparée dans le contexte du recalage affine inter-sujets d'images par RM de cerveau.

Les techniques d'optimisation les plus courantes, ne nécessitant le calcul de la fonction de coût qu'en certains points sans évaluation de son gradient, sont la méthode de Powell et la méthode de Nelder et Mead, dite du simplexe [Press 92]. Ces deux méthodes ont été intensivement utilisées pour le recalage rigide/affine [Maintz 98]. Elles présentent l'avantage d'être relativement rapides. En revanche, elles ne garantissent que la convergence vers un minimum local et sont de fait sensibles à l'initialisation des paramètres de la transformation. Dans certains cas, il est possible d'atténuer ce manque de robustesse par une initialisation judicieuse.

Les méthodes avec la prise en compte du gradient, tels que la descente de gradient ou



le gradient conjugué, nécessitent une estimation de la dérivée du critère. Le gradient d'une fonction donne localement la direction dans laquelle celle-ci décroît le plus. Par nature, les descentes de gradient sont adaptées à la recherche rapide d'un minimum local. Différentes variantes de descente de gradient ont été utilisées pour le recalage rigide/affine d'images de cerveau [Maes 99, Woods 98].

Pour s'affranchir du problème d'initialisation, Hill *et al.* ont proposé d'utiliser les algorithmes génétiques [Hill 94]. Sous certaines conditions, on peut montrer que les algorithmes génétiques garantissent l'obtention du minimum global. Ces conditions impliquent malheureusement des temps de calculs considérables.

Une synthèse des références de recalage d'images de thorax et cardiaques est présentée dans [Mäkelä 02]. Nous présentons également dans les tableaux 2.2 et 2.3 (section 2.3.3) celles pour lesquelles des stratégies d'évaluation ont été mises en oeuvre.

## 1.2 L'importance de l'évaluation

Aujourd'hui, le clinicien dispose d'un ensemble de techniques d'imagerie médicale devenues incontournables dans les démarches diagnostiques et thérapeutiques. Ces diverses modalités d'imagerie fournissent des images de plus en plus précises de nature anatomique (morphologie) et fonctionnelle (métabolisme, perfusion d'organes, etc...) complémentaires pour un même sujet. Cependant, la variété des informations disponibles pour réaliser le diagnostic est suivie d'une augmentation drastique de la quantité des données que le praticien doit intégrer et interpréter.

Afin d'exploiter pleinement l'apport considérable de ces multiples sources d'information et en particulier leur nature quantitative, il est de plus en plus nécessaire de les fusionner et d'en extraire une représentation cliniquement utile. Cette représentation appropriée de l'information nécessite très souvent un post-traitement des images.

Un grand nombre d'outils de traitement d'information et d'images est proposé aux cliniciens sur les consoles d'acquisition d'images. Cependant, les techniques de traitements d'images les plus avancées restent très souvent dans le domaine de la recherche et leur transfert vers les systèmes industriels nécessite préalablement la mise en oeuvre d'une importante étape d'évaluation. Cette évaluation est réalisée en particulier dans les démarches de mise aux normes et d'accréditation des techniques utilisées en routine clinique.

Il existe un décalage entre l'avancement de certains outils de post-traitement et les outils proposés au clinicien. Une des raisons principales de ce décalage est la nécessité de l'évaluation/validation des outils de traitement développés. L'évaluation rigoureuse des performances des techniques de traitement d'images conditionne ainsi leur exploitation effective dans les démarches diagnostiques et thérapeutiques en routine clinique. Cependant, force est de constater que l'évaluation est le maillon faible. Des techniques très sophistiquées et potentiellement très intéressantes d'un point de vue clinique ne seront ainsi pas utilisées par manque d'une évaluation rigoureuse. Ce constat résulte très généralement de la difficulté méthodologique de disposer d'une référence afin de mesurer les

performances des techniques à évaluer. Il est difficile de construire un protocole d'évaluation/validation car, en imagerie médicale, la *vérité terrain* n'est pas disponible.

La communauté de traitement d'images s'interroge de plus en plus sur le problème de l'évaluation empirique lors de groupes de travail, d'éditions spéciales ([Bowyer 98, Gee 00, Shahidi 01, Loew 97, Shtern 99]) et de conférences (*Validations on Medical Image Processing*, *VMIP*<sup>1</sup>, MICCAI'03). Cependant, l'évaluation des performances des algorithmes dans le domaine de la Vision par Ordinateur n'occupe pas la place qu'elle mériterait compte tenu de son importance. D'après Forstner *et al.*, les appels aux conférences internationales incitent plus à présenter de nouveaux algorithmes que d'évaluer les performances d'algorithmes déjà existants [Förstner 96]. La difficulté de développer des algorithmes dont la procédure d'évaluation des performances soit bien décrite est une barrière psychologique importante.

Les évaluations de performances d'un algorithme sont très coûteuses en temps. En effet, Förstner indique un rapport de temps investi de 1 à 10 entre la théorie et l'implantation et entre l'implantation et les tests également [Förstner 96]. On ne peut pas prédire les performances d'un algorithme seulement à travers des évaluations théoriques et empiriques. Les tests empiriques aident à régler les paramètres d'un algorithme (ex : le bruit) qui vont être utilisés pour évaluer ses performances.

Typiquement, on fait appel à des méthodes d'évaluation pour déterminer le système ou la technique d'analyse d'images qui conduit à la meilleure image ou à l'information la plus pertinente dans un contexte donné. Il est fondamental d'identifier les questions auxquelles on se propose de répondre dans le cadre d'une validation.

## Evaluation dans le contexte de la mise en correspondance

Beaucoup d'efforts de recherche en traitement d'images médicales ont été dirigés vers la définition d'outils efficaces et rapides de traitement, de mise en correspondance et de visualisation. La mise en correspondance ou recalage d'images constitue une étape essentielle des systèmes d'aide à l'interprétation d'images médicales morphologiques et fonctionnelles. L'évaluation peut être considérée comme un composant à part entière d'une démarche de qualité. De telles procédures de qualité ne sont pas faciles à établir : la résolution et la précision du recalage sont parfois délicates à quantifier ; il existe peu de jeux de données standards d'évaluation et il n'existe pas, à ce jour, de procédure standardisée d'évaluation. Un état de l'art sur les différentes approches d'évaluation est présenté dans le chapitre 2.

Notre travail de thèse concerne l'évaluation de méthodes de recalage en imagerie cardiaque multimodale pour l'étude de la viabilité myocardique. Dans ce cadre, notre recherche s'est orientée vers le développement d'une stratégie rigoureuse d'évaluation des méthodes de recalage d'images par RM et TEP à partir d'un simulateur Monte Carlo d'images TEP.

---

<sup>1</sup><http://idm.univ-rennes1.fr/VMIP/miccai2003/>

## Chapitre 2

# Évaluation des méthodes de traitement d'images médicales

Les progrès observés dans les techniques d'imagerie médicale s'accompagnent d'un nombre important de travaux concernant le traitement, l'analyse et l'interprétation des images médicales. Un examen de la littérature montre qu'une évaluation rigoureuse des performances des procédures de traitement d'images médicales proposées est souvent absente. Cependant, le problème de l'évaluation des méthodes de traitement est actuellement l'objet d'un intérêt croissant dans une démarche visant à caractériser la qualité d'une méthode vis à vis non seulement de son originalité mais également de ses performances dans l'application ciblée [Bowyer 98]. Conceptuellement, la situation est illustrée par la figure 2.1. Si les hypothèses émises lors de la conception d'un algorithme complexe ne sont pas satisfaites par l'application, les performances de l'algorithme peuvent être significativement réduites.

Nous distinguons dans ce mémoire *l'évaluation* de la *validation*. La *validation* est un terme utilisé principalement dans un contexte médical et clinique et a pour objectif d'évaluer les performances d'une méthode employée dans un objectif pratique très précis. Il s'agit par exemple de démontrer la capacité d'une méthode pour recalibrer des images TEP et IRM du coeur avec une résolution de l'ordre de celle de la taille du voxel des données TEP. La validation est donc un pré-requis avant toute utilisation clinique. Le terme *évaluation* recouvre l'ensemble des concepts généraux ayant trait à l'analyse des performances des méthodes et inclut la *validation*. Nous présentons les différents critères de performances généralement considérés.

Une stratégie d'évaluation requiert une *référence*. Une *référence absolue* (*gold standard* ou *ground truth* en anglais) peut être définie comme un jeu de données test parfaitement décrit associé à des métriques d'erreur qui permettront de quantifier de manière absolue les performances d'une méthode. Des *références absolues*, qui représentent l'idéal dans un but d'évaluation, peuvent exister dans diverses situations, comme en contrôle non-destructif. Dans le domaine de l'imagerie médicale, une telle référence n'existe généralement pas car on n'a pas accès à la *vérité terrain*. On est alors amené à considérer des pseudo-références,

issues de compromis, qui ne reproduisent que partialement la réalité. On parlera alors de *référence approchée* qui est l'équivalent de *bronze standard* (ou *fuzzy standard*) en anglais. On peut attribuer des qualificatifs différents pour désigner la proximité de cette référence vis à vis de la *référence absolue*.

## 2.1 Cadre conceptuel de l'évaluation

### Comment évaluer

Bowyer et Phillips [Bowyer 98] distinguent trois cadres d'évaluation. Dans le cas de l'**évaluation indépendante**, une équipe constitue un ensemble de données test, conçoit le protocole d'évaluation, donne les images à des opérateurs indépendants et analyse les résultats de l'évaluation. Cette approche permet d'atteindre un niveau élevé de standardisation de l'évaluation car les algorithmes sont testés avec les mêmes images et évaluées selon la même méthode. Les évaluations indépendantes permettent en général une comparaison directe des méthodes qui traitent le même problème. Les algorithmes évalués sont contrôlés par leur concepteur ce qui garantit leur optimalité. L'évaluation indépendante faite par un utilisateur quelconque contribue à l'impartialité et l'objectivité des résultats. L'inconvénient principal de cette forme d'évaluation est l'effort continu que doit assurer l'équipe qui mène le projet d'évaluation.

La deuxième catégorie concerne les évaluations par un **groupe unique**. Souvent, le code exécutable des algorithmes n'est pas disponible. Il est alors nécessaire de recoder l'algorithme à partir des publications. Cela introduit la possibilité que la version évaluée ne soit pas tout à fait identique à celle des concepteurs.

Une caractéristique importante des deux premiers cadres d'évaluation concerne l'accès à une *référence absolue*, déterminée avec précision.

Dans la troisième catégorie, la *référence absolue* est difficile à obtenir. Une grande partie du processus d'évaluation est dédiée à la construction d'une *référence approchée*.

Selon Lehman, la plupart des méthodes d'évaluation de traitements d'images médicales proposées peuvent être classifiées en trois groupes [Lehmann 02] : l'**approche analytique** analyse l'algorithme seulement d'une manière théorique et ne nécessite pas son implantation ; les techniques de la seconde classe, désignées par **qualité empirique** ("empirical goodness", en anglais), sont basées sur des propriétés souhaitées souvent évaluées par un expert humain et n'utilisent pas de référence ; enfin, la troisième classe de méthodes, qualifiées de **différence empirique** ("empirical discrepancy", en anglais), s'appuie sur des mesures de distance à une référence souvent appelée *gold standard*. L'importance de la référence pour l'évaluation est soulignée par de nombreux auteurs [Nguyen 00, Bowyer 98, Jannin 01, Lehmann 02].

Une autre classification des procédures d'évaluation des performances d'un algorithme distingue les méthodes **contextuelles** et **non-contextuelles** [Nguyen 00]. Les méthodes **contextuelles** sont utilisées pour une application particulière, tandis que les méthodes **non-contextuelles** sont menées indépendamment de toute application. La plupart des

approches d'évaluation quantitatives sont **non-contextuelles**. Pour l'analyse des performances d'un algorithme de détection de contours, par exemple, les méthodes **non-contextuelles** consistent en l'évaluation de l'algorithme en changeant différentes caractéristiques du détecteur ou des images. Cette évaluation est directement liée à l'algorithme. Une méthode **contextuelle** estime les performances de l'algorithme dans un contexte spécifique comme la reconstruction d'objets à partir de contours préalablement détectés [Nguyen 00].

Les algorithmes de traitement d'images médicales, qui sont généralement développés dans un objectif médical et clinique, devraient être évalués par des méthodes **contextuelles de différence empirique** c'est à dire intégrant une *référence* [Lehmann 02]. Pour prendre en compte la grande variabilité des images médicales, une évaluation sérieuse des algorithmes de traitement d'images doit être basée sur un ensemble significatif d'images *de référence* [Courtney 01, Bowyer 98, Gee 00, Lehmann 02].

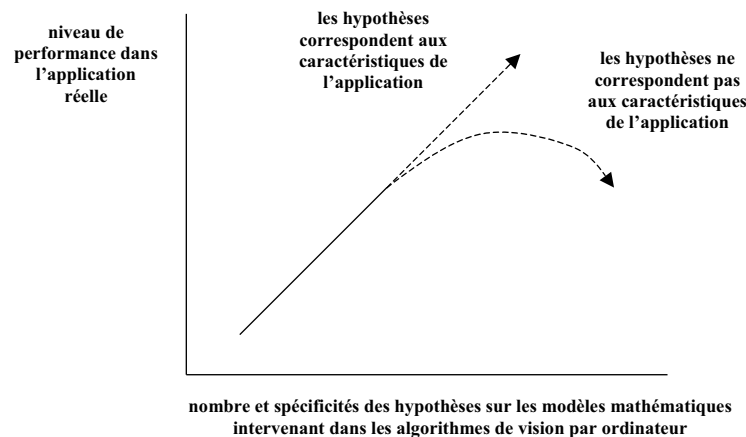


FIG. 2.1: Performance d'un algorithme en fonction de sa complexité théorique (d'après [Bowyer 98])

### Quoi évaluer

L'efficacité des systèmes d'imagerie diagnostique est évaluée selon 6 niveaux principaux [Fryback 91] :

- 1) la performance technique qui traduit la pertinence des résultats obtenus,
- 2) la précision diagnostique qui prouve l'efficacité contextuelle de la méthode pour établir un diagnostic d'une manière précise,
- 3) l'impact sur un diagnostic (amélioration du diagnostic),
- 4) l'impact thérapeutique (influence de la technique dans le choix et la réalisation

du traitement),

5) le bénéfice pour le patient (amélioration de la santé du patient),

6) l'impact social (rapport coût-efficacité).

Une évaluation traite en général un niveau mais, une étude complète devrait considérer tous les niveaux. A l'heure actuelle, il n'existe pas de procédure standardisée pour l'évaluation des méthodes de traitement d'images médicales [Jannin 02]. Il est ainsi difficile d'évaluer objectivement les performances d'une méthode et plus encore de comparer des méthodes entre elles.

## 2.2 Procédures d'évaluation

Les principales conditions pour l'évaluation sont : la standardisation de la méthodologie d'évaluation, la conception des données d'évaluation et les métriques d'évaluation [Loew 97, Shtern 99, Shahidi 01, Buvat 99, Yoo 00, Jannin 02].

### 2.2.1 Standardisation des procédures d'évaluation

Les méthodologies d'évaluation actuelles manquent de standardisation. La standardisation des procédures d'évaluation peut contribuer à la dissémination de bonnes pratiques et à une homogénéisation des résultats propices à une évaluation objective et comparative des méthodes. Dans le domaine médical d'autant plus, le passage d'un développement de la recherche à une utilisation en routine clinique demande, au moins une évaluation rigoureuse sur des populations significatives de patients, au plus une certification d'un organisme accrédité comme la FDA (**F**ood and **D**rug **A**dministration) aux États-Unis. Il est probable qu'à l'avenir ces organismes imposent eux-mêmes des critères de qualité et les seuils associés.

La standardisation de la méthodologie d'évaluation peut être facilitée par la caractérisation dans un cadre commun des méthodes d'analyse d'images, des contextes cliniques et des procédures d'évaluation.

1. *Caractérisation des méthodes de traitement d'images* : la description d'une méthode d'une manière générique et standardisée, à partir de ses principales composantes.

2. *Contexte clinique d'évaluation* : Les deux premiers stades de la *HCTA* (Health Care Technology Assessment<sup>2</sup> : organisme pour la promotion du développement et de la diffusion des technologies du domaine de la santé en tenant compte des impacts au niveau

---

<sup>2</sup>[http ://www.nlm.nih.gov/nichsr/ta101/ta10103.htm](http://www.nlm.nih.gov/nichsr/ta101/ta10103.htm)

social, éthique, légal ou politique) consistent en la définition du contexte clinique d'évaluation et en la spécification de l'objectif d'évaluation [Goodman 98]. Cet objectif peut être formulé sous la forme d'une hypothèse que l'évaluation confirmera ou non. L'hypothèse d'évaluation est définie en fonction du contexte clinique, d'un niveau d'évaluation (voir classification précédente), des caractéristiques des données impliquées (ex. : modalités d'imagerie, résolutions spatiales, dimensions), et par la métrique d'évaluation.

*3. Procédures d'évaluation standardisées :* Le besoin des protocoles d'évaluation standardisés se traduit par la nécessité de la définition d'une "terminologie unique standardisée d'évaluation" [Bowyer 98, Buvat 99]. La conception des modèles d'évaluation contribue à cette standardisation [Yoo 00].

Un cadre méthodologique intéressant pour l'évaluation rigoureuse des méthodes de traitements d'images est le Modèle d'Evaluation Générique, (*Generic Evaluation Model*), proposé dans [Buvat 99]. Le Modèle d'Evaluation Générique liste les composantes d'un protocole d'évaluation et les relations existantes entre ces composantes. Il attire l'attention sur des biais potentiels qui pourraient affecter la validité de l'étude.

La procédure est la suivante (figure 2.2) : en considérant la *méthode* à évaluer et le niveau d'évaluation, on formule l'objectif général de l'évaluation (abstract aim). Il est formulé sous la forme d'une question à laquelle on cherche à répondre. L'objectif général ne peut être analysé que dans un *contexte* particulier. La projection dans le contexte de l'objectif abstrait donne une *hypothèse concrète*. L'*hypothèse* détermine les *données d'entrée* sur lesquelles la méthode sera appliquée. Les *données de sortie* sont déterminées par les données d'entrée et la méthode. La *tâche* précise l'information que l'on cherche à extraire des images. De façon générale, on distingue deux types de tâches : les tâches de détection et les tâches d'estimation. Les tâches de détection sont les tâches qui appellent généralement une réponse binaire, par exemple la détection de la présence ou de l'absence d'une tumeur. Les tâches d'estimation sont celles qui conduisent à l'estimation d'un paramètre (ex : des paramètres cardiaques comme le flux sanguin, le volume ventriculaire, la fraction d'éjection) [Hoppin 02]. La *tâche* est effectuée par un *observateur* qui peut être numérique (algorithme) ou humain (médecin).

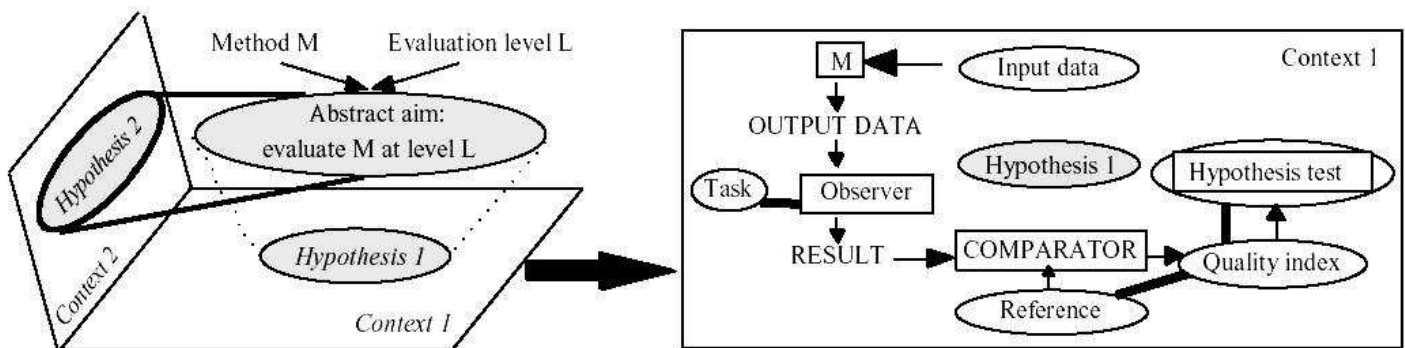


FIG. 2.2: Le modèle générique d'évaluation GEM (d'après [Buvat 99])

A partir de la référence et des résultats produits par la méthode, un *comparateur* fournit des *indices de qualité* ou des figures de mérite, qui quantifient l'écart par rapport à la référence. La dernière étape est le *test de l'hypothèse*. Il utilise les indices de qualité et détermine si l'hypothèse est acceptée ou rejetée à l'aide d'analyses statistiques.

La tâche détermine les techniques d'évaluation à considérer :

- pour les tâches de détection, il s'agit de la famille des méthodes ROC (Receiver Operating Characteristics).
- pour les tâches d'estimation, il s'agit de calculs de biais et de variabilité, basés sur l'écart entre le paramètre estimé et la valeur vraie du paramètre.

Si une *référence absolue* existe, l'analyse de corrélation des biais est très souvent utilisée. En l'absence de *référence absolue*, comme c'est généralement le cas en imagerie médicale, la vérification de l'hypothèse est plus ardue. Des approches statistiques ont été investiguées afin de permettre une évaluation de la robustesse [Hoppin 02, Bland 86]. Pour le recalage inter-modalités par exemple, où la *référence absolue* est absente, les auteurs évaluent souvent les performances de leurs algorithmes en faisant une étude de consistance [Holden 00, Roche 01, Nicolau 03, Pennec 98] : effectuer une série de recalages qui représente une succession des compositions des transformations qui conduisent à une transformation globale qui, en théorie, est la transformation identité. Dans [Jenkinson 01], plusieurs transformations test  $A_n$  ont été appliquées sur une image initiale  $I$ . Les transformations issues des recalages entre les images transformée et l'image de référence  $I_r$  sont notées  $T_n$ . Si la méthode de recalage est consistante, les transformations composées  $T_n \cdot A_n$  doivent être égales à  $T_0$ , transformation qui met en correspondance  $I$  et  $I_r$ . La figure 2.3 illustre ce type d'étude de consistance. La comparaison croisée des différents critères de recalage est employée dans [Hellier 03]. Pennec *et al.* proposent une méthode statistique d'évaluation de l'incertitude sur le recalage basé sur les points ou les repères ("frames", en anglais) [Pennec 98]. Cette méthode fait appel à la matrice de covariance des transformations géométriques.

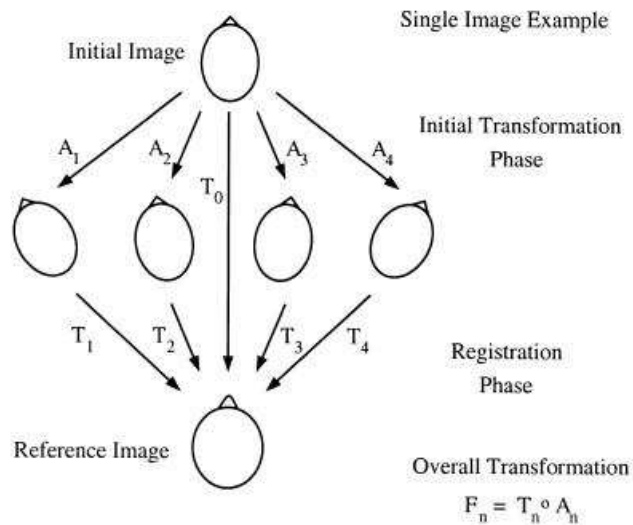


FIG. 2.3: Étude de consistance (d'après [Jenkinson 01])



### 2.2.2 Données d'évaluation

La construction, la caractérisation des données d'évaluation et leur dissémination dans plusieurs groupes de travail constituent une étape importante d'une stratégie d'évaluation. Elles peuvent faire l'objet d'une diffusion dans la communauté scientifique. La base de données test du groupe de recherche de l'Université de Vanderbilt<sup>3</sup>, États-Unis, constitue un excellent exemple dans un contexte d'évaluation de méthodes de recalage en imagerie cérébrale [West 97].

Le terme "gold standard" est très utilisé dans les publications médicales et désigne généralement un système de référence qui permet de valider une méthode. En réalité, la référence idéale, que nous appelons absolue, et qui correspond à l'idéal dans le contexte donné, est très rarement accessible dans le domaine médical. On lui substitue alors une référence plus ou moins approchée. Ce qui explique l'existence d'autres termes comme des références "silver", "bronze" ou "plastique" qui établissent une gradation dans la qualité des références [Lehmann 02, Yoo 00]. Ces données sont caractérisées par le degré de proximité vis à vis de la *référence absolue* pour le cadre applicatif considéré. En analyse d'images médicales, Lehmann propose de classer les références selon les propriétés suivantes [Lehmann 02] :

1. *Reproductibilité* ("reliance", en anglais) : La génération des images tests doit suivre un protocole déterminé et reproductible. Il ne doit pas reposer sur des étapes interactives ou sur l'intervention d'un opérateur.
2. *Réalisme* ("equivalence", en anglais) : les structures, le bruit et d'autres paramètres d'importance dans une référence numérique doivent respecter les caractéristiques des données réelles.
3. *Indépendance* : la construction de la référence doit s'appuyer sur une procédure ou des outils différents de ceux qu'on souhaite évaluer.  
Si une référence possède ces trois propriétés, elle est considérée comme une *référence absolue* ("gold standard"). Une référence qui ne présente que deux des trois, la reproductibilité étant l'une des deux, est qualifiée de "silver standard". Les autres références sont parfois qualifiées de "plastic" et ne devraient pas être utilisées pour l'évaluation. Cette nomenclature est en accord avec les quelques références bibliographiques traitant de l'évaluation en analyse d'images en imagerie médicale. En plus du choix approprié de la référence, deux autres propriétés doivent être satisfaites afin de réaliser une évaluation rigoureuse.
4. *Pertinence* ("relevance", en anglais) : l'algorithme évalué doit être stable en fonction de ses paramètres. Une mesure quantitative extraite d'une image (ex. la taille d'un objet) ne doit pas changer si les paramètres lors de l'acquisition qui structurellement n'affectent pas cette mesure changent (ex. la position de l'objet) et si les conditions biomédicales restent les mêmes (on image le même objet).
5. *Validité* ("significance", en anglais) : afin d'effectuer une analyse statistique signifi-

---

<sup>3</sup>[http ://www.vuse.vanderbilt.edu/ image/registration/](http://www.vuse.vanderbilt.edu/image/registration/)

cative et assurer une généralisation convenable de l'étude, le nombre des *références absolues* doit être suffisamment grand. Ainsi, la variance des résultats obtenus sur l'ensemble des références dans un test de validation croisée doit être faible afin d'assurer la généralité des résultats.

Pour que les conclusions de l'étude d'évaluation soient significatives, il est important d'associer un niveau de confiance à la référence utilisée. Les méthodes utilisées pour comparer les résultats avec la référence devraient tenir compte de l'incertitude liée à la référence. Phelps *et al.* ont fait appel aux méthodes Monte Carlo pour analyser les conséquences sur la précision d'un diagnostic avec les courbes ROC quand on utilise une *référence approchée* [Phelps 95] : si la référence n'est pas absolue, les courbes ROC obtenues représentent une estimation biaisée des courbes ROC réelles.

On peut aussi classer les méthodes d'évaluation en fonction des données d'évaluation utilisées [Jannin 01, Maintz 98] :

1. les **simulations** permettent théoriquement de contrôler tout type de paramètre qui pourrait influencer les performances de la méthode de recalage (comme la transformation géométrique ou le niveau de bruit). La transformation géométrique de référence est directement accessible car elle est intrinsèque à la construction des données. Dans cette catégorie, on trouve deux types de simulations. Les **simulations numériques** sont basées sur des données synthétiques. Elles peuvent être cependant éloignées de la réalité clinique. Les **simulations réalistes à partir des données cliniques** : on utilise des fantômes numériques issus de pré-traitements d'images réelles acquises avec une technique d'imagerie (ex. segmentation d'images par RM) pour simuler une autre modalité afin de créer une référence. Le réalisme des données simulées est difficile à évaluer et, comme les fantômes physiques, elles peuvent difficilement prendre en compte la variété des situations cliniques mais sont utiles quand la *vérité terrain* est difficile à obtenir, voire impossible, comme c'est le cas de la plupart des applications de recalage d'images médicales. Dans le contexte de l'évaluation, les simulations présentent des limitations qui peuvent sous-estimer ou surestimer la précision de la méthode à évaluer. Ces limitations proviennent du fait qu'elles sont basées sur des modèles de la réalité et non sur la réalité elle-même. De plus, il peut exister des dépendances fonctionnelles entre les modèles utilisés pour simuler les données et les modèles (hypothèses) intégrés dans la méthode de traitement d'images [Woods 00a].

Les simulations réalistes utilisent des fantômes numériques pour évaluer ou améliorer les systèmes et techniques d'imagerie médicale. La modélisation précise de l'anatomie humaine et de la technique d'imagerie constitue un excellent outil pour l'analyse des effets physiologiques, physiques et des instruments dans la technique analysée. Les fantômes numériques actuels peuvent être regroupés en deux classes : les **fantômes pixeliques** et les **fantômes géométriques** [Segars 01]. Les **fantômes pixeliques** [Hoffman 90, Zubal 94] sont basés sur des données réelles de patients et donc limités à une anatomie et à des résolutions particulières. L'introduction des variations anatomiques est très limitée et le changement de la résolution à travers

l'interpolation peut introduire des erreurs. Les structures du fantôme ne sont pas définies d'une manière analytique. Les **fantômes géométriques** [Wang 92, Tsui 93, Pretorius 99, Bradely 97] sont basés sur des primitives géométriques qui permettent des variations sur les formes et la génération de multiples résolutions. Ils sont définies analytiquement. Ces fantômes permettent l'accès à des propriétés importantes de l'anatomie (ex : le volume d'un organe). Ils peuvent être assez proches de la réalité, mais les équations mathématiques limitent une modélisation précise des formes des différents organes. Si la complexité des équations qui définissent les différentes structures anatomiques augmente afin d'obtenir des fantômes plus réalistes, la flexibilité du fantôme diminue.

Un fantôme numérique idéal devrait avoir le réalisme des **fantômes pixeliques** et la flexibilité des **fantômes géométriques**. Un exemple d'un tel fantôme est le 4D Mathematical Cardiac-Torso phantom (MCAT) développé à l'Université North Carolina ([Tsui 93, Pretorius 97]).

2. Les **fantômes physiques** sont des modèles anthropomorphiques plus ou moins réalistes, utilisés lorsque le contexte de l'évaluation concerne les aspects liés notamment à la physique d'acquisition. Ici, la *vérité terrain* est intrinsèque au protocole d'expérimentation (elle peut être estimée grâce, par exemple, à la connaissance des caractéristiques géométriques du fantôme ou à la présence de marqueurs). Ces fantômes, bien adaptés à l'évaluation des caractéristiques des systèmes d'acquisition, sont cependant éloignés de la réalité clinique et anatomique. Les méthodes d'évaluation les plus populaires utilisent les fantômes physiques avec des marqueurs extrinsèques ou intrinsèques (repères anatomiques) [Maintz 98].
3. Les **données cliniques**. Lorsque le contexte d'évaluation est lié à l'utilisation clinique des outils, il est préférable d'utiliser des données réalistes, c'est à dire tenant compte des spécificités du contexte clinique, comme la présence de pathologies ou d'artefacts, ou de bougés lors de l'acquisition. Les données cliniques peuvent être classées en deux catégories : les données *prospectives* et *rétrospectives*. Les données *prospectives* sont des données cliniques qui ont été spécialement acquises à des fins d'évaluation. Dans le cadre du recalage d'images, différents systèmes, invasifs ou non-invasifs, sont envisageables pour contrôler la position du sujet dans l'imageur au moment de l'acquisition. La *vérité terrain* est estimée à partir de protocoles dédiés (*e.g.* marqueurs externes invasifs). Dans cette catégorie se trouvent également les données de *référence* qui sont archivées dans des bases des données accessibles via le web pour être utilisées dans différents projets d'évaluation (ex. : Le projet de l'Université Vanderbilt, USA, le projet VHP-Visible Human Project de la National Library of Medicine). Elles permettent de comparer les résultats des procédures d'évaluation.

Les données *rétrospectives* sont des données cliniques pour lesquelles aucun système de contrôle de la position n'a été utilisé lors de l'acquisition. De telles données sont utilisées par les procédures d'évaluation sans *référence absolue*.

4. Les **données post-mortem** sont des données d'évaluation réalistes mais limitées à

l'anatomie car il n'y a pas d'information métabolique ou fonctionnelle. La *vérité terrain* peut être estimée en utilisant des marqueurs invasifs [Hemler 96].

Ces différents types de données d'évaluation sont complémentaires et étudient des facettes différentes d'une méthode ou d'un système. Une évaluation complète de la performance devrait donc être menée en utilisant différents types de données, si possible. En pratique, on est généralement conduit à faire un compromis entre la précision et le réalisme de la référence [Loew 97], les propriétés 1-3 décrites précédemment [Lehmann 02] étant rarement toutes respectées.

Le partage des bases d'images pour des études multicentriques facilite le processus d'évaluation et la comparaison des performances et permet de mener des études de robustesse. Pour le recalage, par exemple, la construction des bases de données de référence peut être réalisée à partir d'ensembles d'images de référence plus ou moins réalistes et précises comme [Loew 97] :

- des données cliniques recalées, prospectives,
- des données *post-mortem* avec des marqueurs insérés,
- des données issues de simulations physiques

Ces bases de données de référence doivent être régulièrement mises à jour compte tenu des nouveaux protocoles et techniques d'imagerie médicale, et de nouvelles applications. Elles doivent inclure des cas difficiles ou rares.

Comme la validation clinique requiert des images cliniques spécifiques aux conditions des institutions cliniques, la mise en oeuvre des validations cliniques reste difficile en l'absence de véritable standardisation des systèmes d'acquisition et des images qu'ils produisent [Yoo 00].

Dans le tableau 2.1, on propose une classification de données d'évaluation, avec leurs avantages et inconvénients respectifs.

### 2.2.3 Métriques d'évaluation

Dans le cadre d'une étude d'évaluation, les résultats d'une méthode sont comparés à la vérité terrain en utilisant des *métriques d'évaluation* choisies afin de valider le critère sélectionné. Elle doivent être des indicateurs cliniques utiles pour les résultats [Yoo 00]. Pour une évaluation globale d'un système de chirurgie assisté par image, il faut estimer les erreurs à chaque niveau du système et prendre en compte leur propagation à travers l'ensemble du processus [Loew 97, Shtern 99, Shahidi 01, Viant 01].

Dans le contexte du recalage des images médicales, la précision est mesurée généralement dans une région cible (*target registration area*, en anglais) [Fitzpatrick 00]. L'erreur quadratique moyenne est la mesure la plus utilisée [Hajnal 01, Fitzpatrick 00, Woods 00a]. Dans le cadre du recalage rigide, les erreurs sur les paramètres de la transformation spatiale, comme les erreurs en rotation et translation suivant les axes de coordonnées sont souvent rapportées [Grova 01]. Cependant, la décomposition d'un mouvement rigide en rotations et translations, n'est pas unique et le résultat dépend de l'ordre d'application

	Avantages	Inconvénients
Simulations numériques idéalistes	<ul style="list-style-type: none"> <li>★ connaissance parfaite de la VT</li> <li>★ réglage des paramètres</li> </ul>	<ul style="list-style-type: none"> <li>★ réalisme</li> <li>★ difficulté d'identifier et de modéliser toutes les sources de variabilité</li> <li>★ il peut exister des dépendances fonctionnelles entre les modèles utilisés pour simuler les données et les modèles (ou hypothèses) sous-jacents à la méthode de traitement d'images étudiée</li> </ul>
Simulations réalistes à partir de données cliniques	<ul style="list-style-type: none"> <li>★ idem que pour les simulations numériques</li> <li>★ meilleur réalisme</li> </ul>	<ul style="list-style-type: none"> <li>★ idem que pour les simulations numériques</li> </ul>
Fantômes physiques	<ul style="list-style-type: none"> <li>★ les conditions réelles d'acquisition d'images sont prises en compte</li> </ul>	<ul style="list-style-type: none"> <li>★ peu de fantômes sont multimodaux ou simulent différentes propriétés physiques</li> <li>★ VT calculée à partir de protocoles dédiés (ex. : marqueurs)</li> </ul>
Données post-mortem	<ul style="list-style-type: none"> <li>★ données réalistes anatomiquement</li> </ul>	<ul style="list-style-type: none"> <li>★ VT estimée à partir de protocoles dédiés</li> <li>★ pas d'information métabolique ou fonctionnelle</li> </ul>
Données de référence partagées	<ul style="list-style-type: none"> <li>★ permettent de comparer les résultats des procédures d'évaluation</li> </ul>	<ul style="list-style-type: none"> <li>★ doivent inclure des cas difficiles ou rares (cas pathologiques)</li> <li>★ devraient posséder une information liée au patient et à sa pathologie</li> <li>★ les bases des données accessibles via web doivent être régulièrement mises à jour</li> </ul>
Données cliniques	<ul style="list-style-type: none"> <li>★ réalisme optimal</li> </ul>	<ul style="list-style-type: none"> <li>★ VT estimée à partir des protocoles dédiés dans le cas des données prospectives ou pas de VT dans le cas de données rétrospectives provenant de la routine clinique</li> </ul>

TAB. 2.1: Données d'évaluation : avantages et inconvénients (VT=*vérité terrain*) (adapté de [Jannin P., Tutorial MICCAI'03])

des transformations élémentaires [Woods 00b].

## 2.3 Évaluation des méthodes de recalage

### 2.3.1 Critères de qualité du recalage

L'évaluation d'une méthode de traitement d'images, et *a fortiori* des méthodes de recalage d'images, porte sur les caractéristiques techniques suivantes [Maintz 98] :

- résolution
- précision
- robustesse/stabilité
- ressources matérielles et humaines
- complexité de l'algorithme et temps de calcul

La *résolution* (“precision”, en anglais) d'un processus est liée à la valeur minimale de l'erreur systématique que l'on peut attendre de ce processus. Cette valeur est généralement exprimée dans l'espace des paramètres. La résolution d'une méthode de recalage est liée à la variation minimale, suivant les paramètres de la transformation, que la méthode est capable de détecter. En appliquant une variation spécifique en entrée de la procédure à valider, on cherche à vérifier la capacité du système à la détecter.

La *précision* (“accuracy”, en anglais) locale, définie pour chaque voxel d'un jeu de données volumiques, est la différence entre ses coordonnées calculées et ses coordonnées théoriques considérées connues à partir d'une vérité terrain. Une valeur globale de la précision d'une méthode peut être calculée, pour la totalité d'un jeu de données, à partir d'une combinaison de valeurs de la précision locale. La précision est le paramètre qui intéresse le plus le clinicien : le chirurgien, par exemple, qui, pour prendre une décision d'une intervention dans une région visualisée sur un écran de contrôle, voudrait savoir quelle est la précision de la position de l'instrument relativement à une région anatomique. L'estimation de la précision peut être **qualitative** et **quantitative**. Pour l'estimation **quantitative** de la précision, il est indispensable d'avoir accès à la *vérité terrain*. L'estimation **qualitative** peut être accomplie par une simple inspection visuelle qui représente une façon triviale de vérifier les résultats des algorithmes de recalage. Un opérateur expert regarde les images recalées à l'aide de différents outils de visualisation et juge le recalage comme un “échec” ou un “succès” si l'erreur estimée se situe au dessus ou en dessous du seuil de tolérance requis par l'application. Des outils de visualisation ont été développés afin d'afficher des données multimodales anatomo-fonctionnelles dans un contexte de fusion, par exemple, des cartographies [Behloul 01], [Nekolla 00] ou des représentations en damier (c'est à dire représentant en alternance une partie de chaque image dans le même repère) [Pauna 03], [Mäkelä 03b]. Ceci est efficace si dans la plupart des cas l'algorithme recalcule avec une erreur acceptable pour l'application et s'il n'échoue qu'occasionnellement. Mais, pour un algorithme qui retourne souvent des erreurs de recalage qui dépassent la précision demandée par l'application, l'opérateur risque de générer trop de faux négatifs (de cas où les images sont suffisamment bien recalées mais qui sont considérées comme

des échecs) ou des faux positifs (des cas où les images ne sont pas recalées d’une façon satisfaisante mais où le recalage est considéré comme un succès) [Hill 00].

Le contrôle visuel peut être réalisé par comparaison de la localisation de points anatomiques connus et identifiables dans les différentes modalités d’imagerie utilisées, en affectant un score à tout point étudié (mise en correspondance parfaite, correcte, passable, mauvaise) [Gallippi 01, Slomka 95, Faber 91, Sturm 03]. Comme l’identification de points anatomiques communs est assez difficile et fastidieuse, une autre approche consiste à segmenter différentes structures des deux images recalées et de les superposer. Des travaux dans le contexte du recalage d’images par RM de cerveau ont montré que les observateurs ont la capacité de détecter visuellement des erreurs de recalage supérieures à 0.2 mm en regardant des images de différence [Holden 00].

La *robustesse* : une méthode est dite robuste si elle fournit des résultats corrects et de façon constante. En d’autres termes, la *robustesse* ou la stabilité signifie que de petites variations en entrée entraînent de faibles variations en sortie d’un algorithme [Maintz 98]. La robustesse peut être vérifiée par l’étude du comportement à la suite de l’introduction dans les données de facteurs perturbateurs (bruit ou données pathologiques comme la présence d’une lésion) qui peuvent, notamment dans le cas du recalage, gêner la convergence. La robustesse peut aussi être étudiée en exécutant la méthode sur un nombre important de données ou des données d’origines diverses qui entrent dans le contexte clinique défini.

La *fiabilité* (“reliability”, en anglais) est la propriété qui confère à un algorithme un comportement prévisible pour une gamme diversifiée de données cliniques en entrée.

Les *ressources* concernent les matériels et l’effort humain investis pour l’utilisation de l’algorithme. L’investissement consenti doit être compensé par l’apport de la méthode dans le contexte clinique [Maintz 98].

La *complexité* et le *temps de calcul* sont des caractéristiques d’une méthode et de son implantation. Leurs impacts seront différents selon les contextes cliniques.

Dans l’idéal, tous ces critères doivent être satisfaits. Mais, en réalité, on est généralement amené à réaliser un compromis et l’impact respectif dépend directement de l’application clinique. Finalement, il s’agit d’évaluer l’utilité clinique de la méthode : est-ce que la méthode apporte un bénéfice à la pratique clinique et peut-elle être considérée supérieure à ses alternatives ? [Maintz 98].

### 2.3.2 Évaluation de la précision des méthodes de recalage

Du point de vue de l’utilisateur, la précision d’un algorithme de mise en correspondance est une des propriétés les plus importantes. On peut énoncer deux raisons principales pour quantifier la précision. La première est de savoir si une méthode de recalage est appropriée pour une application clinique particulière, ou pour comparer plusieurs méthodes entre elles. La deuxième raison est que la connaissance exacte de la précision d’une méthode peut avoir un impact sur la prise d’une décision clinique : une précision élevée peut

être vitale quand on intervient par voie chirurgicale [Holden 00, Woods 00a]. L'évaluation de la précision spatiale des algorithmes est une tâche difficile car la *vérité terrain* n'est généralement pas disponible [Hajnal 01, Fitzpatrick 00, Woods 00a]. Les méthodes de mise en correspondance sont en général évaluées en utilisant des marqueurs externes, des repères anatomiques ou des cadres de localisation externes pour avoir une *référence absolue* [Woods 00a]. Une comparaison directe des mesures d'erreur rapportées dans la littérature n'est pas aisée étant donné que la précision et les méthodes employées pour sa mesure sont différentes [Gilardi 98]. L'inspection visuelle est la méthode d'évaluation la plus évidente pour la précision mais elle est insuffisante.

Pour réduire les erreurs de recalage dues au mouvement cardiaque et à la respiration, l'acquisition est parfois réalisée en apnée avec une synchronisation à l'ECG. Pour le recalage d'images cardiaques, qui utilise des images de thorax comme intermédiaire, le mouvement de la cage thoracique et du diaphragme a des amplitudes différentes selon la région considérée. Les mouvements peuvent produire des artefacts de mouvement qui peuvent introduire des erreurs pendant le recalage [Dahlbom 00]. Les mouvements respiratoires dans le voisinage du diaphragme sont plus importants et de l'ordre de plusieurs centimètres [Goerres 02]. Les images cardiaques TEP et TEMP sont des images intégrées sur une période de temps plus ou moins longue. Cela constitue un problème supplémentaire quand on les recalc avec des images par RM synchronisées sur l'ECG. Dans le cas où les deux images à recaler sont intégrées dans le temps (cas du recalage TEP-TEP ou TEMP-TEMP), on peut considérer que les erreurs dues à la respiration sont similaires [Gilardi 98].

Dans le cas de l'utilisation d'une image intermédiaire (par ex. : image de transmission en TEP), on fait l'hypothèse de complète similarité des conditions d'acquisition [Kim 91], en particulier, cela signifie que le patient ne bouge pas entre les deux acquisitions, ce qui ne peut pas être rigoureusement assuré en pratique. Dans le cas de la TEP, Yu *et al.* affirment que l'alignement rigoureux à l'aide de laser entre deux acquisitions en TEP - d'émission et de transmission, est une méthode adéquate [Yu 94, Yu 95]. Quelques méthodes de recalage entre des images TEP de transmission et d'émission sont présentées dans [Bettinardi 93, Costa 95, Sechet 02]. Les images de transmission sont utilisées pour la correction en atténuation des images d'émission. Si le mouvement entre les images de transmission et d'émission est important (de l'ordre de 2-3 cm), il peut sérieusement affecter la correction d'atténuation de l'image d'émission et, ainsi, sa qualité [Stone 98]. Le deuxième intérêt des images de transmission est qu'elles fournissent davantage d'information sur les structures anatomiques que les images d'émission.

Les fantômes physiques sont plus appropriés pour estimer la précision en imagerie intra-modalité. Pour l'imagerie inter-modalités, qui implique l'utilisation de différents systèmes d'acquisition, on se heurte au même problème qu'avec des images réelles, il n'est pas possible de connaître avec précision la position spatiale du fantôme dans les deux imageurs (pour le contrôle du mouvement, l'orientation à l'aide de laser n'est pas précise). Les données *post-mortem* peuvent fournir une solution s'il ne s'agit pas d'étudier un processus physiologique [Woods 00a].



Introduits récemment, les imageurs intégrés TEP/CT cliniques, IRM/TEP expérimentaux (petit animal) [Slates 99] fournissent des données multimodales automatiquement recalées à 1 mm près (moins de 1 mm pour les études sur des fantômes) [Goerres 02]. Mais ceci ne résout pas le problème de mouvement des organes.

### 2.3.3 Stratégies d'évaluation des méthodes de recalage d'images cardiaques et thoraciques

Nous avons présenté dans le chapitre 1 le contexte clinique dans lequel s'intègre cette thèse. L'imagerie de coeur présente des spécificités et peu de méthodes de recalage ont été proposées et testées dans ce domaine. Il nous a semblé important d'élaborer une stratégie d'évaluation adaptée à ce cas en s'appuyant sur des travaux existant notamment en imagerie cérébrale.

Nous présentons ci-après une synthèse des procédures d'évaluation de méthodes de recalage d'images thoraciques et cardiaques. Pour cela, nous nous appuyons sur les classifications de [Jannin 02, Grova 02, Buvat 99] et l'article de synthèse [Mäkelä 02]. Nous distinguons les trois composantes principales d'une procédure d'évaluation : l'énoncé d'une hypothèse d'évaluation, le choix d'une métrique d'évaluation et la mise en oeuvre de la procédure d'évaluation. L'ensemble classé des références est synthétisé dans les tableaux 2.2 pour l'imagerie de thorax et 2.3 pour l'imagerie de coeur.

#### Hypothèse d'évaluation

Une hypothèse d'évaluation se formalise par la spécification d'un niveau d'évaluation, des caractéristiques des données mises en jeu (modalité, dimensions, résolutions) et d'un contexte clinique. Confirmer ou non une hypothèse d'évaluation se formalise par un test statistique d'hypothèse, nécessitant le recours à une figure de mérite.

Les références bibliographiques en matière de recalage d'images de thorax et de coeur que nous avons étudiées (voir tableaux 2.2 et 2.3) n'adoptent pas une telle démarche, mais étudient plusieurs aspects des méthodes de recalage (précision, robustesse, etc). Lorsqu'un objectif d'évaluation est clairement défini, elles n'y répondent généralement pas à l'aide de tests statistiques d'hypothèse. Dans chaque cas, nous précisons :

- le *niveau d'évaluation* : Un niveau d'évaluation technique (T) correspond à une phase de réglage des paramètres de la méthode, afin d'obtenir des performances optimales. Un niveau d'évaluation clinique (C) étudie les performances d'une méthode dans une application clinique précise, par exemple l'étude de l'influence des zones ischémiques sur la précision du recalage.
- le nombre de techniques d'imagerie employées : cas monomodal ou multimodal.

#### Métriques d'évaluation

L'évaluation quantitative d'une méthode de recalage est obtenue par l'estimation de l'écart entre la transformation géométrique délivrée par la méthode de recalage à évaluer

et la transformation géométrique de référence. Lorsque la transformation géométrique de référence correspond à une *vérité terrain*, alors la mesure de divergence revient à une mesure d'erreur (E) de recalage. Par contre, lorsque la transformation géométrique de référence ne correspond pas à une *vérité terrain* mais, par exemple, au résultat d'une autre méthode de recalage ou à une transformation moyenne, Woods *et al.* introduisent la notion de mesure d'inconsistance (I) [Woods 98]. Considérons trois images  $A$ ,  $B$  et  $C$ . Une méthode de recalage est dite inconsistante si elle ne fournit pas les mêmes résultats quand on recalc directement le couple d'images  $(A, C)$  ou, successivement, les couples  $(A, B)$  et  $(B, C)$ .

### Mise en oeuvre d'une procédure d'évaluation

La mise en oeuvre d'une procédure d'évaluation de méthodes de recalage nécessite de se doter de données d'évaluation et de définir les paramètres en entrée de la procédure. Ces données et paramètres sont utilisés à la fois par la méthode de recalage à évaluer et pour la génération d'une transformation géométrique de référence. Il est alors possible d'évaluer la méthode dans différentes configurations en faisant varier les paramètres. Si la transformation géométrique qui met en correspondance les données est connue, alors la précision de l'algorithme de recalage est estimée en comparant la transformation issue du recalage avec la solution connue, appelée transformation de référence.

### Paramètres de la procédure

Les paramètres en entrée d'une procédure d'évaluation peuvent concerner la qualité des données à recalcr comme par exemple la correction des distorsions géométriques, le recouvrement partiel des données, la segmentation préalable des différentes structures. Pour une même fonctionnelle de coût, ces paramètres concernent aussi l'ajustement des différentes composantes de la méthode de recalage comme par exemple l'algorithme d'optimisation, la stratégie multi-résolution, la méthode de sous-échantillonnage ou de lissage des données, la méthode d'interpolation ou le nombre d'opérateurs lorsqu'il s'agit d'une méthode de recalage semi-automatique.

### Type de référence de la procédure

L'évaluation quantitative des performances d'une méthode de recalage nécessite en général de comparer à une transformation géométrique de référence, la transformation géométrique estimée par la méthode de recalage en question. Ainsi, dans la colonne *Référence* des tableaux 2.2 et 2.3, nous distinguons trois catégories de références définies en fonction de la transformation géométrique. La première repose sur une transformation géométrique parfaitement connue ; on parle de *référence absolue* vis à vis de la transformation géométrique  $(A_p+)$ . Elle est généralement liée à l'utilisation de fantômes physiques ou de simulations numériques. Nous n'avons pas identifié dans la littérature des travaux faisant appel à une *référence absolue* selon les trois critères de Lehmann [Lehmann 02] :

reproductibilité, réalisme, indépendance. La seconde famille concerne l'utilisation de données prospectives acquises afin d'obtenir une transformation géométrique ; on a alors une *référence approchée* ( $A_p$ ). Dans la dernière catégorie d'études dites *sans référence* ( $A_u$ ), la transformation géométrique de référence peut être une transformation géométrique moyenne issue de divers essais de recalage ou une transformation géométrique issue d'une autre méthode de recalage.

On présente dans les tableaux 2.2 et 2.3 une synthèse des stratégies d'évaluation des méthodes de recalage d'images de thorax et de coeur, respectivement. Les notations suivantes sont utilisées :

- Niveau d'évaluation :  
 $C$ =niveau d'évaluation clinique  
 $T$ =niveau d'évaluation technique
- Référence :  
 $A_p$ +référence absolue vis à vis de la transformation géométrique  
 $A_p$ =référence approchée  
 $A_u$ =sans référence
- Type d'erreur :  
 $I$ =mesure d'erreur (la transformation géométrique de référence correspond à une *vérité terrain*)  
 $E$ =mesure d'inconsistance (la transformation géométrique de référence ne correspond pas à une *vérité terrain*)  
 $moy$ =moyenne des écarts absolus  
 $eqm$ =erreur quadratique moyenne calculée sur des points (pts)  
 $SDV$ =écart-type

### Données d'évaluation : images de thorax

- *fantômes physiques* : Des approches d'évaluation des méthodes de recalage d'images de thorax à l'aide de fantômes physiques rigides ont fait l'objet de plusieurs publications. Les plus utilisés sont les fantômes Alderson et Data Spectrum. Des marqueurs extrinsèques, adaptés aux applications cliniques, ont été spécialement conçus pour différents fantômes physiques. Dans le contexte de l'imagerie multimodale, les marqueurs externes doivent être visibles dans l'une et l'autre des modalités à recalcr.

Le fantôme de thorax proposé par Alderson *et al.* [Alderson 62] permet de tenir compte d'acquisitions réelles. Il a été employé pour l'évaluation du recalage d'images CT/TEP par Yu *et al.* [Yu 95], en utilisant des marqueurs extrinsèques en plastique pendant l'acquisition CT et des marqueurs métalliques pendant l'acquisition

en TEP de transmission. Pour l'image TEP d'émission, les compartiments du fantôme physique correspondant aux tissus mous ont été remplis de  $F^{18}$  [Yu 94]. Dans [Gilardi 98], neuf marqueurs remplis de  $Tl^{201}$  et de  $F^{18}$  ont été utilisés pour le recalage d'images TEMP et TEP. Dans le contexte du recalage d'images TEP, Pallotta *et al.* ont pris en compte les surfaces externes du thorax et les surfaces internes des poumons dans la procédure de recalage d'images [Pallotta 95]. Pour guider le recalage, sept marqueurs en plastique remplis de  $F^{18}$  ont été positionnés sur la surface externe du fantôme.

Le fantôme de thorax Data Spectrum<sup>4</sup>, intégrant poumons, coeur et colonne vertébrale, a été utilisé pour l'évaluation de méthodes de recalage monomodal ou multimodal. Dans l'approche d'évaluation décrite dans [Costa 95], le fantôme n'a pas bougé entre les acquisitions TEP de transmission et d'émission. Dans le contexte multimodal CT/TEP, Meyer *et al.* utilisent le fantôme associé à des marqueurs extrinsèques qui sont des capsules vides ou remplies de FDG, selon l'imageur utilisé [Meyer 96]. Dey *et al.* ont comparé la précision de trois méthodes différentes de recalage d'images TEMP/CT de thorax en utilisant l'image de transmission comme médiateur [Dey 99]. Neuf marqueurs ont été définis parmi les multiples vis en plastique du fantôme. Un modèle anthropomorphe de thorax (fantôme de thorax Jaszczak, Data Spectrum) avec des points de contrôle facilement identifiables en TEMP et en TEP a été employé dans [Eberl 96].

Le positionnement des fantômes physiques lors de plusieurs acquisitions est souvent établi à l'aide de lasers qui alignent des points désignés sur le fantôme physique [Cai 99, Bettinardi 93, Kramer 89].

- *simulations numériques* : Des simulations numériques ont été utilisées par Pallotta *et al.* [Pallotta 95]. Un thorax de synthèse a été obtenu par seuillage d'une image de transmission réelle. Trois valeurs différentes d'intensité ont été établies pour les poumons, le reste du thorax et les sept points de référence qui ont été choisis dans la proximité des surfaces externes du thorax.

Dans [Tai 97], un fantôme numérique et anatomique a été élaboré à partir de données réelles CT dans le cadre du projet VHP<sup>5</sup> (Visible Human Project) de la NLM (National Library of Medicine) [Ackerman 95].

Des simulations réalistes ont été employées pour l'évaluation des méthodes de recalage IRM/TEP dans [Pauna 03, Mäkelä 03b]. Un simulateur Monte Carlo a été utilisé pour simuler des acquisitions TEP à partir d'une image par RM réelle. Cette approche est décrite au chapitre 3.

- *données cliniques* : Les données prospectives [Pallotta 95, Yu 94] et rétrospectives

<sup>4</sup><http://www.cspmedicalstore.com/s.nl/c.ACCT88592/sc.2/category.86/.f>

<sup>5</sup>[http://www.nlm.nih.gov/research/visible/visible\\_human.html](http://www.nlm.nih.gov/research/visible/visible_human.html)

[Gilardi 98, Goerres 02, Bacharach 93, Cai 99] ont été aussi utilisées pour l'évaluation des méthodes de recalage.

### Données d'évaluation : images cardiaques

- *fantômes physiques* : Pour l'évaluation d'un algorithme de mise en correspondance d'images TEP, Turkington *et al.* [Turkington 97] ont utilisé un fantôme physique de coeur constitué de deux éléments concentriques en plastique de diamètre 7 et 9 cm, respectivement. Afin de simuler le débit sanguin, les éléments ont été remplis avec une solution de  $F^{18}$  et le fantôme a été imagé à plusieurs reprises avec des laps de temps différents pour obtenir des niveaux de bruit différents.
- *simulations numériques* : Klein *et al.* [Klein 02a] exploitent le fantôme mathématique de coeur 4D MCAT [Pretorius 97] pour évaluer un algorithme de correction de mouvement 4-D pour des images cardiaques TEP. Le fantôme MCAT est composé d'ellipsoïdes qui représentent les structures principales du thorax. Les différentes phases du cycle cardiaque ont été simulées en faisant varier les dimensions des surfaces des ellipsoïdes qui définissent l'épicarde et l'endocarde. Un simulateur numérique TEP Monte Carlo a été utilisé dans [Pauna 03] pour évaluer le recalage d'images par RM et TEP de coeur. L'image métabolique de captation glucidique est issue des images par RM réelles segmentées qui ont servi à construire les fantômes nécessaires au simulateur TEP (voir chapitre 3).
- *données cliniques* : Une autre approche d'évaluation des méthodes de recalage d'images cardiaques s'appuie sur des repères anatomiques. La localisation des repères en imagerie multimodale cardiaque est délicate, compte tenu de la complexité des structures anatomiques. Ceci peut induire des erreurs lors de l'estimation de la précision des méthodes de mise en correspondance basées sur des amers. Leur identification est rendue difficile par les distorsions géométriques dans les images par RM et par l'effet de volume partiel dans les images par RM et TEP. Les marqueurs externes placés sur la peau ne peuvent pas être utilisés car ils ne vont pas se déplacer avec le coeur et ils ne peuvent donc pas mettre en évidence l'évolution de la géométrie du coeur [Sinha 95].

Eberl *et al.* mettent en correspondance des images TEMP pour étudier la perfusion myocardique en choisissant 15 régions d'intérêt rectangulaires et centrées sur des repères anatomo-fonctionnels (ex. : l'apex) [Eberl 96]. Les extrémités (supérieure et inférieure) et le centre du septum constituent des repères sur les images par RM [Bidaut 01]. Sinha *et al.* appariant des repères anatomiques pendant le recalage pour compenser le mouvement dû à la respiration sur les images par RM de perfusion [Sinha 95]. L'évaluation quantitative du recalage peut être également réalisée en mesurant la distance entre les surfaces épicaudiques recalées comme dans

[Sturm 03] où des données multimodales IRM/TEP de 11 patients atteints de maladies cardio-vasculaires ont été analysées. Dans [Hoh 93], les données TEP de coeur sont préalablement recalées avec le même algorithme qui sera validé. Des transformations géométriques rigides sont ensuite appliquées et l'erreur du recalage est calculée comme la différence entre les paramètres des transformations géométriques avant et après le recalage.

## Conclusion

La référence constitue la composante clé d'une procédure d'évaluation. Elle peut être caractérisée en fonction de son réalisme (anatomique et physiologique) et du contrôle de son positionnement. La littérature montre qu'il est nécessaire de faire un compromis entre ces deux aspects et les études actuelles tentent de fournir une référence en termes à la fois de réalisme et de contrôle du positionnement. On peut cependant dire qu'une telle référence absolue, à la fois pleinement réaliste et parfaitement contrôlable dans son positionnement dans diverses modalités d'imagerie, n'existe pas à l'heure actuelle. C'est pourquoi dans les tableaux 2.2 et 2.3 les références absolues n'apparaissent pas. Nous avons identifié celles qui s'approchent le plus par  $A_p+$ . L'évaluation présentée dans [Pallotta 95], dont nous qualifions la référence  $A_p+$ , s'appuie également sur une simulation numérique d'image mais qui procède par filtrage. Elle est donc beaucoup moins réaliste au sens de la physique d'imagerie que la référence proposée dans cette thèse (voir Chapitre 3).

Les systèmes intégrés d'imagerie médicale, comme par exemple les scanners combinés TEP/CT [Beyer 99, Patton 00] pourraient fournir une *référence absolue* pour l'évaluation de recalage d'images [Goerres 02].

			Hyp. d'évaluation		Mise en oeuvre				Métrique d'évaluation			
modal.	Référence	Méthodes de recalage	Niveau d'éval.	Modalités d'imagerie	Référence	Données d'évaluation	$T_{ref}$	Paramètres	Type	Mesurée sur	Type d'erreur	Erreur
1 modalité	[Pallotta 95]	distance chamfrein	T	TEP(TR/EM)	$A_p+$	simul.num.	connue(aléat.)	lissage	E	pts(7)+params.	eqm,moy	2.19±0.52 mm
					$A_p$	fant.phys. Alderson	connue(aléat.)	bruit/lissage	E	pts(7)+params.	eqm	2.68-4.53 mm
			C	3 patients	$A_p$	prospectives	connue	atténuation	E	coups/pixel	corrél.	r=0.93
	[Bacharach 93]	CC	T	TEP/TEP	$A_u$	rétrospectives	connue	correct.atten.	E	params.	moy	1mm 1.5°
	[Yu 94]	head & hat	T/C	TEP(TR/EM)	$A_p$	fant.phys. Alderson	connue	seuillage	E	pts(9)	moy	(x,y)(2.35,1.81mm)
					$A_p$	retrospectives	Id	seuillage	I	params.	moy±SDV	(x)(2.28±2.06) (y)2.34±2.16mm (rot)1.32±1.42°
	[Tai 97]	surface	T	CT	$A_p$	fant. num.	connue	-	E	vect. depl.	min,max	0-4.5mm
	[Costa 95]	MV	T	TEP/TEP	$A_p$	fant.phys. DataSpectrum	connue	bruit statistique	E	params.	moy±SDV	(x)0.07±0.12mm (y)0.16±0.32mm (rot)0.01±0.56°
2 modalités	[Yu 95]	head & hat	T	TEP/CT	$A_p$	fant.phys. Alderson	connue	lissage	E	pts(9)	eqm	(x,y)2.3mm, (z)3mm
	[Meyer 96]	IM	T	TEP/CT	$A_p$	fant.phys. DataSpectrum	connue	-	E	pts(6)+params.	moy,eqm	3.84mm, 0.17°
	[Eberl 96]	SAD	T	TEP/TEMP	$A_p$	fant.phys.	connue	bruit/activité	I	pts(1 5)	moy±SDV	3.1±1.7mm
	[Gilardi 98]	distance chamfrein	T	TEMP/TEP	$A_p$	fant.phys. Alderson	connue	detect. surf.	E	pts(9)	moy	(x,y)4.3mm (z)2.8mm
			C	TEMP/TEP	$A_u$	retrospectives	connue	detect.surf.	I	marqueurs anat.	moy±SDV	(x,y)4.4±2.6mm (z)1.2±1.2
	[Goerres 02]	-	T/C	TEP/CT	$A_u$	rétrospectives	Id	-	I	pts.anat.	moy±SDV	0.4±11.7mm
	[Dey 99]	SAD	T	TEMP/CT	$A_p$	fant.phys. DataSpectrum	connue(aléat.)	-	E	pts(9)+params.	moy±SDV	1.1±0.6mm
	[Cai 99]	distance chamfrein	T	TEP/CT	$A_p$	fant. phys.	connue	detect.surf.	E	params.	moy	2-3mm, 1°
					$A_u$	rétrospectives	connue	detect. surf	E	params.	moy	3-4mm 1.5°
	[Pauna 03]	RC,CC,IM	T	TEP/IRM	$A_p+$	simul.num	connue(aléat.)	transf.	E	params.	eqm±SDV	4.69±4.41mm
	[Mäkelä 03b]	IM, IMN, surface	T/C	TEP/IRM	$A_p+$	simul.num.	connue(aléat.)	transf.	E	params.	eqm±SDV	2.89±0.44mm
			C	TEP/IRM	$A_u$	retrospectives	connue	detect.surf.	I	pts.	eqm.	2.8±0.5mm

TAB. 2.2: Etat de l'art des strategies d'évaluation des methodes de recalage d'images de thorax

			Hyp. d'évaluation		Mise en oeuvre				Métrique d'évaluation			
modal.	Référence	Méthodes de recalage	Niveau d'éval.	Modalités d'imagerie	Référence	Données d'évaluation	$T_{ref}$	Paramètres	Type	Mesurée sur	Type d'erreur	Erreur
1 modalité	[Eberl 96]	SAD	C	TEMP	$A_p$	fant. physique	connue	bruit+ activité	I	pts(15), ROI	moy±SDV	2.1±1.2mm
	[Hoh 93]	SAD,CSS	C	TEP	$A_u$	rétrospectives	connue(aléat.)	activité	I	params.	moy±SDV	(x,y)0.5±0.5mm (z)1.1±1.1mm (rot)0.5±0.4°
	[Bidaut 01]	MSD	C	IRM/IRM	$A_u$	rétrospectives	Id.	mouvement	I	marqueurs anat.	moy	(x,y)(2.35,1.81)mm
	[Klein 02b]	SDC	T	TEP/TEP	$A_p$	fant.num. (MCAT)	connue	mouvement	I	vect.depl.	moy	(x)1.9,(y)2.4,(z)6.8 mm
	[Turkington 97]	CC	T/C	TEP/TEP	$A_p$	fant. phys.	Id	bruit	E	intensite	moy	(x,y)1.7mm (z)4.2mm
2 modal.	[Pauna 03]	RC,CC,IM	T	TEP/IRM	$A_p+$	simul.num	connue(aléat.)	transf.	E	params.	moy±SDV	1.78±2.04mm

TAB. 2.3: Etat de l'art des stratégies d'évaluation des méthodes de recalage d'images de coeur



## Chapitre 3

# Stratégie d'évaluation de la précision des méthodes de mise en correspondance en imagerie thoracique et cardiaque TEP et IRM

Nous avons montré dans le chapitre 1, l'intérêt que présente la confrontation d'images issues de modalités différentes dans l'amélioration du diagnostic des pathologies cardiovasculaires. Les discussions avec la composante médicale du laboratoire nous ont conduit à privilégier prioritairement la fusion de données métaboliques issues de la TEP et de données anatomiques et fonctionnelles issues de l'IRM. Des travaux ont ainsi été entrepris pour le recalage rigide dans un premier temps, d'images de ces deux modalités. Conséquemment, la nécessité de l'évaluation des méthodes est apparue en particulier en termes de précision et de robustesse. Comme en analyse d'images médicales en général, l'absence de *référence absolue* est une difficulté majeure en imagerie cardiaque. Dans le cadre de l'évaluation, nous avons écarté le recours à des fantômes physiques qui, pour les plus accessibles, restent très loin de la réalité. Les avancées dans les connaissances des processus physiques impliqués dans les systèmes d'acquisition rendent possible la simulation *in silico* d'images.

Ainsi, le laboratoire CERMEP<sup>6</sup> (Centre d'Exploration et de Recherche Médicales par Émission de Positons) de Lyon a développé récemment un simulateur Monte Carlo en TEP qui a pu être validé relativement à une caméra TEP haute-résolution. Ce simulateur combiné à des images par RM réelles constitue notre référence pour l'évaluation des méthodes de recalage.

Cette stratégie présente néanmoins des limites et son réalisme peut être discuté ce qui nous empêche de la qualifier de *référence absolue*. Elle n'en constitue pas moins une base solide qui peut évoluer.

---

<sup>6</sup>[www.cermep.fr](http://www.cermep.fr)

### 3.1 Présentation de la stratégie

La stratégie d'évaluation des méthodes de recalage IRM/TEP en imagerie cardiaque repose sur la construction d'une référence constituée d'images TEP et par RM en parfaite correspondance. Les images TEP sont obtenues à partir de la simulation des phénomènes physiques résultant du comportement d'un produit de contraste en interaction avec les tissus d'un objet modèle extrait de données en IRM anatomiques. Afin d'évaluer les performances globales des algorithmes de recalage, un ensemble de transformations est généré aléatoirement. Les résultats de recalage obtenus pour l'ensemble des couples d'images tests peuvent ainsi être directement comparés. La figure 3.1 illustre schématiquement le principe de la stratégie.

Les étapes clés de cette stratégie sont présentées dans les sections suivantes. Il nous semble également important de préciser les évaluations conduites sur notre référence elle-même afin de juger de sa qualité.

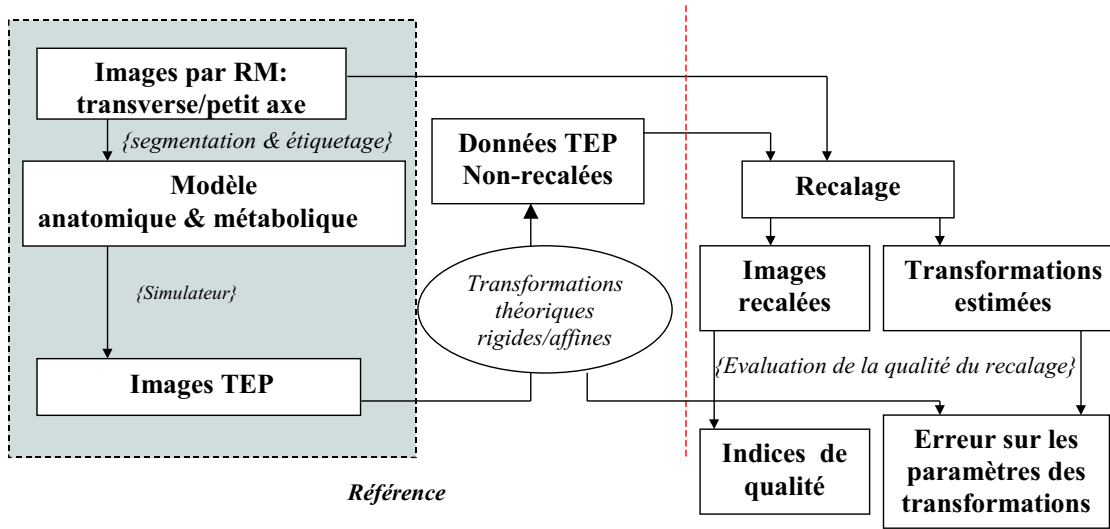


FIG. 3.1: Stratégie d'évaluation : vue générale

### 3.2 Construction de la référence

Notre stratégie d'évaluation s'appuie sur la construction d'une référence constituée d'images par RM et TEP en parfaite correspondance.

Les images par RM acquises chez un sujet sain sont segmentées à l'aide d'un expert médical afin de créer un modèle numérique 3D des principales structures anatomiques du thorax. Ce modèle est ensuite entré dans le simulateur TEP qui calcule par la suite les sinogrammes utilisés pour reconstruire les images TEP.

### 3.2.1 Données IRM de référence

Un simulateur TEP Monte Carlo a été utilisé pour obtenir les images TEP d'émission et de transmission à partir d'images par RM réelles. Pour simuler des acquisitions avec le simulateur Monte Carlo, on crée deux volumes, le premier représentant la distribution spatiale du traceur (fantôme d'émission) et le deuxième la position spatiale des milieux atténuateurs (fantôme de transmission) qui est utilisé par le simulateur Monte Carlo. Ces deux images tridimensionnelles peuvent être obtenues en segmentant manuellement une image par RM T1 de thorax.

Les images par RM ont été acquises à l'Hôpital Neuro-Cardiologique de Lyon avec un scanner 1.5 T (Siemens, Erlangen, Allemagne). Deux acquisitions IRM ont été réalisées.

Une première série a été acquise en apnée avec une séquence multi-coupes TSE HASTE (TR=872 ms et TE=120 ms) synchronisée sur le pic R de l'ECG. Elle est constituée de 15 coupes transverses 512\*512, les dimensions du pixel dans le plan de coupe sont de 0.97\*0.97 mm et l'épaisseur de coupe est de 8 mm. Une deuxième série d'acquisition a été réalisée en synchronisation avec l'ECG. Neuf coupes en petit axe couvrent le coeur ("cine fast long-angle shot two-dimensional", TR=650 ms et TE=4.8 ms) et, pour chacune de ces neuf coupes, neuf phases du cycle cardiaque ont été acquises. Les images ont été reconstruites en des matrices 128\*128 pixels et les dimensions du pixel dans le plan sont de 1.093\*1.093 mm et l'épaisseur de coupe est de 8 mm.

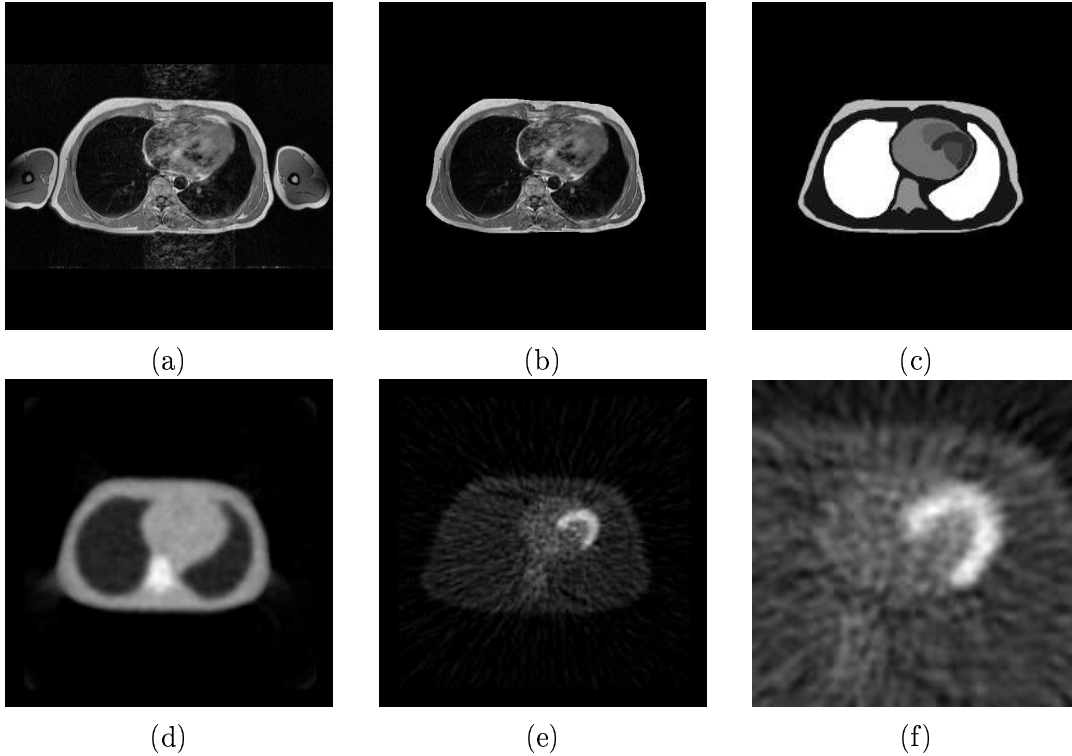


FIG. 3.2: Les données de référence IRM/TEP : (a) image par RM réelle, (b) image par RM sans les bras, (c) image par RM segmentée et étiquetée, (d) image simulée de TEP de transmission, (e) image simulée de TEP d'émission, (f) image simulée du cœur en émission

Organe/Structure	Radioactivité spécifique (Bq/cc)	Milieu atténuant
graisse	3000	<i>graisse</i>
muscle	1800	<i>muscle</i>
poumons	1500	<i>poumon</i>
os	3600	<i>os</i>
péricarde	3500	<i>muscle</i>
VG	11.000	<i>muscle</i>
cavité VG	4500	<i>eau</i>
cavité VD	3800	<i>eau</i>
foie	5000	<i>eau</i>

TAB. 3.1: Valeurs de fixation du FDG et choix des milieux atténuants utilisés lors de la simulation de thorax

### 3.2.2 Modèle numérique 3D (anatomique et métabolique)

Les bras ont été masqués des images par RM transverses du thorax (figure 3.2b) parce que pendant l'acquisition TEP le patient garde les bras positionnés au dessus de la tête pour éviter l'atténuation dans la région du cœur. Les images par RM ont été segmentées en 9 classes tissulaires (muscle, poumons, foie, graisse, os, ventricule gauche (VG), cavité ventriculaire gauche, péricarde, cavité ventriculaire droite) et le fond (figure 3.2 c). La segmentation manuelle a été réalisée à l'aide d'un radiologue en utilisant le programme de traitement d'images *ACTIVIS* développé au CERMEP. Pour chaque classe, on a établi une radioactivité spécifique en Becquerel par  $cm^3$  (Bq/cc) et un milieu atténuant (tableau 3.1).

La TEP de captation glucidique au fluorodesoxyglucose marqué (FDG18) est la technique de référence pour l'étude du métabolisme myocardique [Hartiala 95]. Le modèle métabolique, qui permet d'attribuer une activité spécifique du traceur à chaque tissu, provient d'une étude expérimentale qui a porté sur 6 témoins [Andrei 00]. Pour être capable de comparer les mesures entre les différents patients, les mesures absolues d'activité ont été rapportées à la dose totale (45 minutes après l'injection) du produit injecté au patient. Les doses injectées ont été calculées en fonction du poids du patient. La fixation normale du fluorodesoxyglucose est très variable d'une région du corps à l'autre et d'un patient à l'autre. Les poumons ont un métabolisme qui ne nécessite quasiment pas de glucose, et c'est dans le VG que le produit s'accumule avant d'être éliminé.

### 3.2.3 Simulation des données TEP

Les images TEP de transmission et d'émission ont été simulées à partir du modèle 3D anatomique et métabolique en utilisant le logiciel SORTEO (Simulation Of Realistic Transmission and Emission Object) [Reilhac 99, Reilhac 02]. Ce logiciel repose sur la méthode Monte Carlo qui permet d'intégrer la plupart des phénomènes physiques intervenant

en imagerie TEP. L'utilisateur peut modifier tous les paramètres du système de détection ; le simulateur utilise par défaut les paramètres de la caméra ECAT HR+ (CTI/SIEMENS) actuellement en fonction au CERMEP.

### Le rôle conceptuel du simulateur Monte Carlo

La méthode de Monte Carlo décrit un secteur scientifique très large dans lequel plusieurs processus, systèmes et phénomènes physiques sont simulés par des modèles statistiques en utilisant l'échantillonnage aléatoire d'une variable. L'idée générale d'une analyse de Monte Carlo est de créer un modèle qui soit le plus similaire possible au système physique réel, et de créer des interactions dans ce système basées sur des probabilités connues d'occurrence, avec un échantillonnage aléatoire de la fonction de densité de probabilité.

Plus le nombre d'événements augmente, plus la qualité du comportement moyen du système s'améliore, impliquant une diminution de la statistique de l'incertitude. Si, par expérience, on connaît la distribution des événements qui interviennent dans notre système, on est capable de générer une densité de probabilité et de l'échantillonner d'une manière aléatoire afin de simuler un système réel. En imagerie nucléaire, il y a actuellement un intérêt grandissant pour l'utilisation des techniques de Monte Carlo, pour la simulation de la TEMP et de la TEP [Buvat 02].

### Simulateur Monte Carlo d'acquisitions TEP

Le simulateur TEP Monte Carlo est un outil très puissant pour générer d'une manière précise les projections des distributions d'un traceur en tenant compte des caractéristiques du scanner et de la distribution de milieux atténuants. La simulation de Monte Carlo constitue la manière la plus fiable de générer des projections TEP selon les caractéristiques du système de détection, de la description du fantôme virtuel et des processus aléatoires inhérents à la détection des coïncidences. A partir d'une description spatio-temporelle 3D des milieux émetteur et atténuant, le simulateur permet de mimer l'acquisition de ce volume en simulant, pour chaque désintégration, les phénomènes suivants :

- Le parcours du positon issu de la désintégration du traceur dans la matière.
- L'annihilation d'un positon avec un électron de la matière donnant naissance à une paire de photons diamétralement opposés (avec une incertitude d'angulation également simulée).
- Le parcours d'un photon dans la matière vivante et environnementale du sujet, compte tenu des sections efficaces des phénomènes pouvant intervenir dans cette gamme d'énergie (Compton, Photoélectrique et Rayleigh).
- La caméra est vue comme étant une matrice cylindrique de détecteurs. Un photon est libre d'y déposer son énergie via une ou plusieurs interactions. Ce dernier est détecté lorsqu'il a déposé une énergie suffisante (seuil énergétique) dans un seul détecteur.
- Si les deux photons issus de la même annihilation sont détectés, l'évènement est alors enregistré. Lorsque l'enregistrement (sinogramme) présente un nombre d'évènements

conséquents, le volume peut être reconstruit (retroprojection filtrée ou reconstruction itérative).

- Le temps mort et le taux d'événements aléatoires sont simulés à partir d'une pré-simulation des taux d'événements simples présents sur chaque bloc détecteur à tout instant de l'acquisition.

Ces phénomènes étant pris en compte dans les images TEP, la méthode permet d'obtenir des données simulées réalistes tant du point de vue du bruit statistique présent dans les images que des distorsions spatiales générées.

## **Validation du simulateur TEP Monte Carlo S.O.R.T.E.O.**

### *Distorsion spatiale et taux d'évènements diffusés*

L'étude menée dans [Reilhac 99] a pour but de vérifier que les données TEP simulées contiennent les mêmes distorsions spatiales ainsi que le même taux d'évènements diffusés que les données TEP réelles.

Une première série de tests consiste à vérifier que le simulateur reproduit bien la résolution spatiale du scanner installé sur le site du CERMEP. On évalue la résolution spatiale d'un tel scanner en mesurant la dispersion spatiale sur l'image d'une ligne source émettrice, ou d'un point émetteur placé dans le champ de vue de la camera. Il est nécessaire de reproduire ce genre de mesure car cette dispersion n'est pas constante suivant la position de la ligne source dans le champ de vue (distance radiale par rapport au centre).

Dans un deuxième temps, le taux d'évènements diffusés issu de la simulation TEP a été comparé à des taux réels dans des conditions d'acquisition similaires (même distribution radioactive et même milieu atténuant). Les fantômes généralement utilisés pour réaliser ce genre de mesure sont cylindriques et remplis d'eau. Des inserts permettent d'y disposer une ligne source à plusieurs distances radiales. Le taux de diffusé est alors mesuré directement sur le sinogramme.

### *Taux de comptage*

Le taux de comptage du simulateur TEP Monte Carlo S.O.R.T.E.O. a été validé par rapport au scanner TEP ECAT EXACT HR+ du CERMEP. Pour cela, plusieurs simulations ont été effectuées pour différentes configurations d'acquisition et différents fantômes. Le nombre d'événements enregistrés en coïncidence ainsi que les événements fortuits déterminés par la ligne à retard ont été enregistrés à chaque fois. Le nombre d'événements vrais a été calculé comme la différence entre le nombre total d'événements enregistrés en coïncidence et le nombre d'événements fortuits. Les résultats ont été comparés avec ceux obtenus sur des acquisitions TEP réelles [Reilhac 04]. Pour les acquisitions TEP réelles, le fantôme physique utilisé est constitué d'un thorax et d'un cerveau. Pour les simulations TEP, le fantôme numérique est issu d'une segmentation manuelle des différentes régions d'images TEP réelles (transmission/émission) du fantôme physique. Les détails de la procédure sont décrits dans [Lartizien 02].

#### Données TEP

##### *Captation glucidique normale*

A partir des mesures réalisées sur les données TEP de 6 sujets sains, nous avons défini un modèle de captation glucidique normale, en attribuant à chaque structure de notre modèle spatial de l'anatomie la captation glucidique mesurée au sein de la population étudiée. Pour la simulation, on a supposé qu'on a injecté au patient une dose moyenne de 300 MBq. Comme l'acquisition ne se fait pas tout de suite après injection, en vue de la décroissance logarithmique de l'activité en fonction du temps, on obtient une activité totale de 255 MBq après 45 minutes. Les paramètres du modèle théorique et les simulations correspondantes sont présentées dans le tableau 3.1 et les figures 3.2 (d)-(f) et 3.7 (a).

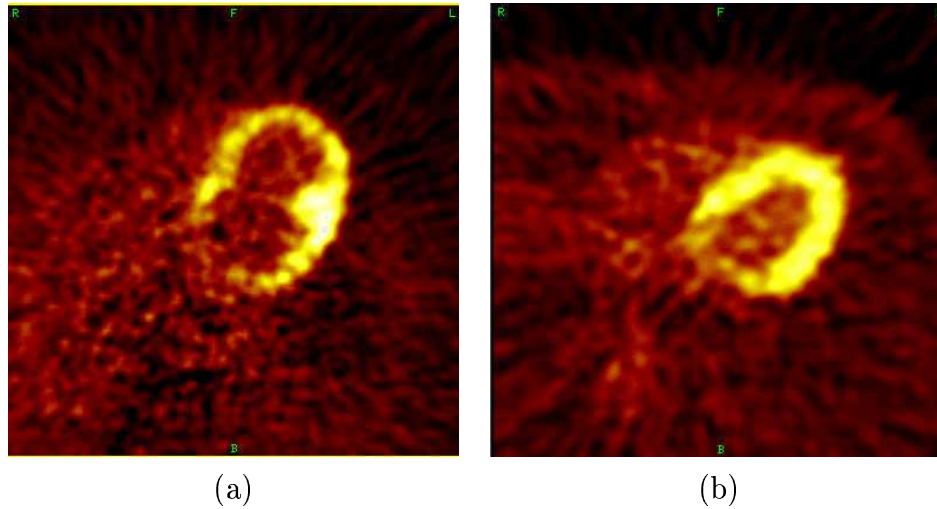


FIG. 3.3: Image de coeur TEP d'émission (a) réelle, (b) simulée

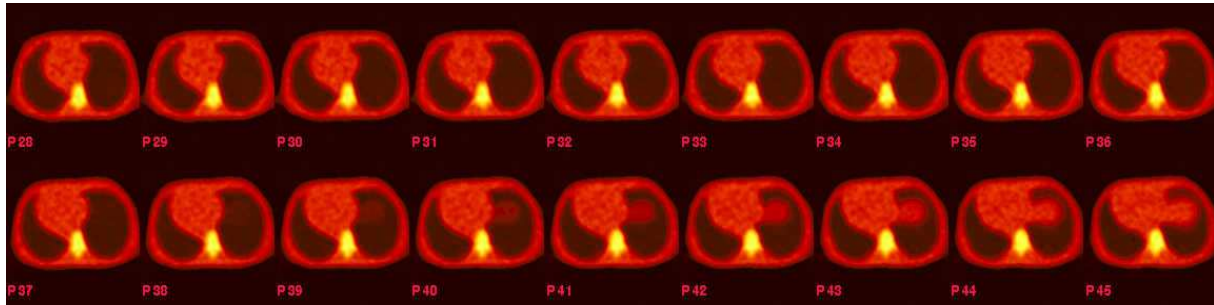


FIG. 3.4: Image 3D de thorax TEP de transmission simulée

La simulation de l'image de transmission a été réalisée en utilisant le Ge-68 comme source externe radioactive. Les images ont été reconstruites en des matrices volumiques de

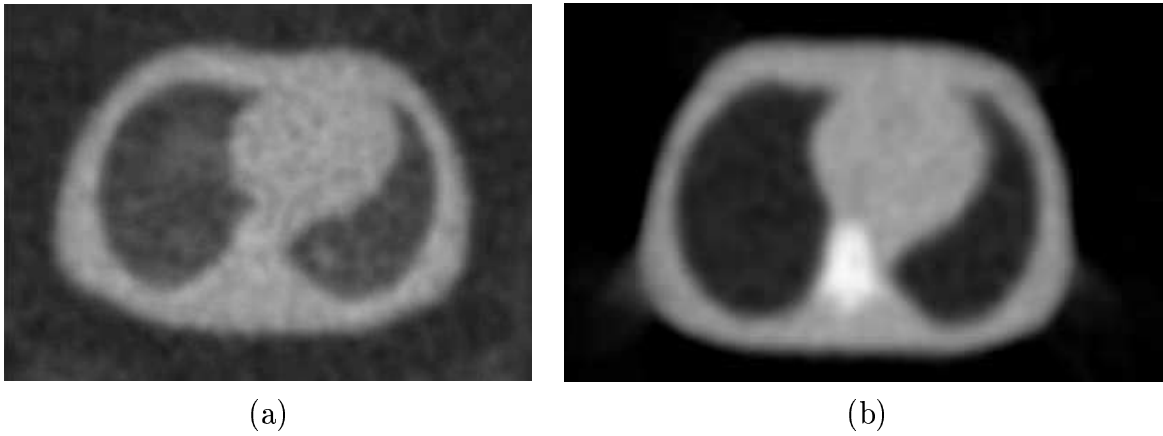


FIG. 3.5: Image de thorax TEP de transmission (a) réelle, (b) simulée

128\*128\*63 voxels. La reconstruction des images de transmission et d'émission est réalisée à partir des projections simulées, et l'algorithme FBP (Filtered Back Projection) avec filtre de Hanning (kernel=6 mm) a été utilisé. La dimension du voxel pour les images transverses de thorax d'émission et de transmission est de  $3.52*3.52*2.43 \text{ mm}^3$ . La dimension du voxel pour l'image d'émission de coeur est de  $1.76*1.76*2.43 \text{ mm}^3$ .

Les figures 3.3, 3.4 et 3.5 montrent que, visuellement, les données TEP d'émission et de transmission simulées sont très similaires aux données réelles.

#### *Captation glucidique dans un cas ischémique*

Un modèle théorique d'hypocaptation glucidique dans un cas ischémique a été défini en modifiant les valeurs d'activité du modèle de captation glucidique normale au niveau du ventricule gauche. Nous avons utilisé le même fantôme numérique anthropomorphique en ajoutant trois nouvelles classes ischémiques associées à des valeurs de radioactivité spécifique correspondant à trois valeurs d'hypocaptation glucidique : on a ainsi ajouté artificiellement une ischémie transmurale, avec un gradient d'hypocaptation de FDG à partir de l'endocarde de 80%, 40% et 20%, respectivement (les valeurs ont été calculées à partir d'une région de contrôle du tissu sain du ventricule gauche). Dans la figure 3.6, on montre la forme du tissu atteint par l'ischémie et l'étendue de chacune des trois zones ischémiques du ventricule gauche. Le nombre des pixels de chaque région ischémique représente 1.4%, 2.51% et 4.12% de l'ensemble pour une diminution de la captation glucidique de 80%, 40% et 20%, respectivement. Les paramètres de reconstruction sont les mêmes que pour le coeur sain. Les images obtenues sont représentées dans la figure 3.7 (b).



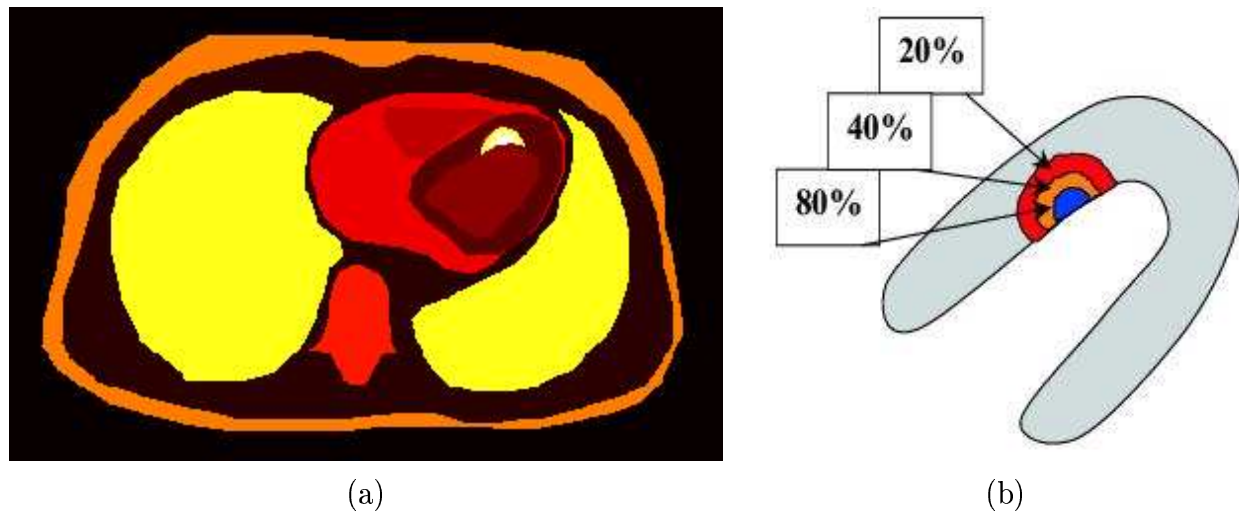


FIG. 3.6: (a) Modèle anatomique pathologique du fantôme numérique, (b) structure des régions ischémiques du ventricule gauche

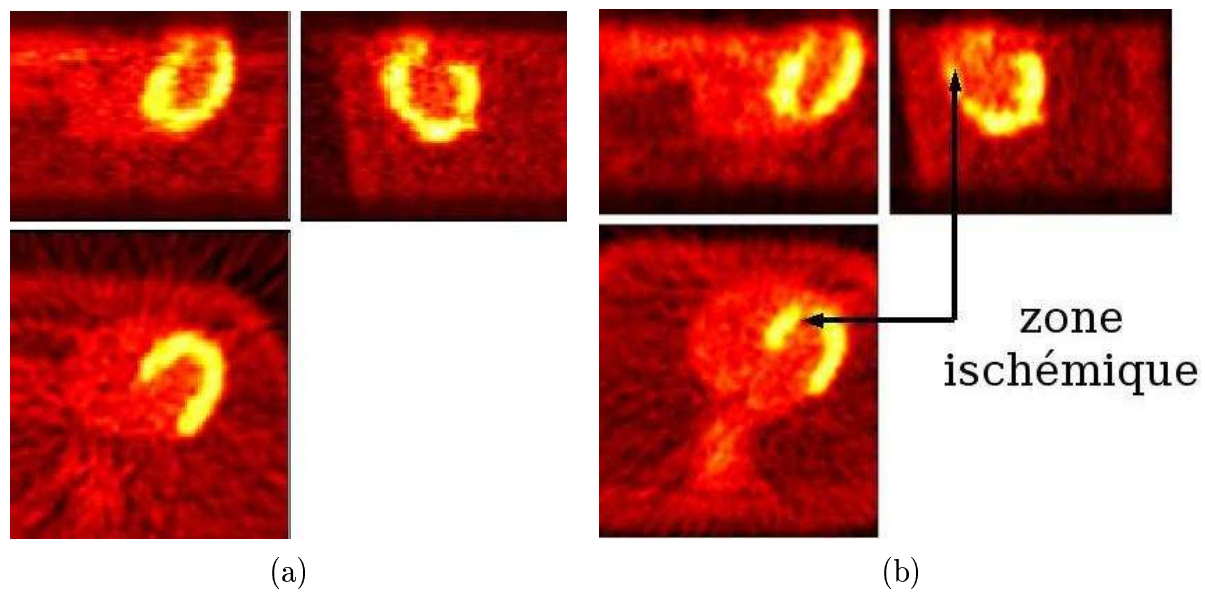


FIG. 3.7: Simulation d'images TEP de captation glucidique (a) normale et (b) dans un cas ischémique

### 3.3 Évaluation des performances des méthodes de mise en correspondance

#### A : EVALUATION DE LA PRÉCISION SPATIALE : CAS DE DONNÉES IRM RÉELLE/TEP SIMULÉE

Dans le contexte des données simulées, les images IRM et TEP sont parfaitement alignées. Dans le schéma d'évaluation proposé, un ensemble de transformations géométriques connues dites de référence est aléatoirement généré. En appliquant ces transformations géométriques sur l'image TEP, on obtient une série de couples d'images IRM/TEP non-recalées présentée à l'algorithme de recalage à évaluer. Les transformations estimées peuvent alors être comparées aux transformations réelles appliquées qui sont connues. La précision du recalage est estimée en calculant des indices de qualité à partir des images recalées et des transformations estimées.

#### 3.3.1 Transformations de référence

Soit  $T_0$  la transformation géométrique entre l'image par RM et l'image TEP simulée. Par construction,  $T_0$  est la transformation identité  $I$ . Pour simuler différentes situations de recalage, nous avons généré aléatoirement  $N$  transformations géométriques théoriques  $T_i$ ,  $i=1\dots n$ . Chacune a été obtenue par tirage aléatoire d'un vecteur de 6 paramètres (3 translations en *mm* et 3 rotations en *degrés*). Les amplitudes maximales des transformations ont été choisies compte tenu de l'expérience des praticiens. Chacune des transformations  $T_i$  a été appliquée aux images TEP avec une interpolation trilineaire.  $T$  représente la transformation géométrique de référence, qui est comparée à la transformation  $A$  estimée par la méthode de recalage à évaluer.

#### 3.3.2 Métriques d'évaluation de la précision

##### Critères liés aux transformations rigides

La mesure de différence entre deux transformations géométriques rigides ne doit pas être biaisée par la sélection de points clés (par ex. : les coins d'un cube). C'est la raison pour laquelle nous avons utilisé la mesure globale des erreurs calculée sur une région sphérique de rayon  $R$  décrite dans [Jenkinson 00], [Jenkinson 01].

Les transformations théorique  $T$  et estimée  $A$  sont comparées quantitativement en calculant la déviation  $d_{RMS}$  entre les transformations  $T^{-1} \times A$  et  $T_0$ . On obtient [Jenkinson 01] :

$$d_{RMS} = \sqrt{\frac{1}{5}R^2\text{Trace}(M^T M) + t^T t}, \quad (3.1)$$

où  $\begin{pmatrix} M & t \\ 0 & 0 \end{pmatrix} = T^{-1} \cdot A \cdot T_0 - I$ ,  $M$  est la matrice de rotation de dimensions  $3 \times 3$  et  $t$  est le vecteur de translation de dimensions  $3 \times 1$ . Pour un ensemble de  $n$  transformations théoriques, la précision d'une méthode de recalage est évaluée par la moyenne des  $d_{RMS}$  et la déviation standard.

#### Boîtes à moustaches

On estime également la différence entre deux transformations géométriques rigides théorique ( $th$ ) et mesurée ( $m$ ) en calculant des erreurs intrinsèques de la méthode en comparant les 6 composantes d'une transformation rigide : trois rotations et trois translations. Pour l'axe  $X$ , par exemple, les erreurs en translation et rotation sont définies par :

$$\Delta tr_X = |tr_X^{th} - tr_X^m|, \Delta r_X = |r_X^{th} - r_X^m| \quad (3.2)$$

La représentation des boîte à moustaches (ou *box and whiskers plot* ou *box-plot* en anglais) est employée pour illustrer les erreurs en valeur absolue de chacun des 6 paramètres de la transformation géométrique [Grova 01].

La boîte à moustaches est une représentation graphique simple mais efficace de la statistique d'un échantillon de données (figure 3.8). Elle est constituée d'un rectangle (*box* en anglais) orienté selon les axes d'un système de coordonnées. L'échelle de l'axe vertical correspond aux données. Les extrémités inférieure (*bas*) et supérieure (*sommet*) du rectangle correspondent au premier et au troisième quartile, respectivement. La différence entre le premier et le troisième quartile d'une distribution s'appelle l'écart inter-quartile, noté  $I_q$ . Le rectangle contient la moitié centrale des données. Le rectangle est partagé en deux par un trait horizontal au niveau de la médiane.

On définit le *pas* comme le segment de longueur  $1.5I_q$  et on considère les données situées entre les extrémités plus ou moins le *pas*. La plus élevée de ces données est appelée *limite supérieure*. Un trait vertical (*whisker* en anglais ou moustache) s'étend du sommet jusqu'à la limite supérieure. De façon similaire, on définit une *limite inférieure* et une moustache qui s'étend du bas jusqu'à la limite inférieure.

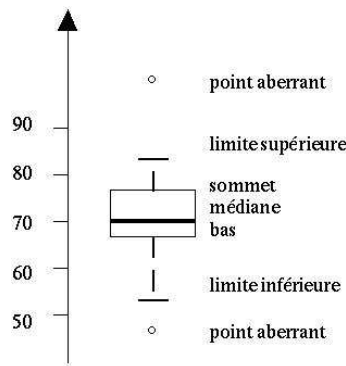


FIG. 3.8: Boîte à moustaches pour la représentation de la distribution des erreurs

Les observations qui dépassent les limites sont considérées comme des points aberrants et marquées individuellement avec le symbole  $o$  (figure 3.8). Cette représentation permet de visualiser la dispersion, l'asymétrie des données et les points aberrants (observations exceptionnelles, parfois erronées) de l'échantillon. Plusieurs échantillons peuvent être représentés simultanément et comparés en alignant ce type de représentation.

## **B : ETUDE DE CONSISTANCE : CAS DES DONNÉES RÉELLES**

Quand on ne connaît pas la transformation géométrique qui met en correspondance les images IRM et TEP (e.g. cas de données réelles) on ne peut pas évaluer la précision d'une méthode de recalage mais on peut étudier sa consistance comme dans [Jenkinson 01] (voir chapitre 2). Comme précédemment, on peut comparer  $T^{-1} \times A$  avec  $T_0$ , où  $T_0$ , représente dans ce cas la transformation géométrique estimée entre les images par RM et TEP réelles en utilisant la meilleure méthode parmi toutes les méthodes évaluées en termes de précision à l'aide d'une référence (cas **A**).

### **3.4 Hypothèse d'évaluation**

#### **3.4.1 Niveau d'évaluation**

Le niveau d'évaluation de cette étude est un niveau technique.

#### **3.4.2 Objectif de l'évaluation**

Les erreurs de recalage doivent être comparées à la résolution spatiale des images, sachant que le paramètre critique est l'épaisseur des coupes.

L'objectif d'évaluation peut être formulé de la manière suivante : “est ce que le recalage multimodal IRM/TEP utilisant une approche iconique (sans segmentation, basée sur les statistiques des images) ou une approche surfacique permet d'atteindre une erreur quadratique moyenne estimée sur le thorax et le coeur, significativement inférieure à la taille du voxel des données TEP (que nous identifions  $V_{TEP}$ ) ?”

#### **3.4.3 Test d'hypothèses**

Afin de répondre à la question posée par notre objectif de validation, on formule deux types de test d'hypothèses unilatéraux sur une moyenne (à gauche,  $T_1$  et à droite,  $T_2$ ) (voir Annexe D pour plus de détails sur la signification d'un test d'hypothèses) :

$$T_1 : \begin{cases} H0 : E[d_{RMS}] \geq V_{TEP} \\ H1 : E[d_{RMS}] < V_{TEP} \end{cases} \quad (3.3)$$

$$T_2 : \begin{cases} H0 : E[d_{RMS}] \leq V_{TEP} \\ H1 : E[d_{RMS}] > V_{TEP} \end{cases} \quad (3.4)$$

où  $E$  indique l'espérance mathématique et  $V_{TEP}$  est la valeur présumée de la moyenne. Le premier type de test,  $T_1$ , est utilisé pour prouver que la moyenne des erreurs d'une méthode de recalage multimodal IRM/TEP est inférieure à la taille du voxel des images TEP, tandis que le test de type  $T_2$  est employé pour le cas où on veut prouver que l'erreur de recalage est supérieure à cette valeur.

Ces tests sont résolus à l'aide d'une statistique de Student à partir de la moyenne  $\overline{rms}$  et l'écart-type  $\sigma$  calculés sur la distribution d'erreurs  $d_{RMS}(i)$ ,  $i = 1 \dots n$ ,  $n$  étant le nombre de tests réalisés :

$$\overline{rms} = \frac{1}{n} \sum_{i=1}^n d_{RMS}(i) \quad (3.5)$$

$$\sigma = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (d_{RMS}(i) - \overline{rms})^2} \quad (3.6)$$

# Chapitre 4

## Evaluation comparative de méthodes de mise en correspondance

La stratégie d'évaluation présentée au chapitre 3 est mise en oeuvre afin d'estimer la précision de trois méthodes de recalage dans le cas de la mise en correspondance d'images TEP et par RM du thorax et du coeur. Deux d'entre elles sont des approches iconiques (les mesures de similarité sont basées sur les intensités des images), tandis que la troisième est une approche surfacique.

Nous présentons tout d'abord le principe et les principales caractéristiques des méthodes de recalage évaluées. Les hypothèses et les résultats de l'évaluation comparative de ces méthodes sont ensuite exposés.

### 4.1 Présentation des méthodes de mise en correspondance évaluées

#### 4.1.1 Les méthodes de recalage iconiques

Les méthodes de recalage iconiques occupent une place privilégiée parmi les méthodes de recalage de données médicales. Ce sont surtout les approches basées sur la théorie de l'information, notamment de l'information mutuelle, qui ont montré leur efficacité et leur robustesse dans un grand nombre de situations, notamment les problèmes de recalage mono-et multi-modalités en imagerie cérébrale. L'avantage de ces méthodes de recalage est qu'elles ne requièrent en général aucun pré-traitement (e.g. segmentation). Leur mise en oeuvre est généralement assez similaire et repose sur un processus itératif d'estimation de la transformation et de calcul de la similarité.

#### Méthode HMR (Home Made Registration)

Nous avons développé une méthode de recalage rigide iconique dont le principe est décrit dans la figure 4.1. Soit  $I_1$  l'image de référence et  $I_2$  l'image flottante à recaler sur  $I_1$ . Le processus itératif est initialisé par une transformation géométrique  $T_0$  : l'image

transformée  $I_2^i$  de  $I_2$  est calculée et comparée à  $I_1$  par une mesure de similarité. Si la mesure de similarité ne respecte pas le critère d'arrêt, une transformation  $T_{i+1}$  est prédite par modification de  $T_i$  et une nouvelle image  $I_2^i = T_i(I_2)$  est créée.  $T^*$  est la transformation géométrique optimale qui met en correspondance les images  $I_1$  et  $I_2$ .

Dans le contexte du recalage multi-modalités, nous avons investigué trois mesures de similarité basées sur les intensités des voxels : le Rapport de Corrélacion (RC) [Roche 98], le Coefficient de Corrélacion (CC) [Roche 01] et l'Information Mutuelle (IM) [Wells 96, Maes 97] (tableau 4.1 et Annexe B). Pour chaque transformation géométrique, l'estimation de tous ces critères est réalisée à partir de l'histogramme joint.

L'optimisation des différents critères suivant la transformation géométrique  $T$  est réalisée à l'aide de la méthode de Powell, associée à l'algorithme de recherche mono-dimensionnelle de Brent [Press 92]. Nous avons également mis en oeuvre une approche multi-résolutions par ondelettes à trois niveaux.

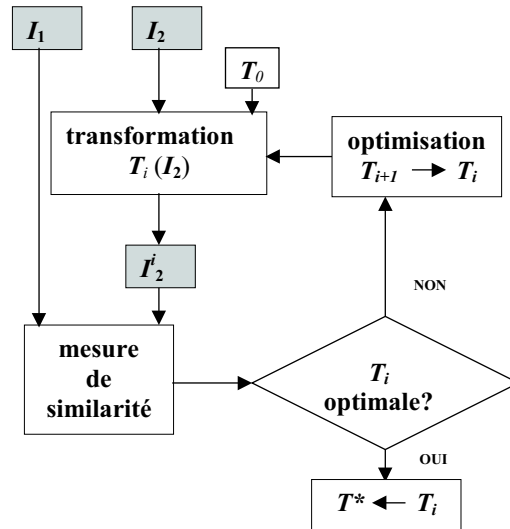


FIG. 4.1: Principe de la méthode de recalage rigide HMR

### Modèle de transformation géométrique

La transformation géométrique utilisée est une transformation rigide 3D à 6 paramètres combinant trois translations et trois rotations autour des axes de coordonnées (angles Euler  $\alpha$ ,  $\beta$ ,  $\gamma$ ).

Pour chaque point  $(x, y, z)$  d'une image, on définit une transformation géométrique conduisant aux nouvelles coordonnées  $(x', y', z')$ . L'ordre des transformations élémentaires choisi est le suivant :

<i>Nature de la dépendance</i>	<i>Mesure de similarité</i>	<i>Définition</i>	<i>Min</i>	<i>Max</i>
Fonctionnelle	Rapport de Corrélation	$1 - \frac{E[Var(I j)]}{Var(I)}$	0	1
Affine	Coefficient de Corrélation	$\sum_{i,j} p_{i,j} \frac{(i-\bar{I})(j-\bar{J})}{\sigma_i \sigma_j}$	-1	1
Statistique	Information Mutuelle	$\sum_{i,j} p_{i,j} \log \frac{p_{i,j}}{p_i p_j}$	0	$\infty$

TAB. 4.1: Définition des mesures de similarité basées sur les niveaux de gris :  $p_{i,j}$  est la probabilité conjointe des images et  $p_i$  et  $p_j$  sont les distributions marginales correspondantes (Annexe B).

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = T_x \cdot T_y \cdot T_z \cdot R_z \cdot R_y \cdot R_x \cdot \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (4.1)$$

où  $T_x, T_y, T_z, R_x, R_y, R_z$  représentent respectivement les matrices de translations et des rotations (Annexe C)

## Mesures de similarité

Les mesures de similarité utilisées sont définies dans le tableau 4.1. Ces mesures sont maximales lorsque les images sont correctement alignées. Pour plus de détails se reporter à l' Annexe B.

## Interpolation

Afin de gagner en précision, la transformation géométrique utilisée pour le recalage est exprimée dans un espace continu (e.g. pour une transformation rigide, les translations sont exprimées en *mm* et les rotations en *degrés*). La plupart des méthodes de recalage iconiques utilisent des méthodes itératives d'optimisation et requièrent ainsi de nombreuses évaluations successives de la mesure de similarité. Chaque estimation de la mesure de similarité nécessite directement ou indirectement, un ré-échantillonnage de l'image flottante dans le référentiel de l'image de référence via la transformation géométrique, et donc le recours à une interpolation des intensités de l'image flottante.

Nous avons utilisé une interpolation par B-splines développée par Thevenaz<sup>7</sup>. L'image flottante sur laquelle porte l'interpolation est l'image TEP.

<sup>7</sup><http://bigwww.epfl.ch/>



### Approche multi-résolutions

Bien souvent, l'optimisation est imbriquée dans un schéma multi-résolutions. Le principe d'un tel schéma repose sur une décomposition pyramidale des deux images de niveaux de résolution élevée à faible. On recalc successivement les couples d'images en allant des niveaux de résolution les plus grossiers vers le plus fin. Il est généralement admis que les approches multi-résolutions offrent le double avantage d'accélérer la convergence et de réduire la sensibilité aux maxima locaux (les fonctions de coût étant d'autant plus convexes que la résolution des images est faible).

Dans notre approche multi-résolutions, une pyramide à trois niveaux est générée par une décomposition par ondelettes (avec les facteurs d'échelle 4, 2 et 0). Des paramètres différents de l'algorithme de recalage seront attribués à chaque niveau de résolution. Il s'agit en particulier du type d'interpolation et du seuil de tolérance (*tol*) qui intervient dans la méthode d'optimisation Powell. Les paramètres sont les suivants :

- premier niveau de décomposition (résolution la plus basse) : interpolation par plus proche voisin et  $tol = 0.4$
- deuxième niveau de décomposition : interpolation linéaire et  $tol = 0.1$
- troisième niveau (pleine résolution) : interpolation quadratique et  $tol = 0.01$

### Recherche de la transformation optimale

La recherche de la transformation optimale est réalisée par la méthode de Powell [Press 92]. La méthode de Powell recherche l'optimum d'une fonction multi-dimensionnelle par optimisations monodimensionnelles successives selon des directions mutuellement conjuguées dans l'espace des paramètres. Chaque optimisation mono-dimensionnelle est réalisée à l'aide de la méthode de Brent, qui approche localement la fonction par une parabole. Si la fonctionnelle de coût à optimiser est quadratique, on montre que la méthode de Powell converge en un nombre fini d'itérations.

L'ordre dans lequel les paramètres de la transformation sont optimisés peut avoir une influence sur le résultat comme le souligne par exemple D. Sarrut [Sarrut 00]. Dans notre application, ils sont optimisés selon l'ordre défini dans l'équation (4.1).

### Méthode IMP

La méthode de recalage IMP (Information Mutuelle-Parzen) est issue de la librairie "open-source" en C++, ITK<sup>8</sup>-*Insight Segmentation and Registration Toolkit*. Cette librairie est dédiée au recalage et à la segmentation d'images. Sa structure orientée objet et sa modularité permettent une construction par blocs d'une méthode de recalage. Nous avons configuré un algorithme basé sur l'Information Mutuelle avec une interpolation trilinéaire et une approche multi-résolutions à 3 niveaux. L'Information Mutuelle est calculée sur un nombre fixe d'échantillons (50 dans notre cas, fenêtre de Parzen) afin d'accélérer la

---

<sup>8</sup>[http ://www.itk.org/](http://www.itk.org/)

convergence. L'optimisation est réalisée par une descente de gradient avec 1000 itérations maximum à chaque niveau de résolution. La rotation est exprimée sous forme de quaternions (voir Annexe C).

La matrice de rotation exprimée à l'aide des composantes du quaternion  $q = [q_0, q_1, q_2, q_3]$  est définie par :

$$Q = \begin{bmatrix} q_0^2 + q_1^2 - q_2^2 - q_3^2 & 2(q_1q_2 + q_0q_3) & 2(q_1q_3 - q_0q_2) \\ 2(q_1q_2 - q_0q_3) & q_0^2 - q_1^2 + q_2^2 - q_3^2 & 2(q_2q_3 + q_0q_1) \\ 2(q_1q_3 + q_0q_2) & 2(q_2q_3 - q_0q_1) & q_0^2 - q_1^2 - q_2^2 + q_3^2 \end{bmatrix} \quad (4.2)$$

La matrice de rotation exprimée en angles d'Euler est :

$$E = R_z * R_y * R_x \quad (4.3)$$

Pour le passage entre les représentations par angles d'Euler et les quaternions, on déduit de (4.2) et (4.3) par identification les angles d'Euler exprimés à l'aide des coefficients de la matrice  $Q$  :

$$\alpha = \text{atan2}(Q(2, 3), Q(3, 3)) ; \beta = \text{asin}(-Q(1, 3)) \text{ et } \gamma = \text{atan2}(Q(1, 2), Q(1, 1))$$

En utilisant ces transformations, on peut facilement déduire les angles d'Euler à partir de la matrice de transformation décrite en termes de quaternions par l'algorithme de mise en correspondance IMP.

#### 4.1.2 Méthode de recalage basée sur des surfaces

Une méthode de recalage rigide d'images TEP et par RM du thorax et du coeur, basée sur la mise en correspondance de surfaces, a été développée dans le cadre des travaux de thèse de Timo Mäkelä en collaboration avec le Laboratory of Biomedical Imaging<sup>9</sup> (Helsinki University of Technology, Finlande) [Mäkelä 03a], [Mäkelä 01]. Nous notons cette méthode SBR (Surface Based Registration). Son principe est illustré dans la figure 4.2. La transformation rigide issue du recalage est le résultat de la minimisation d'une distance entre un ensemble de points appartenant aux structures anatomiques extraites de l'image TEP et une carte de distance construite à partir de l'image par RM segmentée. La segmentation des structures est réalisée par déformation d'un gabarit a priori du thorax selon la technique des pyramides déformables proposée dans [Lötjönen 99]. L'initialisation est réalisée soit manuellement (par adaptation des dimensions de la boîte englobante) ou

<sup>9</sup><http://www.hut.fi/Units/Biomedical>

encore par recalage rigide entre le gabarit et l'image binarisée. Dans notre étude, l'initialisation a été réalisée manuellement. Trois techniques d'optimisation différentes ont été évaluées pour la recherche de la transformation optimale. Ce sont les méthodes de Powell, du simplexe [Press 92] et une méthode non-standard décrite dans [Mäkelä 01], appelée *optimisation sur une grille discrète*.

## 4.2 Évaluation de la précision des méthodes

La stratégie d'évaluation décrite dans le chapitre 3 est mise en oeuvre pour les algorithmes de recalage décrits précédemment. La référence utilisée est constituée du couple d'images par RM-réelle/TEP-simulée de coeur et de thorax. La construction de la référence est détaillée dans la section 3.2. Les images par RM et TEP simulée sont interpolées à la même résolution. Les valeurs des intensités des deux types d'images ont été recadrées linéairement entre 0 et 255 niveaux de gris afin que le calcul de l'histogramme conjoint soit plus court lors de chaque itération. L'image flottante, sur laquelle sont appliquées les transformations tests aléatoires, est l'image TEP simulée. L'image de référence est l'image par RM.

Notre objectif est d'évaluer la précision des méthodes de recalage rigide dans deux situations pratiques qui sont :

- Le recalage d'images par RM transverses et d'images TEP du thorax chez un même patient.
- Le recalage d'images par RM petit axe et d'images transverses TEP du coeur chez un même patient.

Ce second cas correspond aux conditions cliniques en imagerie cardiaque qui sont à l'origine de ce travail. En effet, le protocole initialement utilisé dans l'équipe ne prévoyait pas l'acquisition d'images transverses en IRM. La représentation des *boîtes à moustache* et des tests statistiques permettent de comparer les méthodes.

### 4.2.1 Mise en correspondance d'images de thorax

Le recalage d'images TEP de transmission et par RM transverses est évalué pour les trois méthodes décrites précédemment.

Nous avons généré de manière aléatoire 50 transformations théoriques s'appuyant sur deux distributions gaussiennes (une pour les rotations et une pour les translations). Les experts ont estimé que les mouvements de rotation et de translation du patient sont limités, en général, à  $\pm 5$  degrés et à  $\pm 5$  cm, respectivement. Les images ont été interpolées linéairement et la taille du voxel isotrope est de  $1.95 \text{ mm}^3$  pour une grille de  $256 \times 256 \times 150$  voxels. La méthode HMR utilise trois mesures de similarité différentes, tandis que la méthode SBR a été évaluée avec trois méthodes d'optimisation différentes.

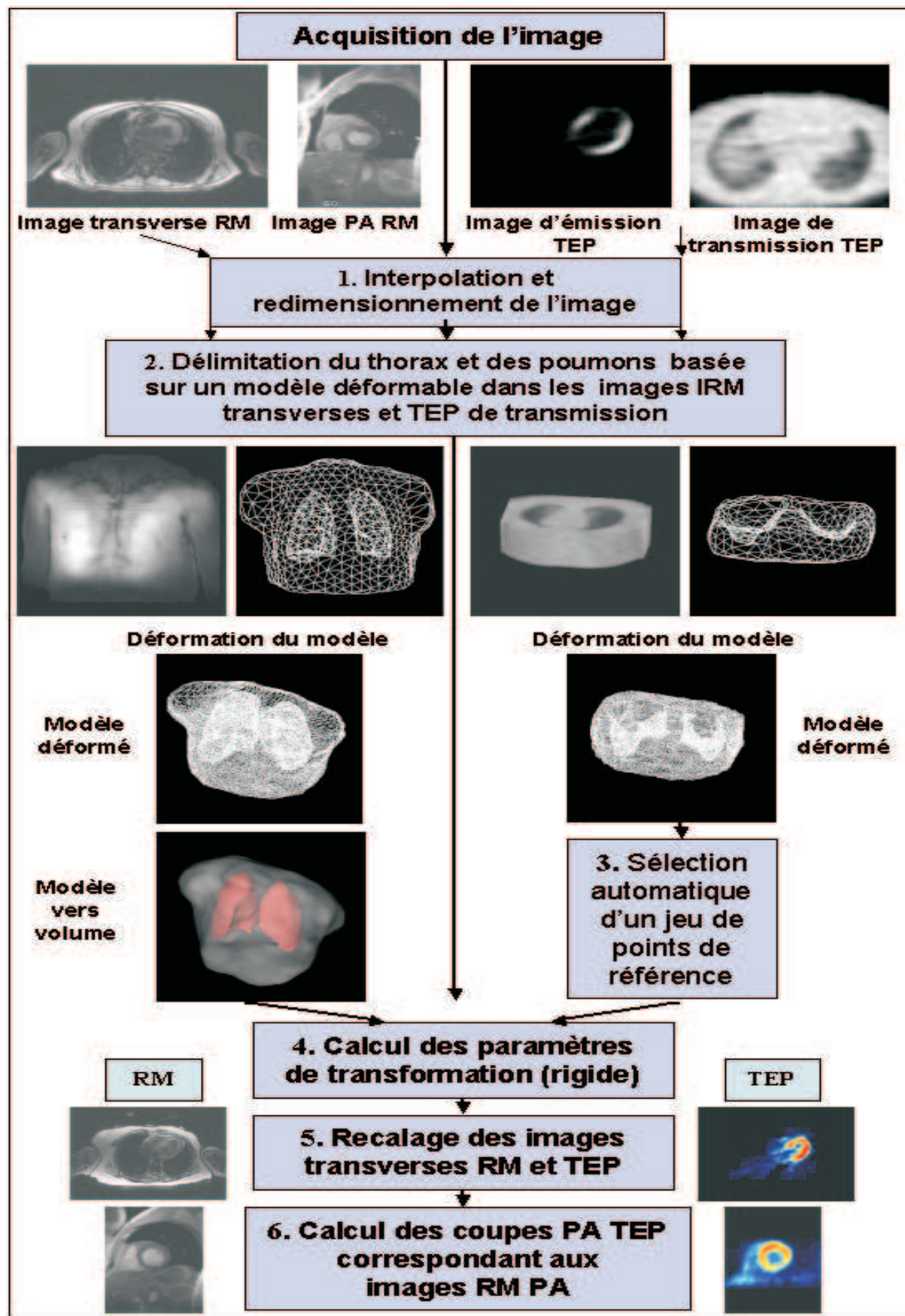


FIG. 4.2: Principe de la méthode de recalage SBR basée sur les surfaces (issu de [Mäkelä 03a])

### Recalage

La figure 4.3 illustre deux cas de recalage d’images transverses de thorax par RM et TEP de transmission ; l’un où l’erreur  $d_{RMS}$  est inférieure à 3.52 mm, la taille des pixels des images TEP (recalage avec IMP, figure 4.3 (d)) et l’autre quand elle est supérieure à 3.52 mm (recalage avec HMR+CC, 4.3 (e)). La simple observation visuelle des images recalées ne permet pas directement de dire si l’un des recalages est supérieur à l’autre. Les deux résultats semblent très similaires. Le calcul de l’erreur  $d_{RMS}$  permet de nuancer l’interprétation.

### Tests d’hypothèses

Pour évaluer si l’erreur  $d_{RMS}$  issue du recalage des différentes méthodes est significativement inférieure ou supérieure à 3.52 mm (taille du pixel des données TEP de thorax dans le plan), on définit les hypothèses nulle et alternative pour les deux types de tests d’hypothèses (section 3.4.3) :

$$T_1 : \begin{cases} H_0 : E[d_{RMS}] \geq 3.52mm \\ H_1 : E[d_{RMS}] < 3.52mm \end{cases} \quad (4.4)$$

$$T_2 : \begin{cases} H_0 : E[d_{RMS}] \leq 3.52mm \\ H_1 : E[d_{RMS}] > 3.52mm \end{cases} \quad (4.5)$$

Ces tests sont résolus à l’aide d’une statistique de Student, avec un *seuil de signification*  $\alpha$  de 1%. Les résultats du test de Student pour toutes les méthodes de recalage évaluées sont présentés dans le tableau 4.2, où la notation “ $< 0.0001(T_1)$ ” signifie que le paramètre *pvalue* est issu d’un test d’hypothèses de type  $T_1$ .

### Résultats de l’évaluation

Le recalage d’images par RM et TEP de transmission de thorax est évalué pour les méthodes IMP et HMR.

Pour l’évaluation du recalage d’images par RM/TEP avec la méthode surfacique (SBR), les figures 4.4 (a)-(c) et le tableau 4.2 montrent clairement que toutes les méthodes d’optimisation donnent des résultats similaires et non satisfaisants (erreurs moyennes de l’ordre de 7 voxels), la méthode de Powell étant légèrement plus précise que les autres. Le test d’hypothèses  $T_2$  rejette l’hypothèse nulle avec une *pvalue* inférieure à 0.0001. Cette méthode n’est donc pas en mesure de recalcr les images par RM et TEP de thorax avec la précision souhaitée.

Sur l’ensemble des expériences de recalage avec les méthodes iconiques (figure 4.4 (d)-(g) et tableau 4.2), le test d’hypothèses  $T_1$  confirme que les méthodes de recalage les plus performantes pour cette application sont IMP, HMR+IM et HMR+RC (erreurs  $d_{RMS}$  de  $1.38 \pm 0.21$ ,  $1.33 \pm 0.04$  et  $1.37 \pm 0.41$  voxels, respectivement) où la *pvalue* est inférieure à 0.0001. La moyenne d’erreurs  $d_{RMS}$  issues des recalages avec la méthode HMR+CC est de  $1.86 \pm 0.27$  voxels et la valeur *pvalue* correspondante issue du test  $T_1$  est égale à 0.001.

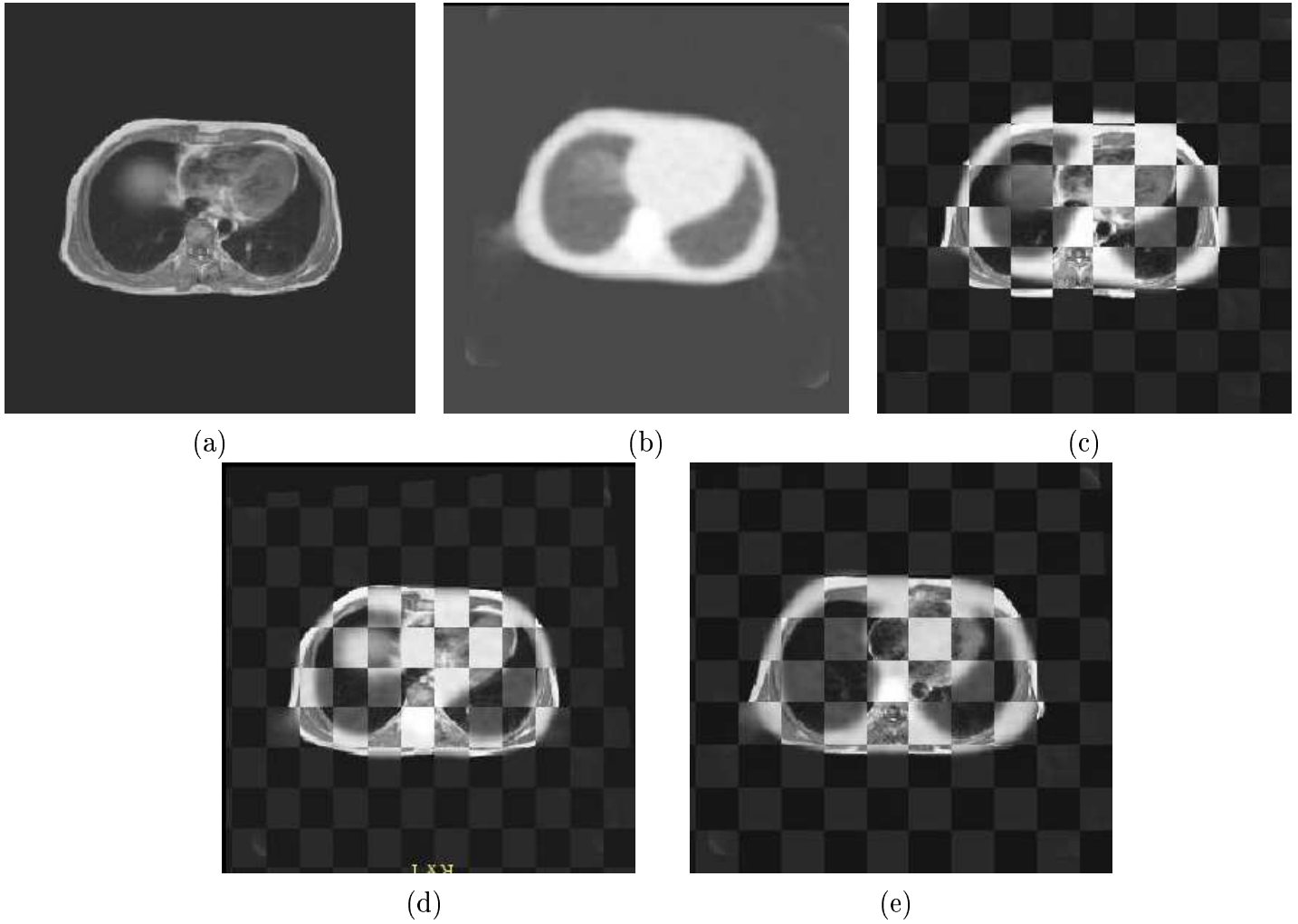


FIG. 4.3: Recalage d'images transverses par RM et TEP de transmission de thorax : (a) image par RM, (b) image TEP transformée ; visualisation en damier des volumes avant le recalage (c) et après le recalage avec les méthodes IMP (d) et HMR+CC (e)

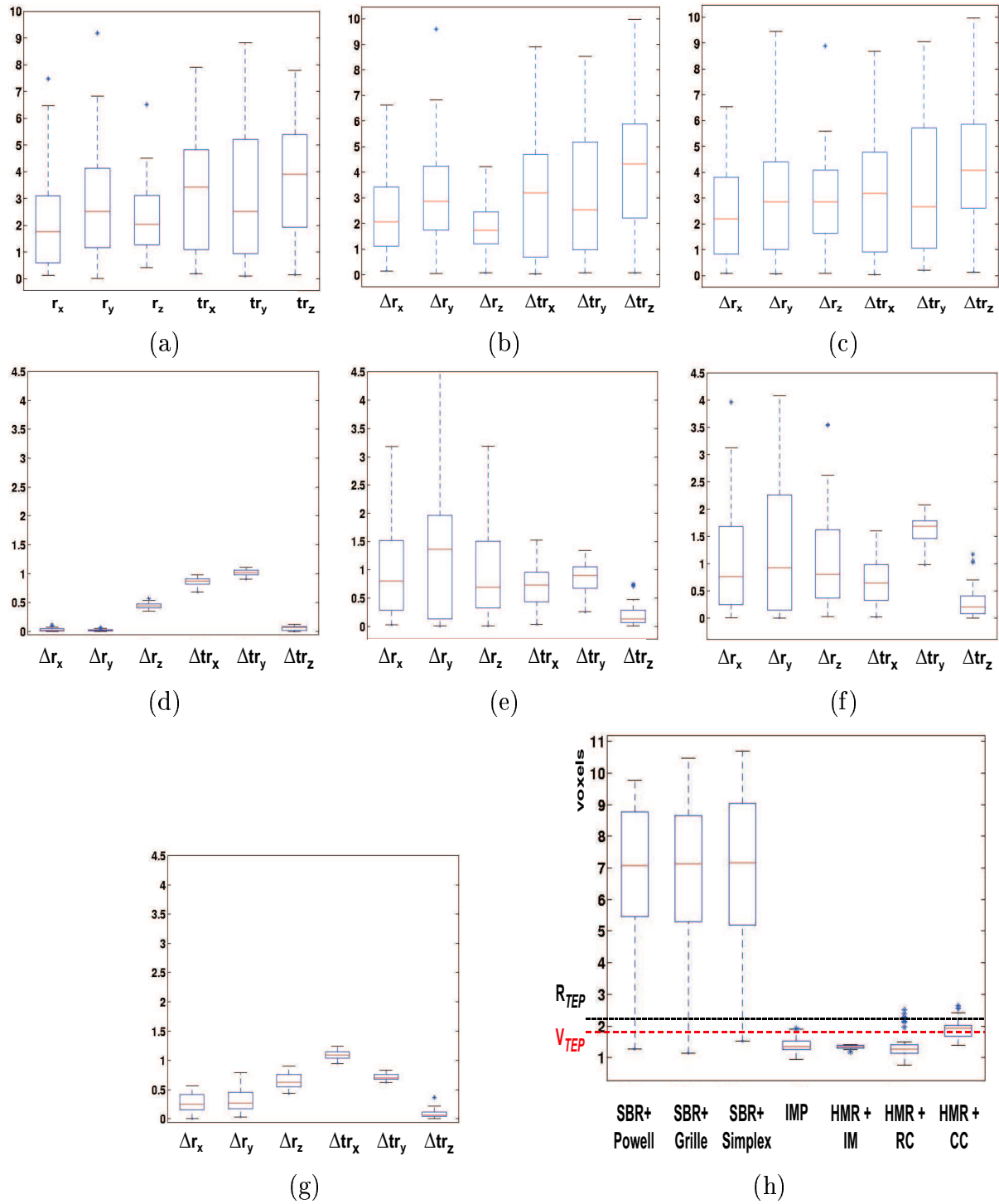


FIG. 4.4: Erreurs de recalage sur les paramètres de la transformation pour les images par RM et TEP de transmission de thorax avec la méthode SBR et les optimisations de Powell (a), grille (b) et Simplexe (c), avec la méthode HMR et les mesures de similarité IM (d), RC (e) et CC (f) et avec la méthode IMP (g). (h) erreurs  $d_{RMS}$  pour chaque méthode. Les erreurs en translation sont en voxels (1 voxel=1.95 mm) et les erreurs en rotation sont en degrés. La taille du voxel en TEP ( $V_{TEP}$ ) et la résolution physique de la caméra TEP sont respectivement égales à 3.52 mm et 4.1 mm.

Ceci indique qu'il faut rejeter l'hypothèse nulle  $H_0$  mais avec un risque de se tromper de 0.001 et que l'incertitude sur la précision de recalage est plus grande que pour les autres méthodes. Elle est moins précise que les méthodes IMP et HMR+RC.

Les erreurs estimées sur les six paramètres de la transformation géométrique indiquent que, parmi les méthodes les plus performantes, les erreurs les plus importantes sont obtenues en rotation pour la méthode HMR+RC ( $1.36 \pm 1.27$  degrés) et en translation pour IMP et HMR+IM ( $1.09 \pm 0.07$  et  $1.02 \pm 0.05$  voxels, respectivement)

L'avantage principal de la médiane, par rapport à la moyenne arithmétique, est qu'elle n'est pas indûment influencée par quelques données exceptionnelles (extrêmes). Elle est donc parfois plus représentative de la précision d'une mesure que la moyenne. Il suffit en effet que l'algorithme ait nettement échoué dans un cas pour que la moyenne soit anormalement élevée. Un tel échec est généralement dû à la convergence vers un minimum local. Les valeurs médianes et moyennes des erreurs de recalage présentées dans le tableau 4.2 sont cependant du même ordre et ceci montre que les résultats issus des expériences individuelles de recalage ne varient pas nettement (pas de points extrêmes). Les statistiques sur l'erreur  $d_{RMS}$  pour l'ensemble des méthodes sont représentées dans la figure 4.4 (h).

Le temps de calcul pour une expérience de recalage d'images TEP/IRM de thorax avec les méthodes HMR et IMP est de l'ordre de 50 minutes et 5 minutes, respectivement (tests effectués sous Linux avec un PC Pentium III, 800Mhz).

## 4.2.2 Mise en correspondance d'images de coeur

Nous avons généré un ensemble de 100 transformations aléatoires de la même façon que pour le cas des images de thorax. Les rotations autour des trois axes sont limitées à  $\pm 5$  degrés et les translations à  $\pm 3$  cm. Les méthodes de mise en correspondance IMP et HMR sont évaluées. Dans un premier temps, nous avons tenté le recalage direct des images mais celui-ci a échoué systématiquement. Nous avons alors procédé à une pré-orientation de l'image TEP en petit axe à partir des informations d'orientation (angles) des coupes contenues dans l'entête des images par RM. Puis, nous avons isolé dans les deux images une région d'intérêt contenant le coeur (figure 4.5). Les images ont été interpolées linéairement et la taille du voxel isotrope est de  $1\text{ mm}^3$  pour une grille de  $161 \times 161 \times 126$  voxels. La région sélectionnée a été positionnée dans une boîte noire afin d'éviter les effets de bord lors de l'application des transformations géométriques.

Les images TEP de coeur sont issues des deux simulations d'un coeur sain et pathologique (voir chapitre 3).

### Recalage

La figure 4.6 est un exemple d'évaluation visuelle en damier pour le recalage d'images par RM et TEP d'émission de coeur avec les méthodes HMR+IM et HMR+CC. La représentation des images après le recalage avec la méthode HMR+CC (figure 4.6 (e), voir les zones fléchées) montre que le recalage n'a pas abouti à un alignement correct.



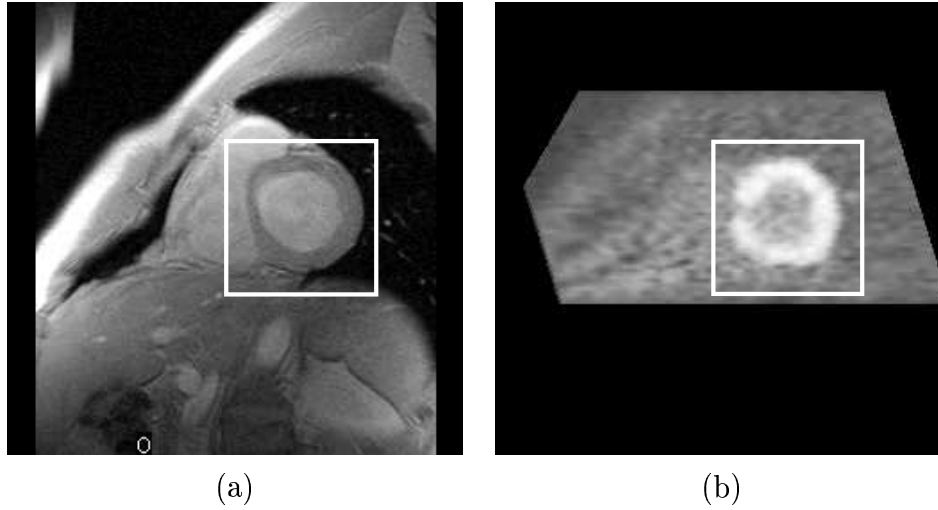


FIG. 4.5: Image du coeur sain en vue petit-axe avec la région d'intérêt contenant le ventricule gauche (a) coupe IRM, (b) coupe TEP simulée correspondante

### Test d'hypothèses

Pour vérifier si l'erreur  $d_{RMS}$  issue du recalage par les différentes méthodes est significativement inférieure à 2.43 mm (épaisseur des coupes pour les données TEP de coeur), on définit l'hypothèse nulle et l'hypothèse alternative de la manière suivante :

$$T_1 : \begin{cases} H0 : E[d_{RMS}] \geq 2.43mm \\ H1 : E[d_{RMS}] < 2.43mm \end{cases} \quad (4.6)$$

$$T_2 : \begin{cases} H0 : E[d_{RMS}] \leq 2.43mm \\ H1 : E[d_{RMS}] > 2.43mm \end{cases} \quad (4.7)$$

Les tests sont résolus à l'aide d'une statistique de Student, avec un *seuil de signification*  $\alpha$  de 1%. Les résultats du test de Student pour toutes les méthodes de recalage évaluées sont présentés dans les tableaux 4.3 et 4.4.

### Résultats de l'évaluation

Dans le contexte du recalage d'images par RM/TEP de coeur sain, les méthodes IMP, HMR+RC et HMR+IM sont les plus précises, avec des erreurs  $d_{RMS}$  de  $1.52 \pm 0.29$ ,  $1.91 \pm 1.31$  et  $0.85 \pm 0.88$  voxels, respectivement (tableau 4.3). Le test d'hypothèses  $T_2$  montre que la méthode HMR+CC ne permet pas d'obtenir une erreur inférieure à 2.43 mm, ( $d_{RMS} = 6.83 \pm 3.31$  voxels,  $pvalue < 0.0001$ ).

Globalement, la valeur médiane de  $d_{RMS}$  est très proche de la valeur moyenne. Ceci indique une répartition homogène des erreurs et le peu d'impact des éventuels points aberrants.

La dispersion des erreurs sur les paramètres de la transformation géométrique estimée, ainsi que sur les erreurs  $d_{RMS}$  pour les méthodes de mise en correspondance IMP et

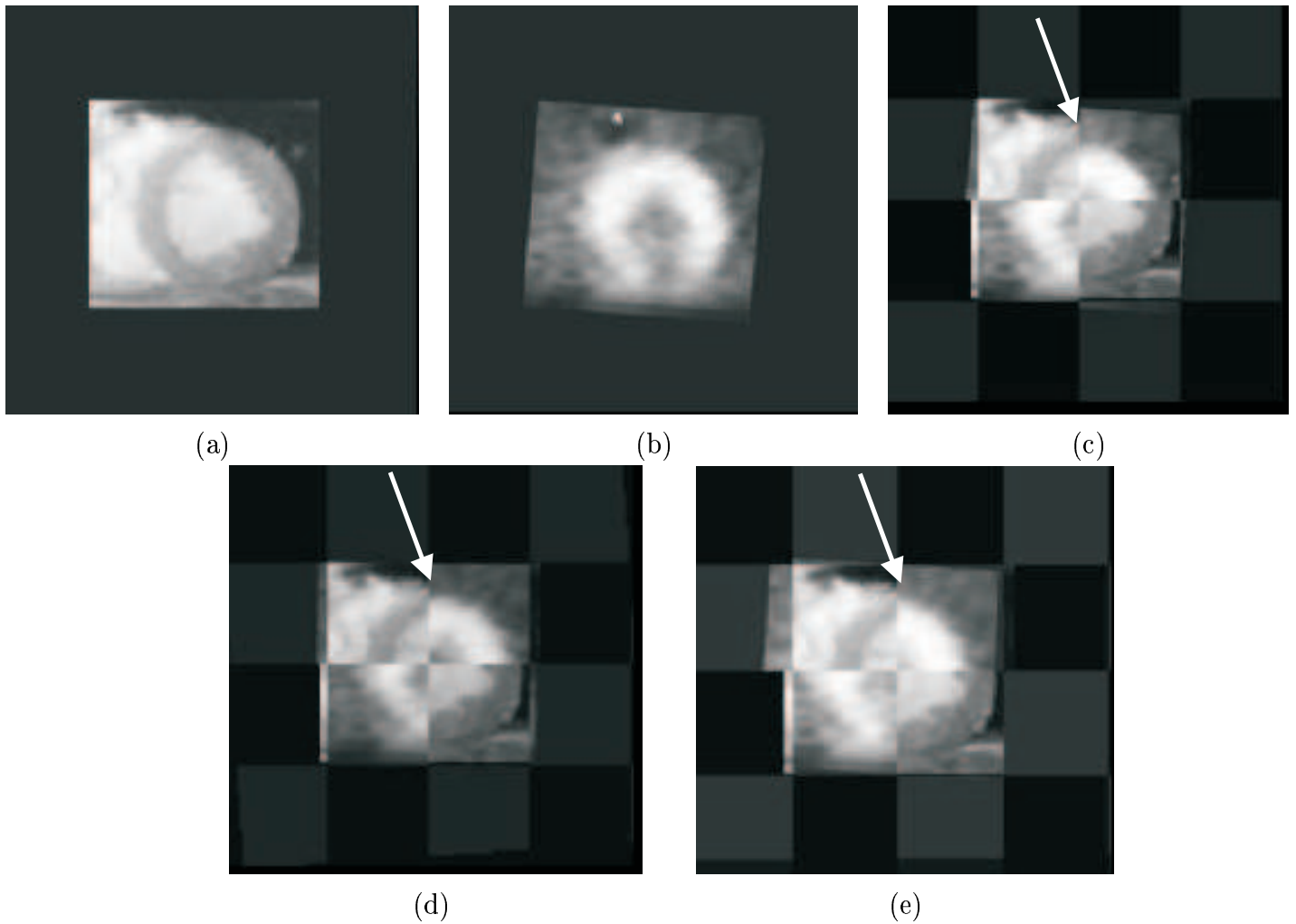


FIG. 4.6: Recalage d'images par RM et TEP d'émission de coeur : (a) image par RM, (b) image TEP transformée, visualisation en damier des volumes avant le recalage (c) et après le recalage avec les méthodes HMR+IM (d) et HMR+CC (e)

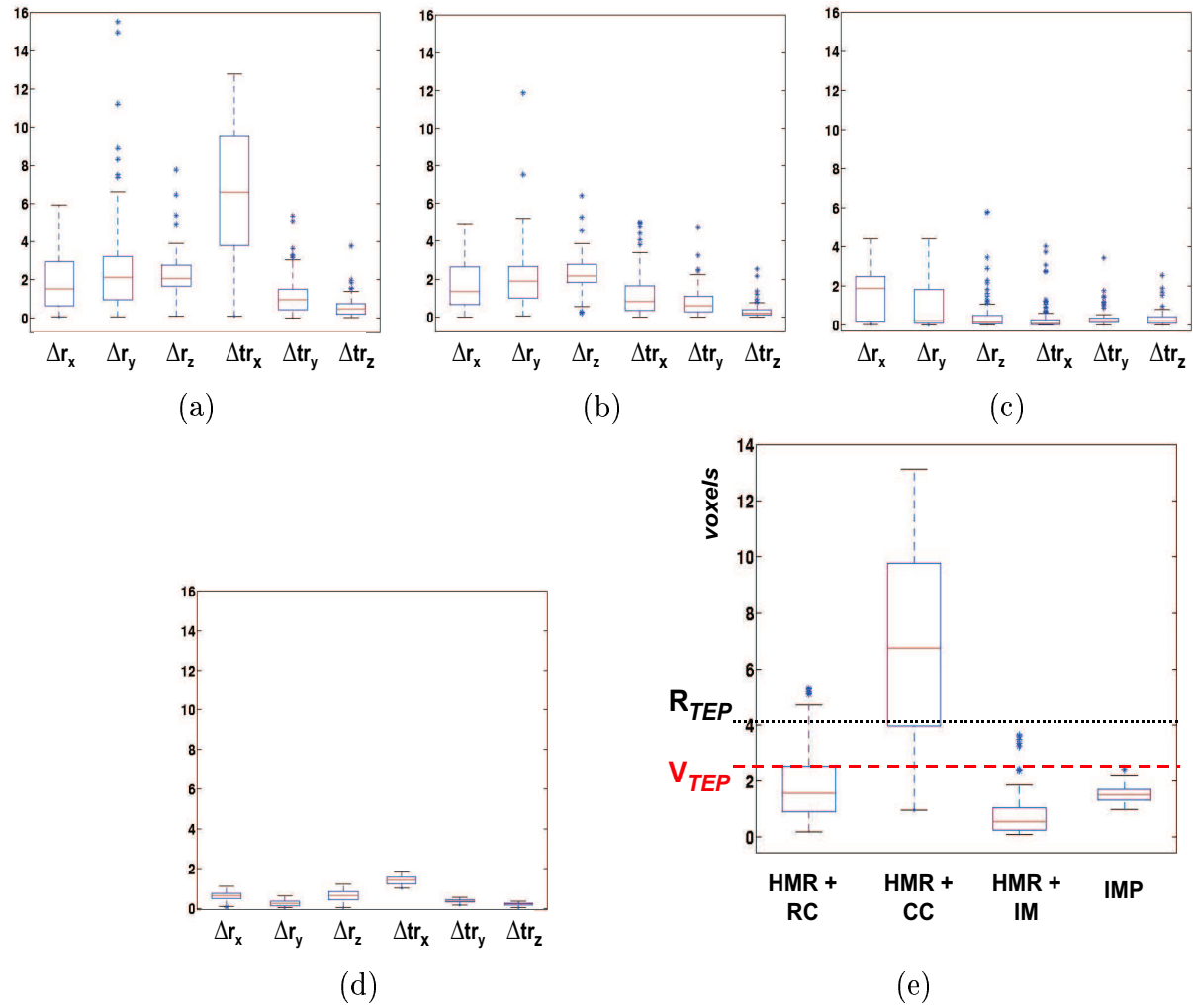


FIG. 4.7: Erreurs de recalage pour les images de coeur sain en utilisant la méthode de recalage HMR avec les mesures de similarité CC (a), RC (b), IM (c) et la méthode IMP (d). (h) erreurs  $d_{RMS}$  pour chaque méthode. Les erreurs en translation sont en voxels (1 voxel=1 mm) et les erreurs en rotation sont en degrés. La taille du voxel en TEP ( $V_{TEP}$ ) et la résolution physique de la caméra TEP sont respectivement égales à 2.43 mm et 4.1 mm.

HMR sont illustrées dans la figure 4.7 pour le coeur sain et la figure 4.8 pour le coeur pathologique.

Pour le cas de la mise en correspondance d'images cardiaques pathologiques, on retrouve les mêmes méthodes de recalage les plus performantes que pour le cas sain, les erreurs étant comparables. L'écart le plus important entre les erreurs issues des recalages des cas sain et pathologique est observé pour la méthode HMR+RC :  $1.91 \pm 1.31$  versus  $1.11 \pm 0.67$  voxels, respectivement. La méthodes IMP fournit des erreurs plus importantes selon la translation en X, tandis que la méthode HMR+RC donne des erreurs importantes en rotation.

Le temps de calcul pour une expérience de recalage d'images TEP/IRM de coeur avec les méthodes HMR et IMP est de l'ordre de 15 minutes et 3 minutes, respectivement (tests effectués sous Linux avec un PC Pentium III, 800Mhz).

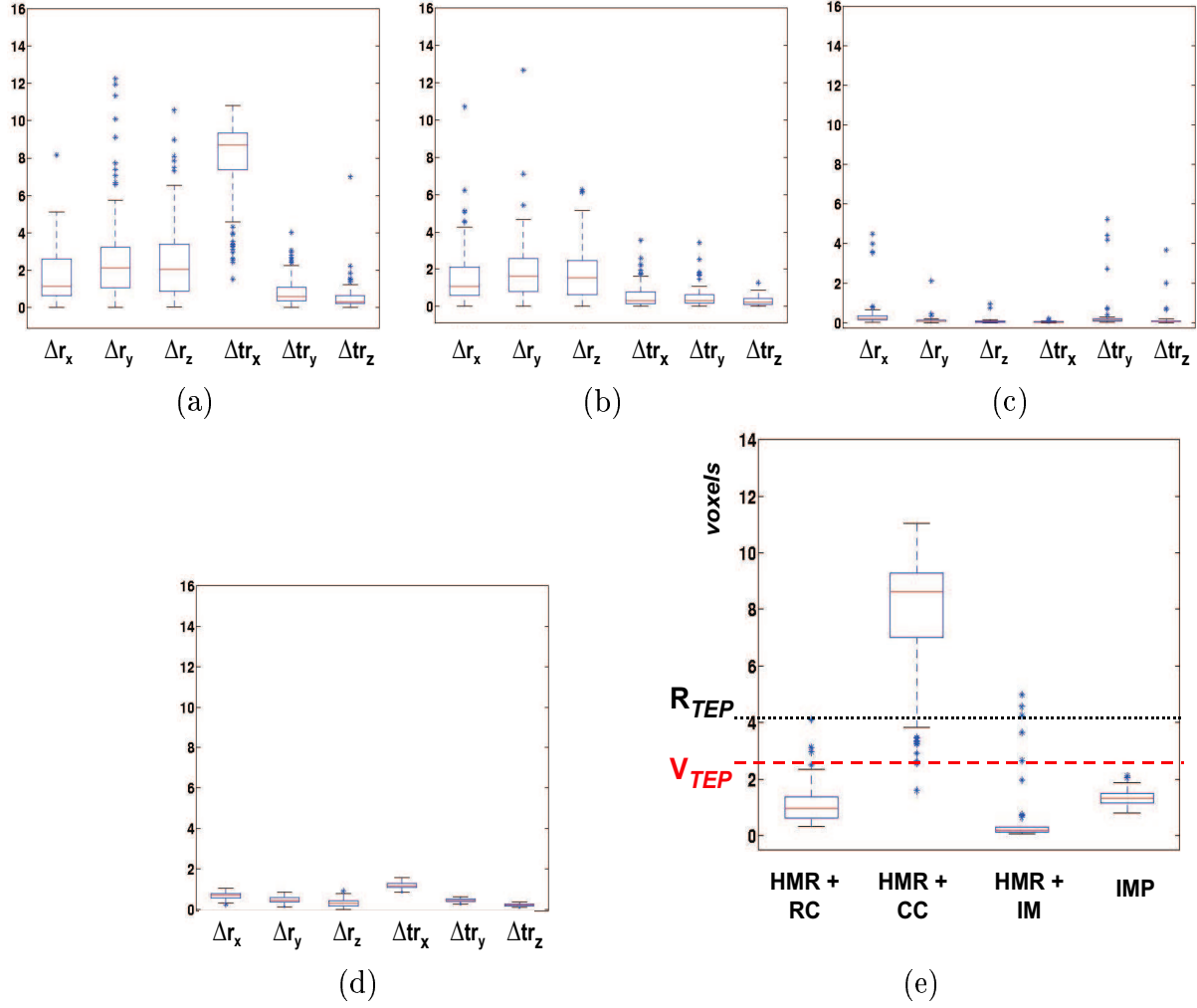


FIG. 4.8: Erreurs de recalage pour les images de coeur pathologique en utilisant la méthode de recalage HMR avec les mesures de similarité CC (a), RC (b), IM (c) et la méthode IMP (d). (h) erreurs  $d_{RMS}$  pour chaque méthode. Les erreurs en translation sont en voxels (1 voxel=1 mm) et les erreurs en rotation sont en degrés. La taille du voxel en TEP ( $V_{TEP}$ ) et la résolution physique de la caméra TEP sont respectivement égales à 2.43 mm et 4.1 mm.

erreur	IMP		SBR+Powell		SBR+Grid		SBR+Simplexe		HMR+IM		HMR+RC		HMR+CC	
$pvalue$	$< 0.0001(T_1)$		$< 0.0001(T_2)$		$< 0.0001(T_2)$		$< 0.0001(T_2)$		$< 0.0001(T_1)$		$< 0.0001(T_1)$		$0.001(T_1)$	
	moy $\pm$ SDV	med	moy $\pm$ SDV	med	moy $\pm$ SDV	med	moy $\pm$ SDV	med	moy $\pm$ SDV	med	moy $\pm$ SDV	med	moy $\pm$ SDV	med
$d_{RMS}$	$1.38 \pm 0.21$	1.34	$6.95 \pm 2.22$	7.25	$7.65 \pm 2.62$	7.79	$7.96 \pm 2.76$	7.67	$1.33 \pm 0.04$	1.34	$1.37 \pm 0.41$	1.27	$1.86 \pm 0.27$	1.91
$\Delta r_X$	$0.26 \pm 0.1$	0.24	$2.12 \pm 1.76$	1.75	$2.46 \pm 1.82$	2.09	$2.57 \pm 1.8$	2.34	$0.02 \pm 0.02$	0.02	$0.99 \pm 0.89$	0.8	$1.35 \pm 1.7$	0.76
$\Delta r_Y$	$0.3 \pm 0.1$	0.26	$2.88 \pm 2.13$	2.5	$3.2 \pm 1.92$	2.87	$2.97 \pm 2.18$	2.81	$0.02 \pm 0.01$	0.01	$1.36 \pm 1.27$	1.35	$1.28 \pm 1.2$	0.92
$\Delta r_Z$	$0.65 \pm 0.12$	0.63	$2.26 \pm 1.34$	2.04	$1.82 \pm 0.97$	1.67	$2.84 \pm 1.79$	2.79	$0.44 \pm 0.04$	0.44	$0.89 \pm 0.75$	0.69	$1.11 \pm 1.04$	0.8
$\Delta tr_X$	$1.09 \pm 0.07$	1.08	$3.32 \pm 2.21$	3.5	$3.26 \pm 2.54$	3.26	$3.51 \pm 2.73$	3.31	$0.86 \pm 0.06$	0.86	$0.71 \pm 0.36$	0.73	$0.64 \pm 0.39$	0.64
$\Delta tr_Y$	$0.7 \pm 0.05$	0.7	$3.16 \pm 2.57$	2.5	$3.39 \pm 2.68$	2.62	$3.53 \pm 2.79$	2.78	$1.02 \pm 0.05$	1.09	$0.85 \pm 0.26$	0.9	$1.62 \pm 0.25$	1.69
$\Delta tr_Z$	$0.07 \pm 0.06$	0.05	$3.9 \pm 2.41$	4	$4.58 \pm 2.95$	4.46	$4.69 \pm 3$	4.18	$0.05 \pm 0.03$	0.06	$0.2 \pm 0.19$	0.12	$0.28 \pm 0.27$	0.2

TAB. 4.2: Recalage d'images de thorax : valeurs moyenne (moy), écart-type (SDV) et médiane (med) des erreurs issues des différentes méthodes de recalage. Les erreurs en rotation ( $\Delta r_X$ ,  $\Delta r_Y$ ,  $\Delta r_Z$ ) sont en degrés et les erreurs en translation ( $\Delta tr_X$ ,  $\Delta tr_Y$ ,  $\Delta tr_Z$ ) et les  $d_{RMS}$  sont en voxels (taille du voxel 1.95 mm)

Erreurs	HMR+CC		HMR+RC		HMR+IM		IMP	
$pvalue$	$< 0.0001(T_2)$		$< 0.0001(T_1)$		$< 0.0001(T_1)$		$< 0.0001(T_1)$	
	moy $\pm$ SDV	med	moy $\pm$ SDV	med	moy $\pm$ SDV	med	moy $\pm$ SDV	med
$d_{RMS}(\text{voxels})$	$6.83 \pm 3.31$	6.72	$1.91 \pm 1.31$	1.54	$0.85 \pm 0.88$	0.55	$1.52 \pm 0.29$	1.51
$\Delta r_X(\text{degrés})$	$1.89 \pm 1.45$	1.52	$1.85 \pm 1.97$	1.33	$1.57 \pm 1.24$	1.87	$0.6 \pm 0.21$	0.6
$\Delta r_Y(\text{degrés})$	$2.65 \pm 2.63$	2.13	$2 \pm 1.59$	1.87	$0.94 \pm 1.16$	0.25	$0.23 \pm 0.14$	0.23
$\Delta r_Z(\text{degrés})$	$2.24 \pm 1.19$	2.07	$2.26 \pm 0.98$	2.17	$0.56 \pm 0.99$	0.18	$0.63 \pm 0.28$	0.63
$\Delta tr_X(\text{voxels})$	$6.61 \pm 3.4$	6.5	$1.26 \pm 1.26$	0.83	$0.38 \pm 0.81$	0.09	$1.4 \pm 0.2$	1.41
$\Delta tr_Y(\text{voxels})$	$1.18 \pm 1.05$	0.96	$0.78 \pm 0.76$	0.58	$0.35 \pm 0.46$	0.2	$0.36 \pm 0.08$	0.35
$\Delta tr_Z(\text{voxels})$	$0.57 \pm 0.52$	0.47	$0.3 \pm 0.38$	0.2	$0.35 \pm 0.41$	0.2	$0.19 \pm 0.07$	0.2

TAB. 4.3: Recalage d'images de coeur sain : valeurs moyenne (moy), écart-type (SDV) et médiane (med) des erreurs issues des différentes méthodes de recalage (taille du voxel 1mm)

Erreurs	HMR+CC		HMR+RC		HMR+IM		IMP	
$pvalue$	$< 0.0001(T_2)$		$< 0.0001(T_1)$		$< 0.0001(T_1)$		$< 0.0001(T_1)$	
	moy $\pm$ SDV	med	moy $\pm$ SDV	med	moy $\pm$ SDV	med	moy $\pm$ SDV	med
$d_{RMS}(\text{voxels})$	$7.79 \pm 2.22$	8.62	$1.11 \pm 0.67$	0.95	$0.63 \pm 1.2$	0.17	$1.33 \pm 0.24$	1.3
$\Delta r_X(\text{degrés})$	$1.72 \pm 1.43$	1.12	$1.75 \pm 2.3$	1.09	$0.54 \pm 1.01$	0.22	$0.66 \pm 0.15$	0.69
$\Delta r_Y(\text{degrés})$	$2.92 \pm 3.23$	2.09	$1.88 \pm 1.63$	1.65	$0.13 \pm 0.29$	0.07	$0.44 \pm 0.15$	0.44
$\Delta r_Z(\text{degrés})$	$2.58 \pm 2.17$	2.04	$1.82 \pm 1.48$	1.53	$0.08 \pm 0.16$	0.04	$0.3 \pm 0.2$	0.27
$\Delta tr_X(\text{voxels})$	$7.88 \pm 2.24$	8.7	$0.54 \pm 0.62$	0.33	$0.03 \pm 0.03$	0.02	$1.17 \pm 0.15$	1.15
$\Delta tr_Y(\text{voxels})$	$0.91 \pm 0.84$	0.62	$0.48 \pm 0.53$	0.31	$0.48 \pm 1.12$	0.1	$0.44 \pm 0.07$	0.44
$\Delta tr_Z(\text{voxels})$	$0.56 \pm 0.78$	0.34	$0.29 \pm 0.24$	0.2	$0.21 \pm 0.58$	0.07	$0.21 \pm 0.06$	0.21

TAB. 4.4: Recalage d'images de coeur pathologique : valeurs moyenne (moy), écart-type (SDV) et médiane (med) des erreurs issues des différentes méthodes de recalage (taille du voxel 1 mm)

## 4.3 Discussion

Nous avons évalué trois méthodes de mise en correspondance d'images thoraciques et cardiaques en imagerie multimodale IRM/TEP : une méthode surfacique (SBR) et deux méthodes iconiques (IMP et HMR).

Des transformations tests générées aléatoirement ont été appliquées sur l'image TEP afin de générer des jeux de données de référence pour la procédure d'évaluation. La comparaison des transformations géométriques tests et celles estimées par l'algorithme de recalage nous permet de quantifier les erreurs commises par chaque méthode.

Pour le recalage d'images de thorax par RM et TEP de transmission, les tests statistiques nous permettent de dire que la méthode surfacique ne remplit pas le critère de précision. Par contre, les méthodes HMR et IMP ont donné des résultats satisfaisants, à l'exception de HMR+CC où la précision semble satisfaisante mais l'incertitude plus grande par rapport aux autres méthodes. L'utilisation du Coefficient de Corrélation implique en effet une hypothèse de dépendance linéaire entre les intensités des images qui n'est pas vérifiée avec les modalités TEP et IRM. Le Coefficient de Corrélation est plutôt adapté au contexte du recalage monomodalité, IRM/IRM par exemple [Holden 00]. Il a cependant été utilisé en recalage multimodal d'images cérébrales, comme par exemple le recalage IRM/TEP [Studholme 96] avec des résultats satisfaisants. Nos résultats ne permettent pas de conclure clairement sur l'adaptation de HMR+CC au recalage d'images de thorax IRM/TEP. L'étude de la dispersion d'erreurs nous fournit des informations sur la robustesse de méthodes étudiées par rapport aux données d'entrée. La méthode HMR+IM présente la plus faible dispersion et la plus grande précision (figure 4.4 (h)) parmi l'ensemble des méthodes. Les méthodes HMR+IM et IMP présentent la dispersion la plus faible sur les erreurs des six paramètres de la transformation géométrique comme le montrent les figures 4.4 (d) et (g), respectivement. Les méthodes iconiques IMP, HMR+IM et HMR+RC ont des performances sensiblement équivalentes, avec une erreur  $d_{RMS}$  moyenne de l'ordre de 2.6 mm, donc largement inférieure à la taille du voxel en TEP ( $V_{TEP}=3.52$  mm), ainsi qu'à la résolution physique de la caméra TEP ( $R_{TEP}=4.1$  mm). La méthode de recalage surfacique conduit à des erreurs de l'ordre de 1.5 cm, donc très supérieures aux  $V_{TEP}$  et  $R_{TEP}$  (figure 4.4(h)).

Les résultats de recalage d'images cardiaques montrent l'échec de la méthode HMR+CC et le bon comportement des méthodes HMR+RC, HMR+IM et IMP qui, toutes trois, remplissent les conditions de précision, avec une erreur  $d_{RMS}$  moyenne en deçà de l'épaisseur de coupe en TEP ( $V_{TEP}=2.43$  mm) (figures 4.7(e) et 4.8 (e)). Les résultats obtenus sur les images de coeur sain et pathologique ne présentent pas de grands écarts en termes d'erreurs pour les méthodes iconiques évaluées. Cela peut être dû au fait que la région ischémique simulée n'est pas très significative : le nombre des voxels de la région ischémique est réduit par rapport à l'étendu du ventricule gauche (8.3% du nombre total des pixels du ventricule gauche). Cela est confirmé par les valeurs très proches des mesures de similarité (tableau 4.5) calculées entre les couples d'images IRM/TEP. De même, les mesures de similarité calculées entre les deux simulations TEP de coeur sain et pathologique recalées

sont proches des valeurs maximales : CC=0.93, RC=0.95 et IM=0.73. Les moyennes des valeurs des mesures de similarité issues de toutes les expériences de recalage menées sur les images de coeur sain et pathologique sont présentées dans le tableau 4.6. Ces valeurs sont proches des valeurs des mesures de similarité avant d'appliquer les transformations tests. Notons que le fait que les mesures avant et après recalage soient proches n'implique pas forcément un bon recalage. Il est indispensable de contrôler les valeurs de la  $d_{RMS}$  et des erreurs sur les paramètres de la transformation géométrique.

	sain	pathologique
CC	0.65	0.69
RC	0.8	0.82
IM	0.61	0.61

TAB. 4.5: Valeurs des mesures de similarité pour les couples d'images par RM/TEP recalées de référence dans les cas du coeur sain et pathologique

	sain	pathologique
CC	$0.67 \pm 0.01$	$0.64 \pm 0.15$
RC	$0.8 \pm 0.19$	$0.75 \pm 0.01$
IM	$0.59 \pm 0.02$	$0.58 \pm 0.02$

TAB. 4.6: Valeurs moyenne et écart-type des mesures de similarité pour les couples d'images par RM/TEP recalées dans les simulations du coeur sain et pathologique pour l'ensemble des expériences de recalage avec la méthode HMR

Les figures 4.7 (e) et 4.8 (e) montrent que, globalement, les erreurs  $d_{RMS}$  pour la méthode HMR sont moins dispersées dans les expériences de recalages avec des images de coeur pathologique que de coeur sain. La méthode HMR+CC présente la plus grande dispersion des erreurs  $d_{RMS}$ , suivie par HMR+RC, pour toutes les expériences de recalages. Comparée aux autres méthodes, HMR+IM dans le cas pathologique présente la plus faible dispersion avec cependant le nombre le plus élevé de cas exceptionnels (moyennes et écarts-type pour le cas sain et pathologique de  $0.85 \pm 0.88$  et  $0.63 \pm 1.22$  voxels, respectivement, tableaux 4.3 et 4.4). La méthode IMP semble avoir un comportement similaire dans les deux cas, avec le plus faible nombre de points exceptionnels sur les erreurs  $d_{RMS}$  et les plus petites erreurs sur les paramètres de la transformation géométrique, mais avec une précision globale  $d_{RMS}$  inférieure à celle de la méthode HMR+IM.

Avant l'implantation de l'approche multi-résolutions de notre méthode de mise en correspondance HMR, nous avons travaillé sur des images en basse résolution (taille du voxel de 8 mm). Les méthodes étudiées convergeaient rapidement (temps de calcul de l'ordre de 5 minutes) mais dans beaucoup de cas le recalage échouait. Le calcul avec les images en haute résolution était très long (5 heures pour recalcr deux volumes de



dimensions 256\*256\*150 en utilisant l'interpolation cubique et un seuil de tolérance  $tol = 0.01$ ).

L'approche multi-résolutions à trois niveaux de décomposition que nous avons implantée a diminué significativement le temps de calcul. De plus, la précision obtenue avec les mesures de similarité RC et IM est nettement supérieure à celle obtenue en basse résolution (taille du voxel de 8 mm). Il est certain que les paramètres que nous avons utilisés pour chaque niveau de décomposition pourraient encore être adaptés afin de trouver le meilleur compromis précision/temps de calcul.

La méthode IMP offre un bon compromis entre précision et rapidité. C'est l'algorithme qui converge le plus vite. Cette méthode travaille en effet sur des sous images de taille réduite pour estimer la densité de probabilité inconnue (fenêtre de Parzen) plutôt que de calculer la mesure de similarité sur les images entières.

## 4.4 Limites et perspectives

D'après les critères de Lehmann [Lehmann 02] énoncés dans le chapitre 2, notre référence IRM/TEP pourrait être améliorée en abordant les aspects suivants :

- la *reproductibilité* de la référence : il ne doit pas y avoir d'étape interactive lors de sa conception. Pour notre cas, la segmentation manuelle de l'image par RM réelle constitue une source de variabilité intra- et inter-opérateurs. Pour s'affranchir de cet inconvénient, deux solutions peuvent être proposées : la première serait de simuler la TEP et l'IRM à partir d'un même modèle anatomique issu de la segmentation d'une image anatomique réelle ; la deuxième solution serait de quantifier l'erreur de segmentation due à la variabilité intra- et inter-opérateur.
- le *réalisme* de la référence :
  - afin d'obtenir des paires d'images IRM/TEP non-recalées, on a appliqué des transformations géométriques théoriques tests sur la TEP déjà simulée (à partir du modèle anatomique). Mais, pour améliorer le réalisme de nos références, on devrait réaliser une simulation TEP pour chaque transformation test appliquée au préalable au modèle anatomique.
  - l'image TEP que nous avons simulée est issue d'un modèle anatomique obtenu à partir d'une image par RM de thorax réelle où le témoin avait les bras le long du corps. Pour se situer dans les conditions cliniques, l'influence de la position du témoin pendant l'acquisition TEP, avec les bras au dessus de la tête, devrait être prise en compte. Pour cela, une image par RM d'un patient avec les bras au dessus de la tête fournira un modèle anatomique expérimental qui permettrait dans un premier temps d'évaluer la déformation entre les deux positions.
  - intégrer le mouvement dû à la respiration et aux battements du coeur. Des résultats préliminaires ont été déjà présentés dans [Pollari 04] où un modèle de respiration a été appliqué aux données de référence.
- la *validité* de la référence : pour une généralisation convenable de l'étude, le nombre

des références doit être suffisamment grand. Les simulations avec des paramètres différents ou les simulations des cas pathologiques avec un niveau variable de gravité peuvent fournir des informations utiles dans le cadre d'une procédure d'évaluation. Cela permettrait aussi de mener une étude de la robustesse des méthodes de mise en correspondance évaluées vis à vis des données d'entrée.

Une autre perspective de ce travail est d'étudier le comportement des méthodes de recalage évaluées avec des données cliniques. Il s'agit de la dernière étape de validation d'une méthode dans un contexte clinique qui exige une étude exhaustive sur des données réelles IRM/TEP.

# Conclusion

L'objectif de cette thèse était le développement d'une stratégie d'évaluation des méthodes de mise en correspondance rigides en imagerie multimodale IRM/TEP thoracique et cardiaque. Nous avons inscrit cet objectif dans un cadre plus large d'évaluation quantitative des méthodes de traitement d'images médicales.

Notre principale contribution réside dans la construction d'une référence pour l'évaluation des méthodes de mise en correspondance. Elle repose sur la simulation réaliste d'images TEP à partir d'images par RM réelles. Pour cela, nous construisons un modèle théorique de captation glucidique caractéristique du contexte clinique étudié (sujet sain et ischémique) à partir de mesures sur des examens réels. Les principaux processus physiques d'une acquisition TEP sont modélisés dans un simulateur de type Monte Carlo. Notre approche fournit ainsi des images TEP de transmission et d'émission simulées qui sont en parfaite correspondance avec les images par RM réelles. Un ensemble de transformations géométriques rigides de référence aléatoirement générées est appliqué sur l'image TEP afin d'obtenir une série de couples d'images IRM/TEP non-recalées. Les transformations estimées sur ces couples d'images peuvent être alors comparées aux transformations appliquées, et l'écart quantifié à l'aide de critères d'erreur.

Dans un deuxième temps, nous évaluons trois méthodes de mise en correspondance rigide IRM/TEP (chapitre 4) à l'aide de la stratégie développée; deux d'entre elles sont des approches iconiques basées sur les intensités des images (IMP et HMR), tandis que la troisième est une approche surfacique (SBR). Une des méthodes iconiques a été développée dans le cadre de cette thèse (HMR). Des tests statistiques ont été utilisés afin de répondre à la question posée par l'objectif de validation dans le contexte clinique de l'étude de la viabilité myocardique. Le travail d'évaluation réalisé nous a permis une première évaluation objective de la précision des méthodes étudiées, de leurs avantages et de leurs limitations. Nous montrons que les méthodes de recalage IMP, HMR+RC et HMR+IM sont adaptées pour le recalage d'images IRM/TEP de thorax et de coeur.

Ce travail réalisé dans le contexte spécifique du recalage d'images met en évidence la difficulté de concevoir une stratégie rigoureuse d'évaluation et le besoin de définir une terminologie et une méthodologie communes pour l'évaluation. La standardisation des protocoles d'évaluation et leur adoption par la communauté scientifique faciliteraient la

comparaison des processus de traitements d'images médicales (e.g. recalage) ainsi que leur exploitation dans un contexte clinique et leur éventuelle valorisation industrielle.

Les perspectives de cette recherche sont d'ordre essentiellement méthodologique. Tout d'abord, l'intégration du mouvement dû à la respiration et aux battements du coeur au modèle utilisé en entrée du simulateur TEP SORTEO doit fournir une référence plus proche de la réalité. L'extension vers le recalage non-rigide est incontournable pour améliorer les résultats de mise en correspondance d'images réelles de thorax et de coeur.

# Annexe A

## Modalités d'acquisition d'images médicales

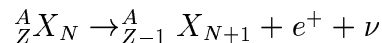
Nous insistons plus particulièrement sur la Tomographie par Émission de Positons et l'Imagerie par Résonance Magnétique qui sont les deux modalités considérées dans cette étude.

### A.1 Tomographie par Émission de Positons

Comme son nom l'indique, la Tomographie par Émission de Positons (TEP) repose sur l'utilisation d'isotopes émetteurs de positons. La tomographie par émission de positons permet d'obtenir, in vivo chez l'homme, la cartographie au sein des organes d'un paramètre physiologique comme le métabolisme du glucose ou le débit sanguin. Cette cartographie est obtenue à partir de la mesure de la distribution volumique et temporelle d'un traceur spécifique injecté au sujet. Le principe de la TEP est assez simple : un traceur radioactif, appelé radio-pharmaceutique, est injecté en quantité infime dans la circulation sanguine. Selon la molécule traceuse, une fonction de l'organisme va être explorée. Grâce à l'imageur 3D et après traitement informatique, les zones denses en radioactivité sont détectées. Ces zones correspondent aux régions possédant une forte concentration du traceur .

#### A.1.1 Émission de positons

L'existence du positon a été proposée pour la première fois par Dirac, à la fin des années 20, et prouvée expérimentalement en 1932. Le positon est l'antiparticule de l'électron et possède la même masse et une charge opposée. L'émission d'un positon  $e^+$  par un atome entraîne pour ce dernier la perte d'un proton et le gain d'un neutron et conduit également à l'émission d'un neutrino :



La masse du neutrino étant très faible, il interagit peu avec les autres particules. Cependant sa présence induit une variabilité de l'énergie du positon émis, contrairement

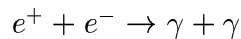
Isotope	$C^{11}$	$N^{13}$	$O^{15}$	$F^{18}$
Demi-vie (minutes)	20	9.97	2.1	110
Energie cinétique la plus probable (MeV)	0.39	0.49	0.73	0.25
Libre parcours moyen dans l'eau (mm)	1.1	1.5	2.7	0.6
Fraction de branchement (%)	99.8	99.8	99.9	96.7

TAB. A.1: Caractéristiques physiques des principaux émetteurs de positons utilisés en TEP. Le libre parcours moyen est déterminé par l'énergie d'émission du positon et constitue une limite en terme de résolution spatiale.

à l'émission  $\gamma$  d'énergie fixe pour un radionucléide donné. Les isotopes les plus communs en TEP sont présentés dans le tableau A.1.

### A.1.2 Annihilation de positons

Les positons subissent les mêmes interactions que les électrons, en particulier la perte d'énergie par ionisation et excitation des molécules et atomes voisins. Après avoir perdu suffisamment d'énergie sur une distance de l'ordre du millimètre (selon l'énergie initiale), le positon s'annihile avec un électron voisin :



La collision matière/antimatière entraîne la disparition des deux particules et la création des deux photons  $\gamma$  emportant l'énergie totale du système. L'énergie totale d'une particule se décompose en son énergie cinétique liée à sa vitesse et son énergie au repos liée à sa masse. Dans le processus d'annihilation, l'énergie provient presque exclusivement des masses de l'électron et du positon, ceux-ci étant quasiment au repos lors de l'interaction. Les lois de conservation de l'énergie totale et de la quantité de mouvement impose que cette énergie (1.022 MeV) soit partagée équitablement entre les deux photons, émis chacun dans des directions diamétralement opposées (figure A.1 (a)).

### A.1.3 Détection en coïncidence

L'émission simultanée des deux photons est à la base de la TEP. Pour recueillir le maximum de paires de photons gamma, la caméra TEP est formée de plusieurs couronnes de blocs détecteurs (figure A.1 (b)). Ces couronnes délimitent un cylindre de détection sur un champ de vue d'une quinzaine de centimètres. Les blocs détecteurs sont constitués de cristaux scintillateurs ayant la capacité d'émettre des milliers de photons secondaires lors de l'absorption d'un photon  $\gamma$  de 511 KeV. Ces photons secondaires sont peu énergétiques et dans le domaine du visible. Ce faible signal lumineux est amplifié par photomultiplication et converti en signal électrique qui peut alors être traité de manière électronique. Si un photon  $\gamma$  est également détecté sur un autre cristal dans un intervalle de temps de quelques nanosecondes (appelé *fenêtre de coïncidence*), la caméra considère ces deux signaux comme étant la conséquence de la même annihilation d'un positon et d'un élec-

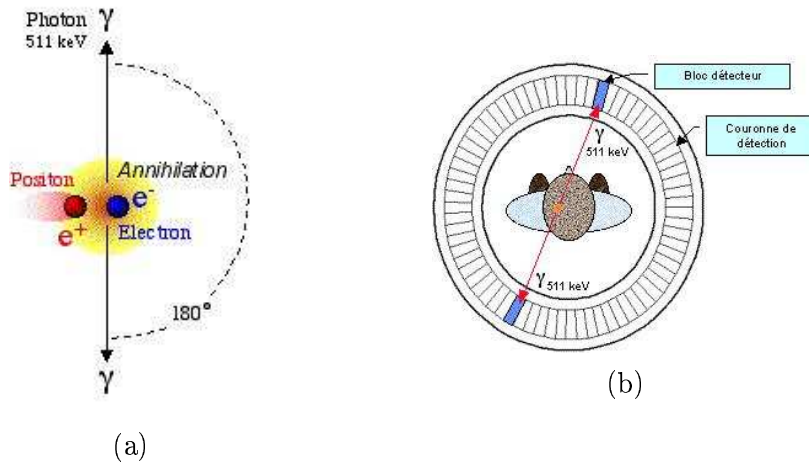


FIG. A.1: (a) annihilation électron/positon, (b) détection en coïncidence des photons

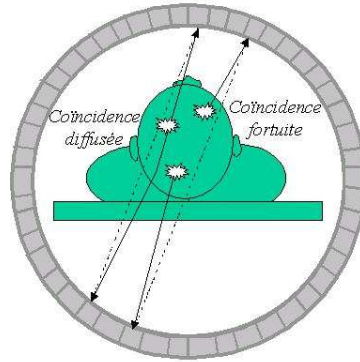


FIG. A.2: Coïncidences fortuite et diffusée

tron. La détection quasi-simultanée de ces deux photons  $\gamma$  est appelé *coïncidence*. Les deux photons ont donc parcouru une droite et ont fini leur course sur les cristaux scintillateurs. Le segment de droite reliant deux cristaux scintillateurs en vis-à-vis est appelé *ligne de réponse*. Le comptage des coïncidences permet de connaître l'intégrale de la radioactivité présente sur la ligne de réponse considérée.

### A.1.4 Projections

L'acquisition TEP consiste à compter le nombre de fois où chaque paire de détecteur est atteint en coïncidence. Pour un anneau de  $n$  cristaux, il y a  $n^2/2$  façon d'apparier les détecteurs, ce qui donne un ordre d'idée de la masse considérable d'information recueillie. La collecte de tous les évènements sur les lignes de réponse liant les milliers de cristaux du système de détection de la caméra TEP (18 432 cristaux pour une caméra Ecat HR+) constitue l'information brute que l'on enregistre dans des fichiers appelés *sinogrammes*. Un *sinogramme* est une matrice dont la  $i^{\text{ème}}$  ligne correspond au  $i^{\text{ème}}$  angle d'acquisition, les pixels de cette ligne contenant le nombre de coïncidences de chaque ligne de réponse

en 2D, perpendiculaire à cette direction. Toute l'information peut être représentée par un sinogramme sur  $180^\circ$ .

La connaissance des désintégrations sur chaque *ligne de réponse* revient à connaître les projections de l'image de radioactivité sur un grand nombre de plans. De ces projections, on peut reconstruire l'image de distribution spatiale de la radioactivité qui a été mesurée. On passe donc, à l'aide d'algorithmes de reconstruction (de type rétro-projection filtrée), des *sinogrammes* à des coupes jointives décrivant un volume, chaque pixel de ces images représentant la mesure locale de radioactivité sur la durée de l'acquisition.

### A.1.5 Facteurs de dégradation des images

#### Diffusé

La diffusion des photons en TEP conduit à deux types d'erreur de localisation de l'annihilation. Si l'un des photons est dévié de sa trajectoire initiale par interaction Compton avant de sortir du patient, il sera capté par un cristal autre que celui attendu pour l'impact initial, du même plan de coupe ou d'un autre plan de coupe, ce qui conduit à des lignes de réponse virtuelles (figure A.2). Il est possible de réduire la diffusion "extra-coupe" par l'utilisation de septa en plomb ou en tungstène qui éliminent géométriquement les radiations extérieures au champ de vue de l'anneau du détecteur. Ces septa agissent comme les collimateurs d'une gamma-caméra et permettent de réduire le diffusé de 30% à 50% mais en réduisant nettement la sensibilité de la caméra. En pratique, le diffusé est estimé à partir de la radioactivité détectée en dehors de l'objet émetteur ou par simulation à partir de la mesure de transmission.

#### Absorption

En TEP, les deux photons émis doivent quitter le corps du patient sans être absorbés pour être détectés. L'effet le plus évident de l'absorption est la diminution du nombre de coups captés, ce qui limite la précision de la quantification de la densité radioactive. Il faut également tenir compte d'une non-uniformité du signal, le milieu du corps ayant une probabilité plus grande d'être atténué. Cependant, la mesure de l'absorption est indépendante de la profondeur de l'émission. On mesure l'absorption due au tissu par l'acquisition préalable d'une image de transmission, en faisant tourner une source de rayons de 511 KeV autour du patient.

#### Coïncidences fortuites

Comme l'annihilation a rarement lieu exactement au centre du champ de vue, les photons peuvent ne pas arriver en même temps sur les détecteurs. Il n'existe pas de détecteurs capable d'avoir la résolution temporelle nécessaire à l'estimation du délai d'arrivée entre deux photons  $\gamma$  (ce délai est de quelques centaines de *pico – secondes*). Le choix de la *fenêtre de coïncidence* est un compromis entre le comptage de toutes les *coïncidences vraies* à l'exclusion des *coïncidences fortuites*. Ces dernières concernent deux photons qui



ne sont pas issus de la même annihilation (figure A.2). Les coïncidences fortuites augmentent le bruit de fond de l'image, notamment en acquisition 3D, avec acceptation de toutes les *lignes de réponse* (gains en sensibilité). Elles sont proportionnelles au carré de l'activité visible par le système de détection. Les coïncidences fortuites sont estimées en comptant pour chaque ligne de réponse le nombre de coïncidences observées dans une fenêtre temporelle décalée.

### Temps mort du cristal

Du fait de la durée de la scintillation d'un cristal, de la bande passante du système électronique de lecture et des tris des coïncidences, les détecteurs TEP possèdent un temps mort fonction du taux d'événements simples arrivant sur le système. Ce temps mort peut atteindre 30 à 60% en début d'acquisition.

En résumé, le nombre de coups mesurés  $M$  contient une portion de diffusés  $D$  et de fortuits  $F$  augmentant le bruit global de l'image. Les méthodes de correction permettent de retrouver les événements vrais  $V$  :

$$V = M - D - F,$$

mais à  $V$  constant il est préférable d'avoir  $D$  et  $F$  faibles dès l'acquisition plutôt que de recourir à une correction avec  $D$  et  $F$  élevés.

## A.2 Imagerie par Résonance Magnétique

L'imagerie par Résonance Magnétique Nucléaire (IRM) est une technique non invasive, d'une totale innocuité, basée sur le principe de la résonance des atomes de certaines molécules (en imagerie diagnostique : l'hydrogène) sous l'action de certaines ondes de radio-fréquences. Le phénomène de résonance magnétique nucléaire (RMN) a été observé par Purcell et Bloch en 1946 (Prix Nobel de Physique en 1952) et appliqué à l'anatomie humaine il y a une vingtaine d'années. L'IRM est utilisée aujourd'hui pour le diagnostic, et le traitement des patients. Il permet d'obtenir des images des organes ou des tissus internes sans introduire le moindre instrument dans le corps, évitant ainsi aux patients des souffrances et réduisant aussi les risques d'infections. Il est aussi notable que cette technique permet d'éviter d'irradier le corps par des rayons X, ou d'y injecter des substances radioactives. Le prix Nobel de Médecine 2003 a été remis conjointement au professeur Britannique Peter Mansfield, et au professeur Américain Paul Lauterburg pour leurs travaux concernant l'imagerie par résonance magnétique.

Selon le champ magnétique auquel ils sont soumis, les noyaux de certains des atomes qui composent la matière peuvent entrer "en résonance" : c'est le principe physique de la résonance magnétique nucléaire. Pour permettre l'enregistrement de ce phénomène, on envoie une onde de radio fréquence courte qui va modifier transitoirement l'orientation des spins qui tournent autour de ces noyaux activés. Lorsque les protons reviennent à leur état initial, ils restituent de l'énergie enregistrable sous forme d'un signal, capté par une antenne réceptrice, puis analysé par de puissants calculateurs. Un traitement informatique

permet de construire une image tridimensionnelle, présentée en coupes successives, dont les informations seront variables en fonction de la séquence d'imagerie utilisée (définie par un ensemble de paramètres). En observant, sous l'effet d'un champ magnétique intense, la résonance des noyaux d'hydrogène, élément présent en abondance dans l'eau et les graisses (80% du corps humain), on peut visualiser la structure anatomique de nombreux tissus (IRM anatomique). On peut suivre également certains aspects du métabolisme ou du fonctionnement des tissus (IRM fonctionnelle). La résonance des noyaux d'hydrogène du sang induite par la présence d'hémoglobine permet par exemple de suivre le trajet du sang dans le cerveau. Il est aussi possible d'accéder à des informations sur des molécules encore plus complexes et d'étudier le déroulement des réactions biochimiques dans un tissu donné : c'est la spectroscopie par résonance magnétique (SRM).

### A.2.1 Principe physique de la résonance magnétique nucléaire

Le principe est basé sur la propriété de résonance des noyaux (hydrogène) de la matière.

Soumis à un champ magnétique statique intense  $B_0$  (régnant dans le tunnel de l'imager), les protons s'orientent selon la direction de ce dernier en précessant à la vitesse angulaire  $\omega_0$  donnée par la relation de Larmor :

$$\omega_0 = \gamma B_0 \quad (\text{A.1})$$

où  $\omega_0$  est la fréquence angulaire de résonance (ou la fréquence angulaire de Larmor) et  $\gamma$  le rapport gyromagnétique. Pour  $B_0 = 0.5$  Tesla (champ "moyen" en imagerie, mais tout de même égal à 10.000 fois le champ magnétique terrestre), la fréquence de Larmor est de 21MHz pour les tissus biologiques (en imagerie protonique) ; elle est située dans le domaine de fréquence des *ondes radiophoniques*. L'alignement selon  $B_0$  peut être parallèle (Sud-Nord) où antiparallèle (Nord-Sud), ces deux états correspondant à des niveaux discrets d'énergie respectifs  $E_p$  et  $E_a$  tels que  $E_p > E_a$ . Les protons se répartissent sur ces deux niveaux selon une distribution de Boltzmann :

$$\frac{N_p}{N_a} = \exp\left(\frac{E_a - E_p}{kT}\right) \quad (\text{A.2})$$

$N_p$  (resp.  $N_a$ ) étant le nombre de protons dans l'état d'énergie  $E_p$  (resp.  $E_a$ ),  $k$  la constante de Boltzmann et  $T$  la température absolue. De plus :

$$E_a - E_p = \gamma \frac{h}{2\pi} B_0 \quad (\text{A.3})$$

$h$  désignant la constante de Planck. Les protons sont donc plus nombreux dans la position parallèle et génèrent un vecteur d'aimantation noté  $M$  de même sens que celui du champ magnétique  $B_0$ . En choisissant pour la direction du champ magnétique l'axe  $z$ , les deux composantes longitudinale  $M_z$  et transversale  $M_{xy}$  sont telles que  $M_z = \| M \|$  et  $M_{xy} = 0$ .

L'application d'un champ magnétique supplémentaire  $B_1$ , induit par un courant sinusoïdal exciteur à la fréquence de Larmor  $\gamma_0$  du proton, modifie l'aimantation globale : le vecteur  $M$  bascule d'un angle d'autant plus élevé que l'excitation est intense et longue.

En général, l'angle de basculement (*flip angle*) vaut  $90^\circ$  ou  $180^\circ$ . L'impulsion RF associée permettant de faire basculer la composante  $M_z$  est dite longitudinale et possède une direction perpendiculaire à celle du champ  $B_0$ .

La suppression de l'excitation s'accompagne d'un retour de l'aimantation  $M$  à sa position d'équilibre : la composante longitudinale  $M_z$  croît vers sa valeur d'équilibre  $M$  et la composante transversale  $M_{xy}$  décroît selon un mouvement de spirale à la vitesse  $\omega_0$ . Ce phénomène de relaxation est gouverné par les équations de Bloch :

$$\begin{cases} M_z = M(1 - e^{-\frac{t}{T_1}}) \\ M_{xy} = M e^{-\frac{t}{T_2}} \end{cases} \quad (\text{A.4})$$

$T_1$  et  $T_2$  étant respectivement les temps caractéristiques de récupération pour la composante  $M_z$  et de décroissance pour la composante  $M_{xy}$ .  $T_1$  est une mesure de la quantité de réalignement des protons sur le champ magnétique externe  $B_0$  ; c'est pourquoi il est appelé *temps de relaxation spin-réseau*.  $T_2$  mesure la quantité de réalignement des protons entre eux ; il est également appelé *temps de relaxation spin-spin*. La diminution de  $M_{xy}$  induit dans une bobine réceptrice un signal sinusoïdal amorti ou FID (*Free Induction Decay*) à la fréquence de résonance  $\omega_0$ , qui est le signal RMN. L'énergie ainsi émise est proportionnelle à la densité de spins du volume étudié. Le spectre de ce signal est obtenu par transformée de Fourier.

### Codage de position des protons

#### La sélection d'une coupe : Codage spatial dans une direction

Soumis au champ magnétique  $B_0$ , tous les protons du volume résonnent à la fréquence  $\omega_0$ . Une discrimination spatiale selon l'axe  $z$  entre ces protons est effectuée en appliquant dans cette direction un gradient de champ magnétique  $G_z = B_0 + g \cdot z \cdot i_z$ ,  $g$  étant une constante et  $i_z$  le vecteur orthonormé directeur selon  $z$  : les protons précessent ainsi à une vitesse  $\omega = \gamma(B_0 + gz)$  proportionnelle à leur ordonnée  $z$  (figure A.3). L'application d'une excitation sélective à la fréquence du plan de coupe souhaité permet dès lors d'obtenir le signal RMN correspondant aux points de cette coupe. L'étape suivante consiste à coder les points du plan sélectionné dans les deux directions restantes.

#### Lecture du plan image : Codage de phase et codage de fréquence

Après sélection du plan de coupe perpendiculairement à l'axe  $z$ , le codage des points de ce plan en  $x$  et en  $y$  est réalisé par l'application successive de deux gradients de champ magnétique  $G_x$  et  $G_y$ .

Pendant l'application du gradient  $G_y$ , les protons précessent à des fréquences différentes, proportionnelles à leur abscisse  $y$  et donc se déphasent. Après suppression de  $G_y$ , les protons précessent de nouveau à la même fréquence, tout en conservant leur déphasage. L'application ultérieure du gradient  $G_x$  entraîne une précession des protons à des fréquences différentes, proportionnelles à leur ordonnée  $x$  : chaque proton peut dès lors

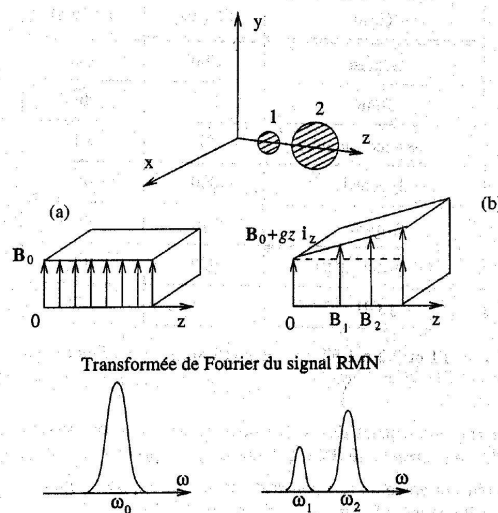


FIG. A.3: Principe du codage spatial dans une direction (a) application d'un champ magnétique uniforme, (b) application d'un gradient de champ magnétique (issu de [Kastler 97])

être caractérisé par sa phase et sa fréquence. La fréquence et la phase de précession sont respectivement proportionnelles aux coordonnées  $x$  et  $y$ . Le gradient  $G_y$  induit ainsi un codage de phase et le gradient  $G_x$  un codage de fréquence.

En pratique, le gradient  $G_y$  est appliqué avant la lecture du signal et le gradient  $G_x$  pendant la lecture du signal. Le temps séparant la première impulsion longitudinale de la lecture du signal est appelé temps d'écho (TE). Le codage de phase  $G_y$  doit être répété autant de fois que l'on désire de lignes, avec une période appelée temps de répétition (TR).

### Influence des constantes physiques et des paramètres d'acquisition

Pour un élément de volume donné, de multiples facteurs conditionnent l'intensité du signal RMN acquis. Il s'agit :

- de paramètres intrinsèques, liés à la nature même des tissus contenus dans ce volume : densité de protons, temps de relaxation  $T_1$  et  $T_2$  ;
- de paramètres extrinsèques, liés à l'appareil de mesure et à la séquence d'acquisition : intensité du champ principal  $B_0$  qui définit en partie l'intensité du signal RMN, nature et grandeurs caractéristiques des séquences utilisées (TE, TR ...).

L'intensité du signal en un point de l'image correspond à la moyenne des intensités des signaux dans les volumes échantillons contenus dans ce point. Ce moyennage tient compte de l'épaisseur de la coupe et peut ainsi être responsable d'effets de volume partiel.

Le choix des paramètres  $T_R$  et  $T_E$  permet d'obtenir des images plus ou moins dépendantes de  $T_1$  ou  $T_2$  (qui sont caractéristiques d'un milieu donné : ils sont plus courts dans un solide que dans un liquide, reflétant le fait que les échanges d'énergie sont plus rapides dans un solide).

La valeur du temps de répétition TR séparant deux impulsions longitudinales à  $90^\circ$  détermine la pondération du contraste en  $T_1$ , en limitant la croissance de la composante

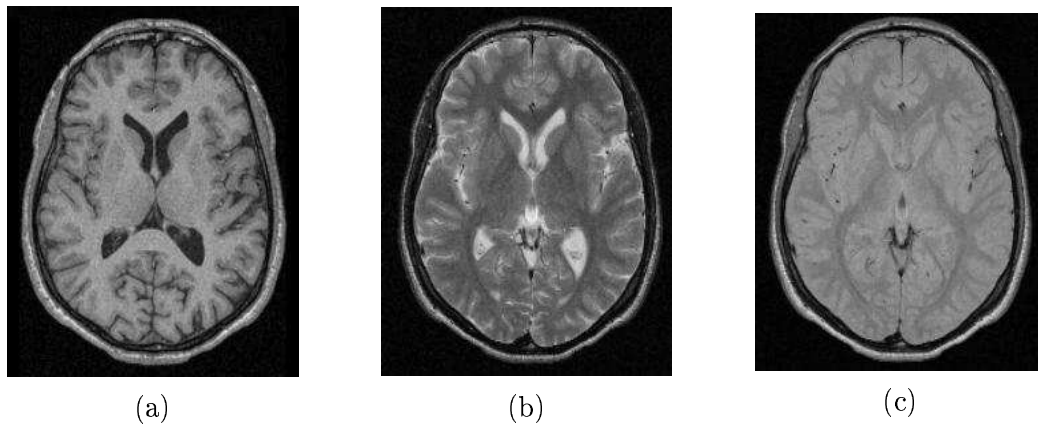


FIG. A.4: Trois images par résonance magnétique d'un même niveau de coupe pondérées en (a) T1, (b) T2, (c) densité protonique

longitudinale. La valeur du temps d'écho TE détermine la pondération du contraste en  $T_2$ , exprimant les variations de la composante transversale. Le contraste en  $T_1$  est d'autant plus fort que TR est court et le contraste en  $T_2$  est d'autant plus fort que TE est long. Un compromis reste à réaliser entre un fort contraste et un bon signal sur bruit : celui-ci est en effet d'autant meilleur que TR est long.

La figure A.4 montre des coupes d'images par résonance magnétique en utilisant différentes séquences IRM.

Pour plus de détails concernant le phénomène de résonance magnétique et ses applications, le lecteur peut de référer au livre de B. Kastler [Kastler 97].

## A.3 Les ultrasons

L'imagerie ultrasonore, où échographie, a été inspirée des sonars. Le principe consiste à émettre une onde ultrasonore à travers le corps humain. L'onde se propage dans les milieux homogènes que constitue chaque tissu mais elle est partiellement réfléchi à l'interface entre les tissus. Les ondes ultrasonores sont produites par des céramiques piézo-électriques vibrant sous l'action d'une tension alternative appliquée à leur bornes. La détection du temps mis par le rayon réfléchi pour revenir à la sonde émettrice permet d'estimer la profondeur des interfaces connaissant la vitesse moyenne de propagation des ultrasons à travers le corps humain. Réciproquement, à la réception, ces mêmes céramiques permettent de transformer l'onde ultrasonore incidente en courant électrique (figure A.5).

Pour construire une image 2D, plusieurs rayons sont acquis par la sonde dans des directions différentes et l'image résultante est interpolée. L'acquisition d'images ultrasonores est peu coûteuse et d'une innocuité totale. Cependant, les images obtenues présentent un phénomène de *speckle* (ou scintillement) provoqué par des phénomènes d'interférence des ondes diffusées à travers les tissus traversés. En outre le traitement et la numérisation de l'onde introduisent un bruit dans l'image. L'atténuation des bruits ultrasonores

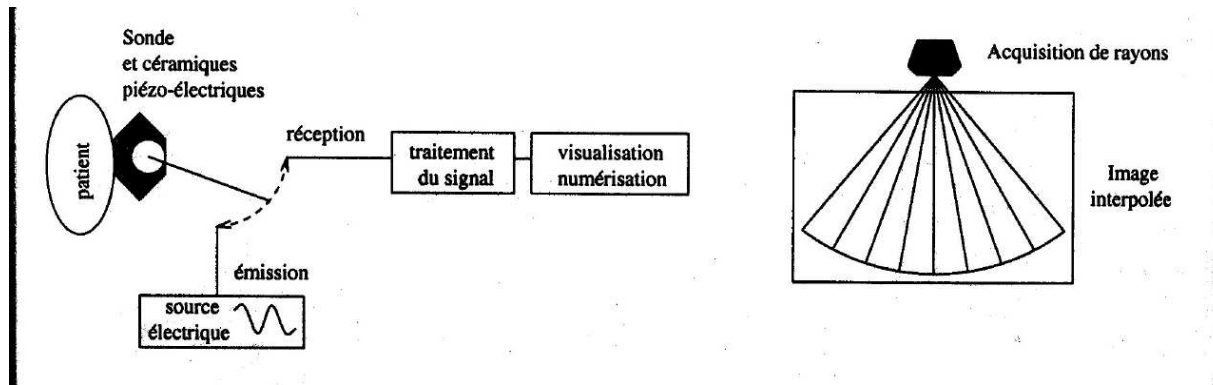


FIG. A.5: Système d'acquisition d'images ultrasonores

dans le corps rend difficile la perception d'organes profonds. Les os et l'air (des poumons) sont particulièrement opaques aux ultrasons et peuvent provoquer des occultations dans l'image.

Des produits de contraste sous forme de micro-bulles d'air enfermées dans des sucres permettent d'améliorer le contraste de certaines images, des vaisseaux sanguins et des cavités cardiaques en particulier.

Des systèmes d'acquisition 3D ultrasonores arrivent aujourd'hui sur le marché. Ils sont fondés sur l'acquisition de plusieurs plans pour constituer un volume. Le mouvement de la sonde est contrôlé mécaniquement ou électroniquement pour réaliser des acquisitions de plans géométriquement bien définis dans l'espace.

## A.4 Les rayons X

### La radiographie conventionnelle

Les images radiographiques, ou images par rayons X, furent les premières images médicales. Les rayons X, découverts par W. C. Roentgen en 1895, gardent actuellement un intérêt diagnostique de premier plan dans beaucoup de domaines de la médecine. Ils peuvent traverser le corps humain, où ils sont plus ou moins atténués suivant la densité électronique des structures traversées. Les rayons résiduels (ceux qui auront traversé le corps) provoquent le noircissement du film placé derrière la table de radiographie (technique radiographique traditionnelle). Ainsi, une structure "aérée" comme celle des poumons paraîtra noire. A l'inverse, une structure dense comme les os paraîtra blanche (les rayons X auront tous été absorbés). Aujourd'hui, les films radiographiques sont remplacés par des détecteurs électroniques, dont les différents points, stimulés ou non par les rayons résiduels, permettent une numérisation et donc un traitement informatique des images obtenues en radiographie.

## La tomographie à rayons X

La tomographie à rayons X, appelée également scanner à rayons X (en anglais “computerized tomography” ou CT), est apparue à la fin des années 60 (G. N. Hounsfield et A. M. Cormack ont reçu le Prix Nobel de Médecine en 1968).

Dérivée important de la radiographie, elle utilise le balayage (d'où le choix du terme scanner, littéralement “balayeur” en anglais) par rayons X pour visualiser un organe en coupes, et l'interprétation immédiate par ordinateur des densités radiologiques obtenues (la densitométrie). Cette technique est également appelée tomodensitométrie ou tomographie X assistée par ordinateur (la zone étudiée peut être “découpée en tranches”, d'où le nom “tomographie”, *tomein* signifiant *couper* en grec). Le scanner permet de pallier l'inconvénient majeur de la radiographie classique qui n'offre “qu'une vue en projection” du volume irradié.

Le principe du fonctionnement du scanner X est illustré dans les figures A.6 et A.7. En faisant tourner la source des rayons X et les détecteurs et en réalisant de multiples acquisitions au cours de la rotation, on construit une coupe 2D plan par plan. Une reconstruction 3D est possible par empilement de coupes mais également avec les scanners hélicoïdaux pour lesquels le déplacement de la source se fait simultanément dans le plan  $xy$  et selon l'axe  $z$ , selon une trajectoire en hélice (voir figure A.7).

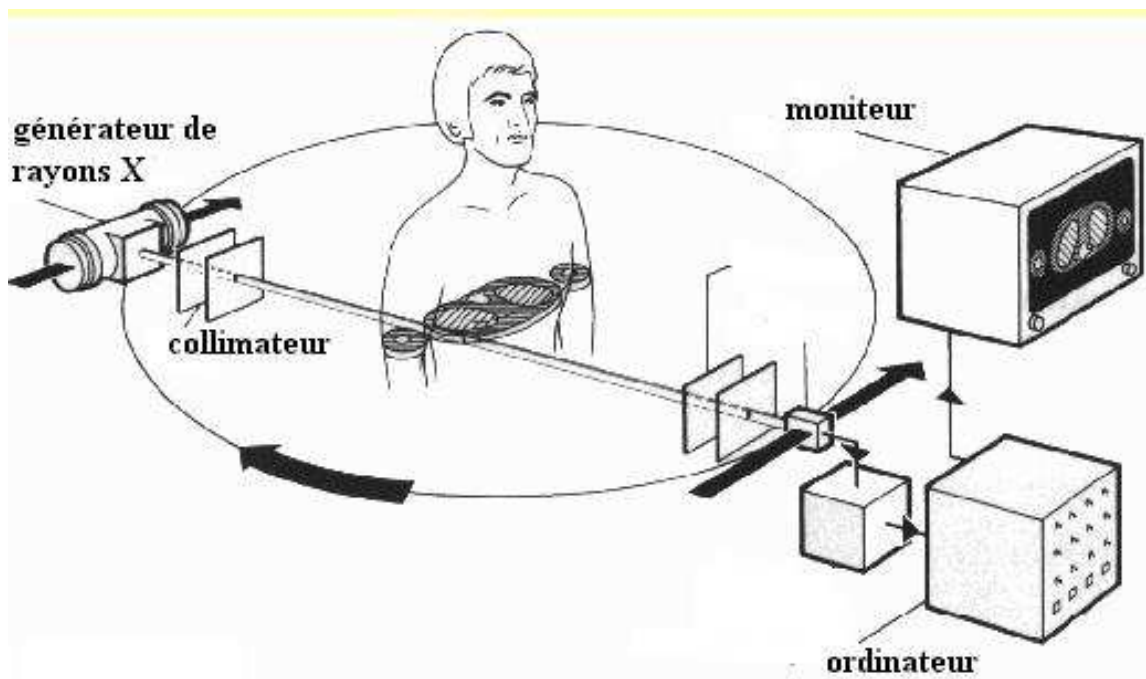


FIG. A.6: Principe d'acquisition en tomodensitométrie

Les scanners hélicoïdaux permettent l'acquisition plus rapide d'un volume au détriment de la qualité puisqu'une interpolation plus importante est nécessaire à la reconstruction des coupes. Le temps d'acquisition par un scanner hélicoïdal d'un volume correspondant à une centaine de plans sur une quinzaine de centimètres est de l'ordre de

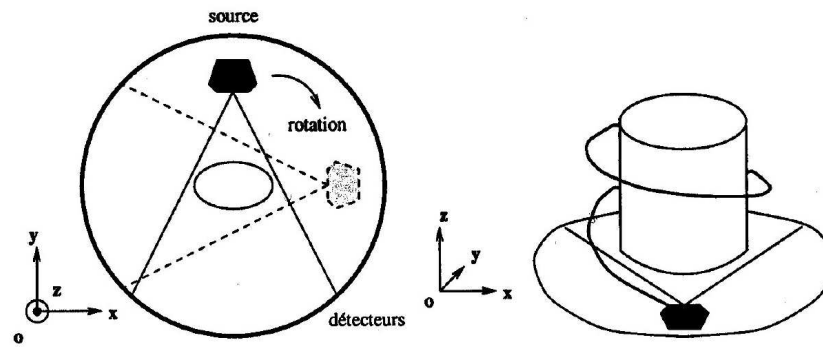


FIG. A.7: Tomographie à rayons X : parcours de l'espace en rotation simple et avec une acquisition spiralée par scanner hélicoïdal

30 secondes contre quelques minutes pour les scanners classiques.

En radiographie ou en tomographie par rayons-X, des substances radio-opaques (agents de contraste) peuvent être injectées dans le corps pour améliorer la visualisation de certains organes.



# Annexe B

## Mesures de Similarité

Il existe un nombre considérable de mesures de similarité dans la littérature du recalage d'images. Il s'avère que la plupart des mesures existantes peuvent être définies à partir de l'histogramme conjoint de l'image de référence et de l'image flottante [Hill 94, Sarrut 00].

Plusieurs mesures peuvent être appréhendées indépendamment de la notion d'histogramme conjoint. Celle-ci implique une interprétation probabiliste, mais d'autres interprétations sont possibles. De plus, le calcul effectif d'un histogramme conjoint peut accroître inutilement la complexité calculatoire d'une mesure. C'est le cas, en particulier, pour toutes les mesures basées sur des différences d'images.

Nous allons présenter les mesures de similarité qui s'inscrivent dans le formalisme de l'histogramme conjoint.

### L'histogramme conjoint

Une mesure de similarité relative à une transformation est calculée à partir des images. Étant donnée une classe de transformation, une procédure d'optimisation est appliquée afin de maximiser cette mesure de similarité dans l'espace de transformation considéré. Généralement, une image dite de *référence*, est fixe et les transformations sont appliquées à l'autre image, désignée par l'image *flottante*. Nous considérons les notations suivantes :

- $I_r = \Pi_{i=1}^d [1...n_i] \rightarrow D_r$  est l'image de *référence*,
- $I_f = \Pi_{i=1}^d [1...n_i] \rightarrow D_f$  est l'image *flottante*,  
 $d$  désigne la dimension des images (2D ou 3D),  $n_i$  représente le nombre d'échantillons suivant le  $i^{\text{ème}}$  axe.  $D_r$  et  $D_f$  sont les domaines des intensités des images, typiquement 256 niveaux de gris.
- la transformation spatiale est notée  $T : \mathbb{R}^d \rightarrow \mathbb{R}^d$
- $S$  est la mesure de similarité

L'histogramme conjoint  $H_T$  est défini en fonction de  $T$  comme le tableau de contingence des couples d'intensités associés aux voxels appariés :

$$H_T : D_r \times D_f \rightarrow \mathbb{R}^+$$

À la fin de la procédure de calcul,  $H_T(i, j)$  représente le nombre de voxels, tels que  $I_f(x) = j$  et  $I_r(T(x)) = i$ . Le calcul est effectué de manière itérative et à l'aide d'une procédure d'interpolation. Nous notons  $n_{ij}$  la valeur de  $H_T(i, j)$ ,  $n_i = \sum_j n_{ij}$  et  $n_j = \sum_i n_{ij}$  les effectifs marginaux, et  $n = \sum_i \sum_j n_{ij}$  la taille totale (voir Figure B.1).

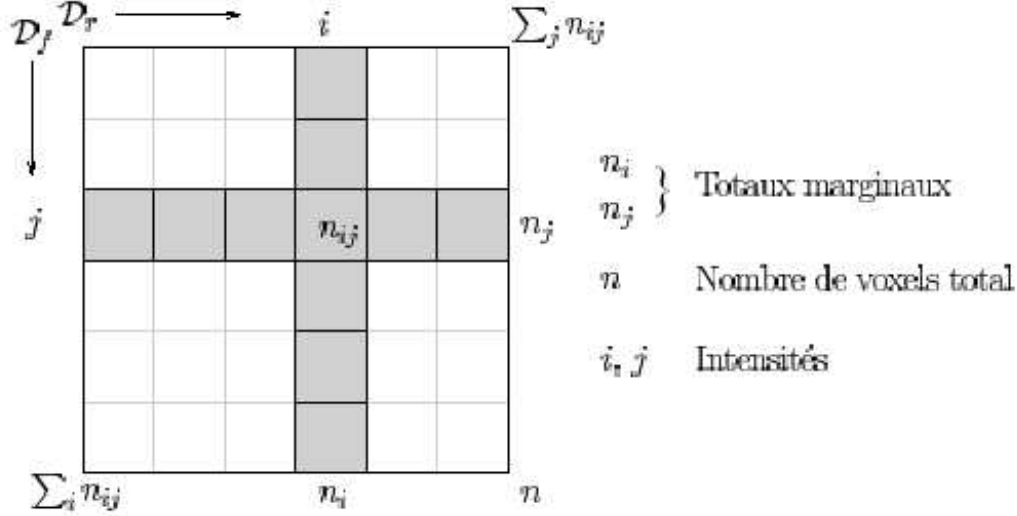


FIG. B.1: Notations pour un histogramme joint d'après [Sarrut 00]

Les quantités  $n_i, n_j, n_{ij}$  sont ensuite normalisées afin d'obtenir des *distributions*, élémentaires et marginales. Nous les désignerons par  $p_i, p_j, p_{ij}$  ( $p_{ij} = \frac{n_{ij}}{n}$ ).

## Mesures statistiques de similarité

Dans le contexte décrit précédemment, une mesure de similarité est une fonction à valeurs réelles dont l'argument est l'histogramme conjoint, lui-même fonction de la transformation spatiale. Le fondement commun aux nombreuses mesures proposées dans la littérature est l'idée que les intensités de deux images manifestent une cohérence d'autant plus forte que les images sont bien alignées. Le rôle de la mesure de similarité est précisément de donner une signification quantitative à cette notion de cohérence.

## Notations

Quelques notations préalables doivent être introduites (moyennes, écarts types, moments, variances) :

Moyennes :  $m_I = \sum_i i p_i$ ,  $m_J = \sum_j j p_j$

Moyennes conditionnelles :  $m_{I|j} = \frac{1}{p_j} \sum_i i p_{ij}$ ,  $m_{J|i} = \frac{1}{p_i} \sum_j j p_{ij}$

Écarts types :  $\sigma_I^2 = \sum_i (i - m_I)^2 p_i$ ,  $\sigma_J^2 = \sum_j (j - m_J)^2 p_j$

Écarts types conditionnels :  $\sigma_{I|j}^2 = \frac{1}{p_j} \sum_i (i - m_{I|j})^2 p_{ij}$ ,  $\sigma_{J|i}^2 = \frac{1}{p_i} \sum_j (j - m_{J|i})^2 p_{ij}$

Covariance :  $Cov(I, J) = \sum_i \sum_j (i - m_I)(j - m_J)p_{ij}$   
 Variances :  $Var(I) = \sigma_I^2$ ,  $Var(J) = \sigma_J^2$

## Coefficient de Corrélation

Le *coefficient de corrélation* mesure une relation linéaire (affine plus précisément) entre les intensités. Il est défini par :

$$\rho(I, J) = \frac{\sum_i \sum_j (i - m_I)(j - m_J)p_{ij}}{\sqrt{\sigma_I^2 \sigma_J^2}} \quad (B.1)$$

soit

$$\rho^2(I, J) = \frac{Cov^2(I, J)}{Var(I)Var(J)} \quad (B.2)$$

Ce coefficient peut être interprété comme la variation de la distribution par rapport à la droite de régression des moindres carrés. Ainsi,  $\rho^2$  est compris entre 0 et 1. C'est un critère symétrique, les images jouant le même rôle. Il est à maximiser, un coefficient de corrélation de 1 indiquant une parfaite correspondance.

## Rapport de Corrélation

Cette mesure de similarité a été récemment proposée par Roche *et al.* dans [Roche 98]. Le *rapport de corrélation* est défini comme suit :

$$\eta^2(I | J) = 1 - \frac{1}{\sigma^2} \sum_j p_j \sigma_{I|j}^2 \quad (B.3)$$

Ce critère mesure l'intensité d'une relation fonctionnelle entre les intensités. De plus, la variance peut être décomposée en la somme de deux termes positifs : la variance des moyennes conditionnelles et la moyenne des variances conditionnelles. En notant  $E[X]$  l'espérance (ou la moyenne) d'une distribution  $X$ , nous avons alors [Sarrut 00] :

$$\begin{cases} Var(I) = Var(E[I | j]) + E[Var(I | j)] = \sum_j p_j (m_{I|j} - m_I)^2 + \sum_j p_j \sigma_{I|j}^2 \\ Var(J) = Var(E[J | i]) + E[Var(J | i)] = \sum_i p_i (m_{J|i} - m_J)^2 + \sum_i p_i \sigma_{J|i}^2 \end{cases} \quad (B.4)$$

Ainsi,  $\eta^2$  peut être défini de la manière suivante :

$$\eta^2(I | J) = \frac{Var(E[I | j])}{Var(I)} \quad (B.5)$$

ou

$$\eta^2(J | I) = 1 - \frac{E(Var[I | j])}{Var(I)} \quad (B.6)$$

Ce critère est égal à 1, son maximum, lorsque la *moyenne des variances conditionnelles* est nulle. Ce n'est pas un critère symétrique, une image doit être choisie comme étant un *modèle* de l'autre (en termes de prédiction).

## Information Mutuelle

Dans le cas multimodal, les critères statistiques basés sur l'entropie des images ont fait leur preuve [Pluim 03]. L'entropie de Shannon d'une image  $I$  peut être définie par :

$$H(I) = -\sum_i p_i \log p_i \quad (\text{B.7})$$

L'entropie conjointe de deux images  $I$  et  $J$  est définie par :

$$H(I, J) = -\sum_i \sum_j p_{ij} \log \frac{p_{ij}}{p_i p_j} \quad (\text{B.8})$$

L'*information mutuelle* [Wells 96, Maes 97] est un critère de similarité très utilisé dans le contexte du recalage multimodal. L'*information mutuelle*  $IM$  entre deux distributions  $I$  et  $J$  de niveaux de gris, est donnée par :

$$IM(I, J) = H(I) + H(J) - H(I, J) \quad (\text{B.9})$$

La version *normalisée* de l'information mutuelle (IMN) s'écrit :

$$IMN(I, J) = \frac{H(I) + H(J)}{H(I, J)} \quad (\text{B.10})$$

# Annexe C

## Matrices des transformations rigides 3D

### C.1 Angles d'Euler

Afin de présenter la rotation sous forme matricielle, nous utilisons la représentation en coordonnées homogènes. Une transformation rigide 3D est une composition de transformations élémentaires de translation et de rotation où les translations sont définies sous forme matricielle par :

$$T_x = \begin{bmatrix} 1 & 0 & 0 & x_{trans} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}; T_y = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & y_{trans} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}; T_z = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & z_{trans} \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

La matrice de rotation est classiquement représentée par les angles d'Euler, c'est à dire par des rotations successives autour des trois axes principaux  $X, Y, Z$  :

$$R_x = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(\alpha) & \sin(\alpha) & 0 \\ 0 & -\sin(\alpha) & \cos(\alpha) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}; R_y = \begin{bmatrix} \cos(\beta) & 0 & -\sin(\beta) & 0 \\ 0 & 1 & 0 & 0 \\ \sin(\beta) & 0 & \cos(\beta) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix};$$
$$R_z = \begin{bmatrix} \cos(\gamma) & \sin(\gamma) & 0 & 0 \\ -\sin(\gamma) & \cos(\gamma) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

L'ordre dans laquelle les transformations sont appliquées est important (le produit matriciel n'est pas commutatif). Une rotation de  $\alpha$  autour de l'axe  $X$  suivie d'une translation suivant  $Z$  ne donne pas le même résultat que la translation suivie de la rotation.

## C.2 Représentation des quaternions

L'inconvénient de la représentation des rotations avec les angles d'Euler est que dans certains cas on perd un degré de liberté (effet *Gimbal Lock*). Par exemple, supposons qu'un objet est tourné dans l'ordre  $Z, Y, X$  et que la rotation autour de l'axe  $Y$  est de  $90^\circ$ . Après la rotation suivant l'axe  $Y$ , les axes  $X$  et  $Z$  sont confondus. Ainsi, chaque rotation lors de la composition d'Euler dépend des rotations précédentes. Les quaternions éliminent ce problème. Au lieu de tourner un objet via une série de rotations successives, les quaternions permettent de tourner un objet autour d'un axe arbitraire et d'un angle quelconque. La rotation est cependant toujours réalisée par des calculs matriciels. Néanmoins, au lieu de multiplier les matrices, les quaternions les représentant sont multipliés. Le résultat final est reconverti en la matrice désirée.

Un quaternion  $q$  est un élément de l'espace à 4 dimensions constituant une généralisation des nombres complexes. Comme mentionné plus haut, la représentation par angles d'Euler requiert la définition de l'ordre dans lequel sont appliquées les rotations. Au contraire, la représentation par quaternions permet d'exprimer une rotation complexe quelconque par une transformation unique traduisant une rotation autour d'un axe.

Un quaternion  $q$  peut être exprimé sous la forme :  $q = q_0 + \mathbf{i}q_1 + \mathbf{j}q_2 + \mathbf{k}q_3$  et son conjugué  $\bar{q} = q_0 - \mathbf{i}q_1 - \mathbf{j}q_2 - \mathbf{k}q_3$ , où  $\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = -1$ ,  $\mathbf{ij} = -\mathbf{ji} = \mathbf{k}$ ,  $\mathbf{jk} = -\mathbf{kj} = \mathbf{i}$ ,  $\mathbf{ki} = -\mathbf{ik} = \mathbf{j}$ . L'amplitude d'un quaternion est :  $\|q\| = \sqrt{q \star \bar{q}} = \sqrt{q_0^2 + q_1^2 + q_2^2 + q_3^2}$ , où " $\star$ " représente la multiplication des quaternions ; le quaternion unitaire :  $\|q\| = 1 \Rightarrow q^{-1} = \bar{q}$ .

Le produit des quaternions est associatif mais il n'est pas commutatif ; l'inverse d'un quaternion est son conjugué.

On peut représenter un quaternion sous la forme d'un scalaire pour sa partie réelle et un vecteur pour les termes imaginaires :

$$q = (a, v), \text{ où } a = q_0 \text{ et } v = [q_1 q_2 q_3]$$

Ainsi, le produit des deux quaternions  $p = (a, v)$  et  $q = (b, w)$  dans la représentation vectorielle est défini par :

$$p \star q = (ab - v \cdot w, aw + bv + v \times w)$$

### Formalisme angle/axe

Le quaternion qui effectue une rotation d'angle  $\theta$  autour du vecteur unitaire  $u$  s'exprime :

$$q = (a, v) \text{ avec } a = \cos \frac{\theta}{2} \text{ et } v = u \sin \frac{\theta}{2}$$

Un point  $m$  est représenté dans l'espace par le quaternion  $M = (0, m)$ . Une rotation d'angle  $\theta$  autour du vecteur unitaire  $u$  appliquée sur  $m$  est décrite par la formule suivante :

$$M_{rotated} = qMq^{-1}$$

### Matrice de rotation

La matrice de rotation correspondante aux angles d'Euler peut être exprimée à l'aide des composantes du quaternion  $q$  sous la forme :

$$Q = \begin{bmatrix} q_0^2 + q_1^2 - q_2^2 - q_3^2 & 2(q_1 q_2 + q_0 q_3) & 2(q_1 q_3 - q_0 q_2) \\ 2(q_1 q_2 - q_0 q_3) & q_0^2 - q_1^2 + q_2^2 - q_3^2 & 2(q_2 q_3 + q_0 q_1) \\ 2(q_1 q_3 + q_0 q_2) & 2(q_2 q_3 - q_0 q_1) & q_0^2 - q_1^2 - q_2^2 + q_3^2 \end{bmatrix}$$

Les angles d'Euler sont exprimés à partir de la matrice  $Q$  de la façon suivante :

$$\begin{aligned} \alpha &= \text{atan2}(Q(2, 3), Q(3, 3)) ; \\ \beta &= \text{asin}(-Q(1, 3)) ; \\ \gamma &= \text{atan2}(Q(1, 2), Q(1, 1)) ; \end{aligned}$$

# Annexe D

## Test d'hypothèses

Un test d'hypothèses consiste à définir une règle de décision concernant la validité d'une hypothèse portant sur la valeur caractéristique (moyenne, proportion, variance, ...) d'une variable aléatoire dans une population dont on observe un échantillon aléatoire.

La procédure générale d'un test requiert les éléments suivants :

- définir la caractéristique  $X$  étudiée, sa loi et les paramètres connus
- définir l'*hypothèse nulle* notée  $H_0$  : cette hypothèse porte sur la nature des valeurs expérimentales, la forme des lois de distribution ou les paramètres de ces lois. Elle suppose que les variations de ce que l'on étudie (moyenne) sont uniquement liées aux fluctuations aléatoires résultant de la variation de l'échantillonnage. C'est l'hypothèse qu'on veut rejeter lors d'un test d'hypothèses.
- définir l'*hypothèse alternative* notée  $H_1$  : c'est l'hypothèse que l'on retiendra si on ne peut pas conserver l'hypothèse nulle. Cette hypothèse suppose que le hasard de l'échantillonnage ne peut pas expliquer l'écart que l'on trouve par rapport aux valeurs attendues données par la loi de probabilité
- fixer *a priori* le *seuil de signification*  $\alpha$  (ou l'erreur de première espèce) qui représente la probabilité de rejeter  $H_0$  alors qu'elle est vraie
- définir un *critère statistique* dont on connaît la loi quand  $H_0$  est vraie
- définir une région critique telle que  $(1 - \alpha)$  soit la probabilité, si  $H_0$  est vraie, pour que le critère statistique appartienne à cette région
- énoncer la règle de décision correspondant à la valeur numérique prise par le critère statistique, à savoir rejeter  $H_0$  si ce critère est dans la région critique, ou accepter  $H_0$  si le critère est dans la région d'acceptation, région complémentaire de la précédente.

### D.1 Test d'hypothèses unilatéral à gauche pour une moyenne

Question : *Est ce bien d'une population de moyenne  $m > a$  qu'est issu l'échantillon observé de taille  $n$  et de moyenne  $\bar{x}$  ?*

On considère le cas d'un échantillon de taille  $n > 30$  (ou la loi de  $X$  est une loi normale) issu d'une population avec une écart-type  $\sigma$  inconnu. On formule le test d'hypothèses



unilatéral à gauche sur une moyenne, adapté à l'application testée dans ce manuscrit :

$$\begin{cases} H0 : m \geq a \\ H1 : m < a \end{cases} \quad (D.1)$$

On définit le *critère statistique calculé* :

$$t_c = \frac{\bar{x} - a}{\frac{\sigma}{\sqrt{n}}} \quad (D.2)$$

On compare le *critère statistique calculé*  $t_c$  à un *critère statistique théorique*  $t_t$  obtenu dans la table de la loi Normale pour la valeur  $(1 - \alpha)$ .

Si  $|t_c| < t_t$  : on garde  $H0$ .

Si  $|t_c| \geq t_t$  : on rejette  $H0$  avec  $\alpha\%$  chance de se tromper.

Une autre manière de tester des hypothèses est d'utiliser la valeur *pvalue*, qui est le *seuil descriptif* du test qui permet de juger dans quelle mesure les données permettent d'accepter l'hypothèse nulle  $H0$ . Pour la valeur *pvalue*, il faut trouver la probabilité associée à la valeur suivante :  $pvalue = P(t_t \leq t_c)$ . La valeur *pvalue* se lit sur la table de la loi normale. Ainsi, pour un seuil de signification  $\alpha$  :

- si  $pvalue > \alpha$ , on ne rejette pas  $H0$
- si  $pvalue \leq \alpha$ , on rejette  $H0$

*Discussion* : Dans le contexte d'un test unilatéral à gauche, si  $t_c \ll -t_t$ , on estime la  $pvalue < 0.0001$ , ce qui va nous conduire à rejeter l'hypothèse nulle  $H0$  (car  $pvalue < \alpha$ ). Par contre, si  $t_c \gg -t_t$ , on peut estimer qu'on doit accepter l'hypothèse nulle  $H0$  mais on ne peut pas estimer la valeur *pvalue*. Dans ce cas on doit formuler le test d'hypothèses à droite suivant :

$$\begin{cases} H0 : m \leq a \\ H1 : m > a \end{cases} \quad (D.3)$$

Pour plus de détails, le lecteur peut se référer à l'ouvrage de S. K. Kachigan [Kachigan 91].

# Bibliographie

- [Acharya ] ACHARYA R., WASSERMAN R., STEVENS J. et HINOHOSA C. *Computerized Medical Imaging and Graphics*, n ° 1.
- [Ackerman 95] ACKERMAN M. J., SPITZER V. M., SCHERZINGER A. L. et WHITLOCK D. G. The Visible Human data set : an image resource for anatomical visualization. *Medinfo*, 1995, vol. 8, n ° 2, pp 1995–1998.
- [Alderson 62] ALDERSON S. W., LANZL L. H., ROLLINS M. et SPIRA J. An instrumented phantom system for analog computation of treatments plans. *Am. J. Roentgenol. Radium Ther. Nucl. Med.*, 1962, vol. 87, pp 185–195.
- [Andrei 00] ANDREI A. Mise en oeuvre d'un protocole de simulation pour l'évaluation des méthodes de reconstruction d'images TEP corps entier en oncologie. *Rapport de stage, CERMEP, Lyon*, 2000.
- [Audette 00] AUDETTE M., FERRIE F. et PETERS T. An algorithmic overview of surface registration techniques for medical imaging. *Med. Image Anal.*, February 2000, vol. 4, n ° 4, pp 201–217.
- [Bacharach 93] BACHARACH S.L., DOUGLAS M.A., CARSON R.E., KALKOWSKI P.J., FREEDMAN N.M., PERRONE P. et BONOW R.O. Three-dimensional registration of cardiac positron emission tomography attenuation scans. *Computer Vision, Graphics and Image Processing*, 1993, vol. 34, n ° 2, pp 311–321.
- [Barillot 99] BARILLOT B. *Fusion de Données et Imagerie 3D en Médecine*. Thèse : Habilitation à diriger des Recherches, IRISA, Université de Rennes, France, 1999.
- [Behloul 01] BEHLOUL F., LELIEVELDT B. P. F., BOUDRAA A., JANIER M., REVEL D. et REIBER J. H. C. Neuro-fuzzy systems for computer-aided myocardial viability assessment. *IEEE Trans. Med. Imaging*, December 2001, vol. 20, n ° 12, pp 1302–1313.
- [Besl 92] BESL P. J. et MCKAY N. D. A method for registration of 3-D shapes. *IEEE Trans. Pattern Anal. Machine Intell.*, February 1992, vol. 14, n ° 2, pp 239–256.
- [Bettinardi 93] BETTINARDI V., GILARDI M.C., LUCIGNANI G., LANDONI C. et RIZZO G. A procedure for patient repositioning and compensation for

- misalignment between transmission and emission data in PET heart studies. *J. Nucl. Med.*, 1993, vol. 34, n° 1, pp 137–142.
- [Beyer 99] BEYER T., TOWNSEND D.W., BRUN T., KINAHAN P.E., CHARRON M., RODDY R. et JERIN J. A combined PET/CT scanner for clinical oncology. *J. Nucl. Med.*, 1999, vol. 41, n° 8, pp 1369–1379.
- [Bidaud 01] BIDAUD L.M et VALLEE J.-P. Automated registration of dynamic MR images for the quantification of myocardial perfusion. *Journal of magnetic resonance imaging*, 2001, vol. 13, pp 648–655.
- [Bland 86] BLAND J. et ALTMAN D. Statistical methods for assessing agreement between two methods of clinical measurements. *Lancet*, 1986, n° 1, pp 307–310.
- [Borgefors 88] BORGEFORS G. Hierarchical chamfer matching : A parametric edge matching algorithm. *IEEE Trans. Pattern Anal. Machine Intell.*, 1988, vol. 10, n° 6, pp 849–865.
- [Bowyer 98] BOWYER K. W. et PHILLIPS P. J. Overview of work in empirical evaluation of computer vision algorithms. IEEE Computer Press, 1998.
- [Bradely 97] BRADELY C.P. et PULLAN A.J. Geometric modeling of the human torso using cubic hermite elements. *Annals of Biomedical Engineering*, 1997, vol. 25, pp 96–111.
- [Brown 92] BROWN L.G. A survey of image registration techniques. *ACM Computing Surveys*, 1992, vol. 24, n° 4, pp 325–376.
- [Buvat 99] BUVAT I., CHAMEROY V., AUBREY F., PÉLÉGRINI M., FAKHRI G. EL, HUGUENIN C., BENALI H., POKROPEK A. T. et PAOLA R. DI. The need to develop guidelines for evaluations of medical image processing procedures. *SPIE Vol. 3661 Medical Imaging*. 1999. pp 1466–1477.
- [Buvat 02] BUVAT I. et CASTIGLIONI I. Monte Carlo simulations in SPECT and PET. *Journal of Nuclear Medicine*, 2002, vol. 46, pp 48–61.
- [Cai 99] CAI J., CHU J.C., RECINE D., SHARMA M., NGUYEN C., RODEBAUGH R., SAXENA A. et ALI A. CT and PET lung image registration and fusion in radiotherapy treatment planning using the chamfer-matching method. *Int. J. Radiation Oncology Biol. Phys.*, 1999, vol. 43, n° 4, pp 883–891.
- [Carrillo 01] CARRILLO A., DUERK J.L., LEWIN J.S. et WILSON D.L. Semiautomatic 3-D image registration as applied to interventional MRI liver cancer treatment. *IEEE Trans. Med. Imaging*, 2001, vol. 19, n° 3, pp 175–185.
- [Costa 95] COSTA W.L., HAYNOR D.R. et LEWELLEN T.K. Registration of segmented attenuation and emission data in PET. *Proc. of Nucl.*

- Science Symposium and Medical Imaging Conference*. 1995. pp 1407 – 1411.
- [Courtney 01] COURTNEY P. et THACKER N. A. Imaging and vision systems : Theory, assessment and applications, chapitre Performance Characterization in Computer Vision : The Role of Statistics in Testing and Design. NOVA Science Books, 2001.
- [Dahlbom 00] DAHLBOM M. et HUANG S.C. Handbook of medical imaging : Processing and analysis, chapitre : Physical and biological bases of spatial distortions in positron emission tomography images, pp 439–448. Academic Press, 2000.
- [Declerck 97] DECLERCK J., FELDMAR J., GORIS M.L. et BETTING F. Automatic registration and alignment on a template of cardiac stress and rest reoriented SPECT images. *IEEE Trans. Med. Imaging*, 1997, vol. 16, n ° 6, pp 727–737.
- [Dey 99] DEY D., SLOMKA P. J., HAHN L. J. et KLOIBER R. Automatic three-dimensional multimodality registration using radionuclide transmission CT attenuation maps : A phantom study. *J. Nucl. Med.*, 1999, vol. 40, n ° 3, pp 448–455.
- [Duncan 94] DUNCAN J., SHI P., AMINI A., CONSTABLE R., STAIB L., DIONE D., SINGER M. et SINUSAS A. Towards reliable, noninvasive measurement of myocardial function from 4D images. *SPIE Vol. 2168 Medical Imaging*. 1994. pp 149–161.
- [Eberl 96] EBERL S., KANNO I., FULTON R.R., RYAN A., HUTTON B.F. et FULHAM M.J. Automated interstudy image registration technique for SPECT and PET. *J. Nucl. Med.*, 1996, vol. 37, n ° 1, pp 137–145.
- [Faber 91] FABER T. L., MCCOLL R. W., OPPERMAN R. M., CORBETT J. R. et PESHOCK R. M. Spatial and temporal registration of cardiac SPECT and MR images : methods and evaluation. *Radiology*, 1991, vol. 179, n ° 3, pp 857–861.
- [Fitzpatrick 00] FITZPATRICK J.M., HILL D.L.G. et MAURER C.R. Handbook of medical imaging, volume 2, chapitre Image registration, pp 375–435. SPIE Press, 2000.
- [Förstner 96] FÖRSTNER WOLFGANG. 10 pros and cons against performance characterisation of vision algorithms. *Proc. of the workshop on "Performance Characteristics of Vision Algorithms"*, Cambridge. 1996.
- [Fryback 91] FRYBACK D. G. et THORNBURY J. R. The efficacy of diagnostic imaging. *Med. Decis. Making*, 1991, vol. 11, pp 88–94.
- [Gallippi 01] GALLIPPI C. M. et TRAHEY G. E. Automatic image registration for MR and ultrasound cardiac images. *Lecture Notes in Computer*

- Science 2082 : Information Processing in Medical Imaging, IPMI01.* Eds. INSANA M. F. et LEAHY R. M. 2001. pp 141–147.
- [Gee 00] GEE J. C. Performance evaluation of medical image processing algorithms. *Medical Imaging 2000 : Image Processing.* Eds. HANSON K. M. 2000.
- [Gilardi 96] GILARDI M.C., RIZZO G., SAVI A. et FAZIO F. Registration of multi-modal biomedical images of the heart. *Q. J. Nucl. Med.*, 1996, vol. 40, n ° 1, pp 142–150.
- [Gilardi 98] GILARDI M. C., RIZZO G., SAVI A., LANDONI C., BETTINARDI V., ROSSETTI C., STRIANO G. et FAZIO F. N. Correlation of SPECT and PET cardiac images by a surface matching registration technique. *Computerized Medical Imaging and graphics*, December 1998, vol. 22, pp 391–398.
- [Goerres 02] GOERRES G. W., KAMEL E., HEIDELBER T.-N. H., SCHWITTER M. R., BURGER C. et v SCHULTHESS G. K. PET-CT image co-registration in the thorax : influence of respiration. *Eur. J. Nucl. Med.*, 2002, vol. 29, n ° 3, pp 351–360.
- [Goodman 98] GOODMAN C. S. Introduction to health care technology assessment. *Nat. Library of Medecine/NICHSR*, 1998.
- [Grova 01] GROVA C., BIRABEN A., SCARABIN J.-M., JANNIN P., BUVAT I., BENALI H. et GIBAUD B. A methodology to validate MRI/SPECT registration methods using realistic simulated SPECT data. *Lecture Notes in Computer Science 2208 : Medical Image Computing and Computer-Assisted Intervention, MICCAI01.* Eds. NIESSEN W.J. et VIERGEVER M.A. 2001. pp 275–282.
- [Grova 02] GROVA C. *Simulations réalistes de données de tomographie d'émission monophotonique TEMP pour l'évaluation des méthodes de recalage TEMP/IRM utilisant des mesures statistiques de similarité : application dans le contexte de la fusion de données en épilepsie.* Thèse : Laboratoire IDM, Faculté de Médecine, Rennes, France, 2002.
- [Hajnal 01] HAJNAL J.V., HILL D.L.G. et HAWKES D.J. *Medical image registration.* CRC Press, 2001.
- [Hartiala 95] HARTIALA J. et KNUUTI J. Imaging of heart by MRI and PET. *Ann. Med.*, 1995, vol. 27, pp 35–45.
- [Hellier 03] HELLIER P., BARILLOT C., COROUGE I., GIBAUD B., GOUALHER G. LE, COLLINS L., EVANS A., MALANDAIN G. et AYACHE N. Retrospective evaluation of inter-subject brain registration. *IEEE Trans. Med. Imaging*, February 2003, vol. 22, n ° 9, pp 1120–1130.
- [Hemler 96] HEMLER P. F., SUMANAWEEERA T., V. D ELSEN P. A., NAPEL S. et ADLER J. R. Quantified registration error versus the accuracy

- of registered surfaces for a multimodality surface-based registration system. *SPIE Vol. 2710 Medical Imaging :Image processing*. 1996. pp 348–357.
- [Hill 94] HILL D. L. G., HAWKES D. J., GLEESON M. J., COX T. C. S., STRONG A. J., WONG W.-L., RUFF C. F., KITCHEN N., THOMAS D. G. T., CROSSMAN J. E., STUDHOLM C., GANDHE A. J., GREEN S. E. M. et ROBINSON G. P. Accurate frameless registration of MR and CT images of the head : Applications in surgery and radiotherapy planning. *Radiology*, 1994, vol. 191, pp 447–454.
- [Hill 00] HILL D.L.G. et HAWKES D.J. Handbook of medical imaging : Processing and analysis, chapitre : Across-modality registration using intensity-based cost functions, pp 537–553. Academic Press, 2000.
- [Hoffman 90] HOFFMAN E.J. et CUTLER P.D. 3-D phantom to simulate cerebral blood flow and metabolic images for PET. *IEEE Transactions on Nuclear Science*, 1990, vol. 37, n ° 2, pp 616–620.
- [Hoh 93] HOH C.K., DAHLBOM M., HARRIS G., CHOI Y., HAWKINS R.A., PHILPS M.E. et MADDAHI J. Automated iterative three-dimensional registration of positron emission tomography images. *J. Nucl. Med.*, 1993, vol. 34, n ° 11, pp 2009–2018.
- [Holden 00] HOLDEN M., HILL D. L. G., DENTON E. R. E., JAROSZ J. M., COX T. C. S., ROHLFING T., GOODEY J. et HAWKES D. J. Voxel similarity measures for 3-D serial MR brain image registration. *IEEE Trans. Med. Imaging*, February 2000, vol. 19, n ° 2, pp 94–102.
- [Hoppin 02] HOPPIN J. W., KUPINSKI M. A., KASTIS G. A., CLARKSON E. et BARRETT H. H. Objective comparison of quantitative imaging modalities without the use of a gold standard. *IEEE Trans. Med. Imaging*, May 2002, vol. 21, n ° 5, pp 441–449.
- [Horn 87] HORN B. K. P. Closed-form solution of absolute orientation using unit quaternions. *J. Opt. Soc. Am.*, 1987, vol. 4, n ° 4, pp 629–642.
- [Janier 98] JANIER M. *Etude de la perfusion et du métabolisme myocardiques par la tomographie d'émission de positons et l' Imagerie par Résonance Magnétique*. Thèse : Université Claude Bernard, Lyon 1, 1998.
- [Jannin 01] JANNIN P., GROVA G. et GIBAUD B. Fusion de données en imagerie médicale : revue méthodologique basée sur le contexte clinique. *ITMB-RBM Innovation et Technologie en Medecine et Biologie*, 2001, vol. 22, n ° 4, pp 196–215.
- [Jannin 02] JANNIN P., FITZPATRICK J. M., HAWKES D. J., PENNEC X., SHAHIDI R. et VANNIER M. W. Validation of medical image processing in image-guided therapy. *CARS*. Eds. INAMURA K., FARMAN A. G., DOI K. et REIBER J. H. C. 2002. pp 299–305.

- [Jenkinson 00] JENKINSON M. *Measuring Transformation error by RMS Deviation*. Technical Report TR99 MJ1, Oxford Centre for Functional Magnetic Resonance Imaging of the Brain, Departement of Clinical Neurology, Oxford University, Oxford, UK, 2000.
- [Jenkinson 01] JENKINSON M. et SMITH S. A global optimization method for robust affine registration of brain images. *Medical Image Analysis*, 2001, vol. 5, pp 143–156.
- [Kachigan 91] KACHIGAN S. K. *Multivariate statistical analysis*. Radius Press, 1991.
- [Kastler 97] KASTLER B. *Comprendre l'IRM*. Masson, 1997.
- [Kim 91] KIM R., AW T., BACHARACH S. et BONOW R. Correlation of cardiac MRI and PET images using lung cavities as landmarks. *Proc. IEEE Conf. Computers in Cardiology*. 1991. pp 49–52.
- [Klein 02a] KLEIN G. J. et HUESMAN R. H. Four-dimensional processing of deformable cardiac PET data. *Medical Image Analysis*, 2002, vol. 6, pp 29–46.
- [Klein 02b] KLEIN G. J., REUTTER B. W. et HUESMAN R. H. Four-dimensional affine registration models for respiratory-gated PET. *IEEE Trans. Medical Imag.*, 2002, vol. 48, n° 3, pp 756–760.
- [Kramer 89] KRAMER E.L., NOZ M.E., SANGER J.J., MEGIBOW A.J. et MAGUIRE G.Q. CT-SPECT fusion to correlate radiolabeled monoclonal antibody uptake with abdominal CT findings. *Radiology*, 1989, vol. 172, n° 10, pp 861–865.
- [Lartizien 02] LARTIZIEN C., COMPTAT C., KINAHAN P. E., FERREIRA N., BENDRIEM B. et TREBOSEN R. Optimization of the injected dose based on noise equivalent count rates for 2D and 3D whole-body PET. *Journal of Nuclear Medicine*, 2002, vol. 43, pp 1268–1278.
- [Lehmann 02] LEHMANN T. M. From plastic to gold : A unified classification scheme for reference standards in medical image processing. *Medical Imaging 2002 : Image Processing*. Eds. HANSON K. M. 2002.
- [Loew 97] LOEW M. H. Medical Imaging Registration Study Project. *Report of NASA Image Registration Workshop Nov.*, 1997.
- [Lötjönen 99] LÖTJÖNEN J., REISSMAN P.-J., MAGNIN I. E. et KATILA T. Model extraction from magnetic resonance volume data using the deformable pyramid. *Med. Image Anal.*, 1999, vol. 3, n° 4, pp 387–406.
- [Lötjönen 01] LÖTJÖNEN J. et MÄKELÄ T.J. Elastic matching using a deformation sphere. *Lecture Notes in Computer Science 2208 : Medical Image Computing and Computer-Assisted Intervention, MICCAI01*. Eds. NIESSEN W.J. et VIERGEVER M.A. 2001. pp 541–548.

- [Maes 97] MAES F., COLLIGNON A., VANDERMEULEN D., MARCHAL G. et SUETENS P. Multimodality image registration by maximization of mutual information. *IEEE Trans. Med. Imaging*, 1997, vol. 16, n° 2, pp 187–198.
- [Maes 99] MAES F., VANDERMEULEN D. et SUETENS P. Comparative evaluation of multiresolution optimisation strategies for multimodality image registration by maximization of mutual information. *Medical Image Analysis*, 1999, vol. 4, n° 3, pp 373–386.
- [Maintz 98] MAINTZ J. B. A. et VIERGEVER M. A. A survey of medical image registration. *Med. Image Anal.*, 1998, vol. 2, n° 1, pp 1–36.
- [Mäkelä 01] MÄKELÄ T. J., CLARYSSE P., LÖTJÖNEN J., SIPILÄ O., LAUERMA K., HÄNNINEN H., PYÖKKIMIES E.-P., NENONEN J., KNUUTI J., KATILA T. et MAGNIN I. E. A new method for the registration of cardiac PET and MR images using deformable model based segmentation of the main thorax structures. *Lecture Notes in Computer Science 2208 : Medical Image Computing and Computer-Assisted Intervention, MICCAI01*. Eds. NIESSEN W.J. et VIERGEVER M.A. 2001. pp 557–564.
- [Mäkelä 02] MÄKELÄ T. J., CLARYSSE P., O.SIPILÄ , PAUNA N., PHAM Q. C., KATILA T., LÖTJÖNEN J. et MAGNIN I. E. A review of cardiac registration methods. *IEEE Trans. Med. Imaging*, September 2002, vol. 21, n° 9, pp 1011–1021.
- [Mäkelä 03a] MÄKELÄ T. J. *Data registration and fusion for cardiac applications*. Thèse : Helsinki University of Technology, Finland, 2003.
- [Mäkelä 03b] MÄKELÄ T. J., POLLARI M., PAUNA N., REILHAC A., CLARYSSE P., MAGNIN I. et KATILA T. Evaluation and comparaison of surface and intensity based registration methods for thorax and cardiac MR and PET images. *Lecture Notes in Computer Science 2674 : Second International Workshop on Functional Imaging and Modeling of the Heart, FIMH, Lyon, France*. Eds. MAGNIN I.E., MONTAGNAT J., CLARYSSE P., NENONEN J. et KATILA T. 2003. pp 224–233.
- [Maurer 93] MAURER C. et FIZPATRICK J. M. A review of medical image registration. *Interactive image-guided neurosurgery*, 1993, pp 17–44.
- [Mazzadi 03] MAZZADI A. *Apport de l'imagerie fonctionnelle à l'étude des couplages entre la perfusion, le métabolisme glucidique et la fonction contractile dans l'ischémie myocardique*. Thèse : Université Claude Bernard, Lyon 1, 2003.
- [Meyer 96] MEYER C. R., BOES J. L., KIM B., BLAND P. H., ZASADNY K. R., KISON P. V., KORAL K., FREY K. A. et WAHL R. L. Demonstration of accuracy and clinical versatility of mutual information for automatic



- multimodality image fusion using affine and thin-plate spline warped geometric deformations. *Medical Image Analysis*, 1996, vol. 1, n° 2, pp 195–206.
- [Nekolla 00] NEKOLLA S., IBRAHIM T., BALBACH T. et KLEIN C. Understanding cardiac imaging techniques - from basic pathology to image fusion, chapitre : Coregistration and fusion of cardiac magnetic resonance and positron emission tomography studies, pp 144–154. IOS Press, 2000.
- [Nenonen 01] NENONEN J., PESOLA K., HÄNNINEN H., K. LAUERMA K., TAKALA P., MÄKELÄ T. J., MÄKIJÄRVI M., KNUUTI J., TOIVONEN L. et KATILA T. Current-density estimation of exercise-induced ischemia in patients with multivessel coronary artery disease. *Journal of Electrocardiography*, 2001, vol. 34, n° suppl., pp 37–42.
- [Nguyen 00] NGUYEN T. B. et ZIOU D. Contextual and Non-Contextual Performance Evaluation of Edge Detectors. *Pattern Recognition Letters*, 2000, vol. 21, pp 805–816.
- [Nicolau 03] NICOLAU S., PENNEC X., SOLER L. et AYACHE N. Evaluation of a new 3D/2D registration criterion for liver radio-frequencies guided by augmented reality. *Lecture Notes in Computer Science 2673 : International Symposium on Surgery Simulation and Soft Tissue Modeling IS4TM'03, juan-les-Pins, France*. Eds. AYACHE N. et DELINGETTE H. 2003. pp 270–283.
- [Pallotta 95] PALLOTTA S., GILARDI M. C., BETTINARDI V., RIZZO G., LANDONI C., STRIANO G., MASI R. et FAZIO F. Application of a surface matching image registration technique to the correlation of cardiac studies in positron emission tomography by transmission images. *Physics in Medicine and Biology*, 1995, vol. 40, pp 1695–1708.
- [Patton 00] PATTON J., DELBEKE D. et SANDLER M.P. Image fusion using an integrated, dual-head coincidence camera with X-ray tube-based attenuation maps. *J. Nucl. Med.*, 2000, vol. 41, n° 8, pp 1364–1368.
- [Pauna 03] PAUNA N., CROISILLE P., REILHAC A., MÄKELÄ T. J., JANIER M., COZAR O. et CLARYSSE P. A strategy to quantitatively evaluate MRI/PET cardiac rigid registration methods using a Monte Carlo simulator. *Lecture Notes in Computer Science 2674 : Second International Workshop on Functional Imaging and Modeling of the Heart, FIMH, Lyon, France*. Eds. MAGNIN I.E., MONTAGNAT J., CLARYSSE P., NENONEN J. et KATILA T. 2003. pp 194–204.
- [Pelizzari 89] PELIZZARI C. A., GHEN G. T. Y., SPELBRING D. R., WEICHSELBAUM R. R. et CHEN C. T. Accurate three-dimensional registration of CT, PET and/or MR images of the brain. *J Comput Assist Tomogr*, 1989, vol. 13, n° 1, pp 20–26.

- [Pennec 98] PENNEC X., GUTTMANN C.R.G. et THIRION J.-P. Feature-based registration of medical images : Estimation and validation of the pose accuracy. *Lecture Notes in Computer Science 1496 : Medical Image Computing and Computer-Assisted Intervention, MICCAI98, Cambridge, USA*. Eds. WELLS W. M., COLCHESTER A. et DELP S. 1998. pp 1107–1114.
- [Phelps 95] PHELPS C. E. et HUSTON A. Estimating diagnostic test accuracy using a "fuzzy" gold standard. *Med. Decis. Making*, 1995, vol. 15, pp 44–57.
- [Pinet 02] PINET F., BABUTY D., CARRIER L., DUPERRAY A., GRYNBERG A., LOIRAND G. et SAMUEL J.L. Biologie et pathologie du coeur et des vaisseaux. Médecine-Sciences, Flammarion, 2002.
- [Pluim 03] PLUIM J.P., MAINTZ J.B. et VIERGEVER M.A. Mutual information based registration of medical images : A Survey. *IEEE Trans. Med. Imaging*, 2003, vol. 22, n ° 8, pp 986–1005.
- [Pollari 04] POLLARI M., LÖTJÖNEN J., MÄKELÄ T., PAUNA N., REILHAC A. et CLARYSSE P. Evaluation of cardiac PET-MRI registration method using a numerical breathing phantom. *IEEE International Symposium on Biomedical Imaging*. 2004. pp 1447–1450.
- [Press 92] PRESS W. H., TEUKOLSKY S. A., VETTERLING W. T. et FLANNERY B. P. Numerical recipes in C : The art of scientific computing 2nd edition. Cambridge Univ. Press, Cambridge, 1992.
- [Pretorius 97] PRETORIUS P. H ., XIA W., KINAG M. A., TSUI B. M. W., PAN T. S. et VILLEGAS B. J. Evaluation of right and left ventricular volume and ejection fraction using a mathematical cardiac torso phantom for gated blood pool SPECT. *J. Nucl. Med.*, 1997, vol. 38, n ° 10, pp 1528–1535.
- [Pretorius 99] PRETORIUS P. H . et KING M.A. A mathematical model of motion of the heart for use in generating source and attenuation maps for simulating emission imaging. *Medical Physics*, 1999, vol. 26, pp 2323–2332.
- [Reilhac 99] REILHAC A., GREGOIRE M. C., COSTES N., LAVENNE F., PIERRE C., DIOU A. et PUJOL J.-F. A PET Monte Carlo simulator from numerical phantom : Validation against the EXACT ECAT HR+ scanner. *Proc. IEEE Nuclear Sciences Symposium, Seattle, Washington, USA*. 1999. pp 1527–1531.
- [Reilhac 02] REILHAC A., LARTIZIEN C., COSTES N., SANS S., COMTAT C. et EVANS A. Accounting for single rates related phenomena in PET Monte Carlo based simulations. *In Proc. IEEE Nuclear Sciences Symposium and Medical Imaging Conference, Norfolk*, 2002, vol. 3, pp 1617–1621.

- [Reilhac 04] REILHAC A., LARTIZIEN C., COSTES N., SANS S., COMTAT C., GUNN R. N. et EVANS A. C. PET-SORTEO : A Monte Carlo-based simulator with high count rate capabilities. *IEEE Transaction on Nuclear Sciences*, 2004, vol. 51, n° 1, pp 46–52.
- [Roche 98] ROCHE A., PENNEC X., MALANDAIN G. et AYACHE N. The correlation ratio as a new similarity measure for multimodal image registration. *Lecture Notes in Computer Science 1496 : Medical Image Computing and Computer-Assisted Intervention, MICCAI98, Cambridge, USA*. Eds. WELLS W. M., COLCHESTER A. et DELP S. 1998. pp 1115–1124.
- [Roche 01] ROCHE A., PENNEC X., MALANDAIN G. et AYACHE N. Rigid registration of 3-D ultrasound with MR images : A new approach combining intensity and gradient information. *IEEE Trans. Med. Imaging*, 2001, vol. 20, n° 10, pp 1038–1049.
- [Sarrut 00] SARRUT D. *Recalage multimodal et plate-forme d'imagerie médicale à accès distant*. Thèse : Université Lumière, Lyon 2, 2000.
- [Sechet 02] SECHET S., REILHAC A., GUNN R., EVANS A. C. et DAGHER A. Frame misalignment-induced errors in PET studies : an investigation on strategies for correction. *In Proc. IEEE Nuclear Sciences Symposium and Medical Imaging Conference, Norfolk*, 2002, vol. 2, pp 1330–1334.
- [Segars 01] SEGARS P. *Development and application of a new dynamic NURBS-BASED cardiac-torso (NCAT) phantom*. Thèse : University of Carolina, Chapel Hill, USA, 2001.
- [Shahidi 01] SHAHIDI R. White paper : Challenges and opportunities in computer-assisted interventions january 2001. *Comput. Aided Surg.*, 2001, vol. 6, n° 3, pp 176–181.
- [Shtern 99] SHTERN F. *Report of the joint working group on image guided diagnosis and treatment, Washington, DC*, 1999.
- [Siltanen 88] SILTANEN P. Comprehensive electrocardiology, chapitre Magnetocardiography. Oxford : Pergamon Press, 1988.
- [Sinha 95] SINHA S., SINHA U., CZERNIN J., PORENTA G. et SCHELBERT H.R. Noninvasive assessment of myocardial perfusion and metabolism : feasibility of registering gated MR and PET images. *Am. J. Roentgenol.*, 1995, vol. 36, pp 301–307.
- [Slates 99] SLATES R. B., FARAHANI K., SHAO Y., MARSDEN P. K., TAYLOR J., SUMMERS P. E., WILLIAMS S., BEECH J. et CHERRY S. R. A study of artefacts in simultaneous PET and MR imaging using a prototype MR compatible PET scanner. *Phys. Med. Biol.*, 1999, vol. 44, pp 2015–2027.

- [Slomka 95] SLOMKA P.J., GILBERT A.H., STEPHENSON J. et CRADDUC T. Automated alignment and sizing of myocardial stress and rest scans to three-dimensional normal templates using an image registration algorithm. *J. Nucl. Med.*, 1995, vol. 36, pp 1115–1122.
- [Slomka 00] SLOMKA P. J., DEY D., PRZETAK C. et BAUM R. Automated 3-D spatial integration of 18-F FDG wholebody PET with CT. *J. Nucl. Med.*, 2000, vol. 41, n° 6, pp 1010–1017.
- [Stone 98] STONE C.D., MCCORMICK J.W., GILLAND D.R., GREER K.L, COLEMAN R.E. et JASZCZAK R.J. Effect of registration errors between transmission and emission scans on a SPECT system using sequential scanning. *J. Nucl. Med.*, 1998, vol. 39, n° 2, pp 365–373.
- [Studholme 96] STUDHOLME C., G.HILL D.L. et HAWKES D.J. Automated 3D registration of MR and CT images of the head. *Medical Image Analysis*, 1996, vol. 1, n° 2, pp 163–175.
- [Sturm 03] STURM B., POWELL K. A., STILLMAN A. E. et WHITE R. D. Registration of 3D CT angiography and cardiac MR images in coronary artery disease patients. *The International Journal of Cardiovascular Imaging*, 2003, vol. 19, pp 281–293.
- [Tai 97] TAI Y.-C., LIN K.P., HOH C.K., HUANG S.C. et HOFFMAN E.J. Utilization of 3-D elastic transformation in the registration of chest X-ray CT and whole body PET. *IEEE Trans. Nucl. Med.*, 1997, vol. 44, n° 4, pp 1606–1612.
- [Tsui 93] TSUI B.M.W. et TERRY J.A. Evaluation of cardiac cone-beam SPECT using observer performance experiments and ROC analysis. *Investigative Radiology*, July 1993, vol. 28, n° 12, pp 1101–1112.
- [Turkington 97] TURKINGTON T.G., DEGRADO T.R., HANSON M.W. et R.E.COLEMAN . Alignment of dynamic cardiac PET images for correction of motion. *IEEE Trans. Nucl. Sci.*, 1997, vol. 44, n° 2, pp 235–242.
- [van den Elsen 93] V. D ELSEN P.A., POL E.-J.D. et VIERGEVER M.A. Medical image matching - A review with classification. *IEEE Engineering in Medicine and Biology*, 1993, vol. 12, n° 2, pp 16–39.
- [Viant 01] VIAN T W. J. The development of an evaluation framework for the quantitative assessment of computed-assisted surgery and augmented reality accuracy performance. *Stud. Health. Technol. Inform.*, 2001, vol. 81, pp 534–540.
- [Viola 95] VIOLA P. A. *Alignment by maximization of mutual information*. Thèse : Massachusetts Institute of Technology, USA, 1995.

- [Wang 92] WANG H. Solid geometry-based object model for monte carlo simulated emission and transmission tomographic imaging systems. *IEEE Trans. Med. Imaging*, 1992, vol. 11, n ° 3, pp 361–372.
- [Wells 96] WELLS W.M., VIOLA P., ATSUMI H., NAKAJIMA S. et KIKINIS R. Multi-modal volume registration by maximization of mutual information. *Med. Image Anal.*, 1996, vol. 1, n ° 1, pp 35–51.
- [West 97] WEST J., FITZPATRICK J. M., WANG M. Y., DAWANT B. M., MAURER C. R., KESSLER R. M., MACIUNAS R. J., BARILLOT C., LEMOINE D., COLLINGNON A., MAES F., SUETENS P., VANDERMEULEN D., v. D ELSEN P. A., NAPEL S., N B. HARKNESS T. S. SUMANAWEEERA, HEMLER P. F., HILL D. L. G., HAWKES D. J., STUDHOLME C., MAINTZ J. B. A., VIERGEVER M. A., MALANDAIN G., PENNEC X., NOZ M. E., MAGUIRE G. Q., POLLACK M., PELIZZARI C. A., ROBB R. A., HANSON D. et WOODS R. P. Comparison and evaluation of retrospective intermodality image registration techniques. *J. Comput. Assist. Tomogr.*, 1997, vol. 21, pp 554–566.
- [Woods 92] WOODS R.P., CHERRY S.R. et MAZZIOTTA J.C. Rapid automated algorithm for aligning and reslicing PET images. *J. Comput. Assist. Tomogr.*, 1992, vol. 17, n ° 4, pp 620–633.
- [Woods 93] WOODS R.P., MAZZIOTTA J.C. et CHERRY S.R. MRI-PET registration with automated algorithm. *J. Comput. Assist. Tomogr.*, 1993, vol. 17, n ° 4, pp 536–546.
- [Woods 98] WOODS R.P., HOLMES S.T., MAZZIOTTA J.C. et CHERRY S.R. Automated image registration I : General methods and intrasubject, intramodality validation. *J. Comput. Assist. Tomogr.*, 1998, vol. 22, n ° 4, pp 141–154.
- [Woods 00a] WOODS R.P. Handbook of medical imaging : Processing and analysis, chapitre Validation of registration accuracy, pp 491–497. Academic Press, 2000.
- [Woods 00b] WOODS R.P. Handbook of medical imaging : Processing and analysis, chapitre Spatial transformation models, pp 465–490. Academic Press, 2000.
- [Yoo 00] YOO T. S., ACKERMAN M. J. et VANNIER M. Toward a common validation methodology for registration and segmentation algorithms. *Proc. of the Second Int. Conf. on Medical Image Computing and Computer-Assisted Intervention - MICCAI 2000*. Lecture Notes in Computer Science 1935, 2000. Lecture Notes in Computer Science 1935. pp 422–431.
- [Yu 94] YU J.N., FAHEY F.H., GAGE H.D., EADES C.G. et KEYES J.W. Evaluation of emission-transmission registration in thoracic PET. *J. Nucl. Med.*, 1994, vol. 35, n ° 11, pp 1777–1781.

- [Yu 95] YU J. N., FAHEY F. H., GAGE H. D., EADES C. G., HARKNESS B. A. et PELIZZARI C. A. Intermodality, retrospective image registration in the thorax. *J. Nucl. Med.*, December 1995, vol. 36, n ° 12, pp 2333–2338.
- [Zerhouni 88] ZERHOUNI E.A., PARISH D.M., ROGERS W.J., YANG A. et SHAPIRO E. P. Human heart : tagging with MR imaging-a method for noninvasive assessment of myocardial motion. *Radiology*, 1988, vol. 169, pp 59–63.
- [Zitova 03] ZITOVA B. et FLUSSER J. Image registration methods : a survey. *Image and Vision Computing*, May 2003, vol. 21, pp 977–1000.
- [Zubal 94] ZUBAL I.G. et HARRELL C. R. Computerized three-dimensional segmented human anatomy. *Medical Physics*, 1994, vol. 21, pp 299–302.

