

SEGMENTATION FÉMORALE PAR MODÈLE DÉFORMABLE ET PROGRAMMATION DYNAMIQUE

THÈSE N° 3833 (2007)

PRÉSENTÉE LE 26 JUILLET 2007

À LA FACULTÉ DES SCIENCES ET TECHNIQUES E L'INGÉNIEUR

Laboratoire de systèmes robotiques 2

SECTION DE MICROTECHNIQUES

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

PAR

Gaëtan MARTI

ingénieur informaticien diplômé EPF
de nationalité suisse et originaire de Lausanne (VD)

acceptée sur proposition du jury:

Prof. J. Jacot, président du jury
Dr C. Baur, directeur de thèse
Dr M. A. Gonzalez Ballester, rapporteur
Prof. J.-Ph. Thiran, rapporteur
Dr J. Troccaz, rapporteur



ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Lausanne, EPFL

2007

TABLE DES MATIÈRES

| | |
|--|-----------|
| LISTE DES ABRÉVIATIONS, SIGLES ET ACRONYMES | 1 |
| RÉSUMÉ | 2 |
| 0.1 Mots clefs | 2 |
| ABSTRACT | 3 |
| 0.2 Keywords | 3 |
| CHAPITRE I | |
| AVANT-PROPOS | 4 |
| 1.1 Note sur l'état de l'art | 5 |
| 1.2 Remerciements | 5 |
| CHAPITRE II | |
| INTRODUCTION | 6 |
| CHAPITRE III | |
| RECALAGE RIGIDE | 11 |
| 3.1 Introduction | 11 |
| 3.2 Recalage point à point | 12 |
| 3.3 Algorithme du plus proche point itéré (<i>ICP</i>) | 15 |
| 3.3.1 Notation | 16 |
| 3.3.2 Métrique entre les jeux de données | 16 |
| 3.3.3 Méthodes d'appariement | 17 |
| 3.3.4 Types de mise en correspondance | 21 |
| 3.3.5 Minimisation de l'erreur métrique | 26 |
| 3.3.6 Limitations et améliorations d'ICP | 28 |
| 3.4 Conclusion | 32 |
| CHAPITRE IV | |
| RECALAGE NON RIGIDE | 35 |
| 4.1 Introduction | 35 |
| 4.2 Déformations globales | 36 |
| 4.2.1 Déformations linéaires | 36 |
| 4.2.2 Déformations polynômiales | 37 |
| 4.3 Déformations locales | 38 |
| 4.3.1 Déformation localement affine | 38 |
| 4.3.2 Interpolation par des fonctions à base radiale | 38 |
| 4.3.3 Déformation de forme libre | 42 |

| | | |
|-----|--|----|
| 4.4 | Approche hiérarchique | 45 |
| 4.5 | Approche progressive | 46 |
| 4.6 | Contribution | 48 |
| 4.7 | Variations autour du recalage progressif | 55 |
| 4.8 | Conclusion | 55 |

CHAPITRE V LOCALISATION AUTOMATIQUE DE PARTIES ANATOMIQUES DU FÉMUR 57

| | | |
|-------|--|----|
| 5.1 | Introduction | 57 |
| 5.2 | Etat de l'art | 58 |
| 5.3 | Localisation par rayons de recherche | 58 |
| 5.4 | Classification des données scanner | 59 |
| 5.4.1 | Recherche de la diaphyse | 62 |
| 5.4.2 | Analyse des profils diaphysaires | 64 |
| 5.4.3 | Recherche de l'axe diaphysaire | 66 |
| 5.4.4 | Localisation du grand trochanter | 67 |
| 5.4.5 | Recherche de la tête fémorale | 69 |
| 5.4.6 | Recherche de l'axe du col du fémur | 72 |
| 5.5 | Résultats | 72 |
| 5.6 | Conclusion | 74 |

CHAPITRE VI RECALAGE NON RIGIDE PAR MODÈLE DÉFORMABLE 76

| | | |
|-------|---|----|
| 6.1 | Introduction | 76 |
| 6.2 | Modèle articulé | 77 |
| 6.3 | Du squelette à la peau | 79 |
| 6.4 | Cylindres généralisés | 80 |
| 6.4.1 | Système de coordonnées cylindriques | 81 |
| 6.4.2 | Modèle du fémur | 83 |
| 6.4.3 | Opérations de déformation | 86 |
| 6.5 | Validation du modèle | 86 |
| 6.6 | Recalage non-itératif | 88 |
| 6.7 | Conclusion | 90 |

CHAPITRE VII SEGMENTATION FINE PAR PROGRAMMATION DYNAMIQUE 92

| | | |
|-----|--------------------------------------|----|
| 7.1 | Introduction | 92 |
| 7.2 | Segmentation par seuillage | 93 |
| 7.3 | Segmentation par snakes | 93 |

| | | |
|---|---|------------|
| 7.4 | Rayons actifs | 95 |
| 7.5 | Représentation de l'énergie | 96 |
| 7.6 | Programmation dynamique | 97 |
| 7.7 | Contribution - Optimisation de l'énergie par programmation dynamique | 99 |
| 7.7.1 | Heuristique pour la sélection du point initial | 100 |
| 7.7.2 | Paramètres | 101 |
| 7.8 | Validation | 104 |
| 7.9 | Résultats | 104 |
| 7.10 | Modèle déformable et programmation dynamique | 106 |
| 7.11 | Conclusion | 109 |
| CHAPITRE VIII | | |
| CONCLUSION | | 112 |
| 8.1 | Synthèse | 112 |
| 8.2 | Contributions | 113 |
| 8.3 | Travaux futurs | 114 |
| 8.4 | Publications et posters | 115 |
| APPENDICE A | | |
| APPARIEMENT FLOU | | 117 |
| APPENDICE B | | |
| REPRÉSENTATION DES ROTATIONS | | 120 |
| 8.5 | Angles d'Euler | 120 |
| 8.5.1 | Gimbal Lock | 121 |
| 8.6 | Quaternions | 121 |
| 8.7 | Cartes exponentielles et formule de Rodrigues | 123 |
| APPENDICE C | | |
| LE FÉMUR | | 125 |
| 8.8 | Anatomie du fémur | 125 |
| 8.8.1 | Partie supérieure | 125 |
| 8.8.2 | La tête (<i>caput femoris</i>) | 128 |
| 8.8.3 | Le col(<i>collum femoris</i>) | 128 |
| 8.8.4 | Les trochanters | 129 |
| 8.8.5 | Tubercule | 130 |
| 8.8.6 | Ligne intertrochantérique | 130 |
| 8.9 | L'architecture interne de la partie proximale du fémur | 130 |
| 8.10 | Pathologies du fémur | 130 |
| 8.10.1 | Ostéoporose | 131 |

| | | |
|--------|---|-----|
| 8.10.2 | Arthroses de hanche ou coxarthroses | 131 |
| 8.10.3 | Fractures | 132 |
| 8.10.4 | Autres pathologies | 133 |

APPENDICE D

LOCALISATION DE CENTRE D'ELLIPSES PAR LANCER

DE RAYONS 134

| | | |
|--------|-------------------------|-----|
| 8.10.5 | Proposition | 134 |
| 8.10.6 | Démonstration | 134 |
| 8.10.7 | Proposition | 134 |
| 8.10.8 | Démonstration | 135 |

GLOSSAIRE 136

RÉFÉRENCES 138

LISTE DES ABRÉVIATIONS, SIGLES ET ACRONYMES

2D Bidimensionnel

3D Tridimensionnel

CFAO Conception et Fabrication Assistée par Ordinateur

CHU Centre Hospitalier Universitaire

CT *Computer Tomography*

CTA *Computer Tomography Angiography*

DDL Degrés De Liberté

FEM *Finite Element Method*

GC *Generalized Cylinders*

ICP *Iterative Closest Point*

LM-ICP ICP par résolution Levenberg-Markardt

PD Programmation Dynamique

PTH Prothèse Totale de Hanche

RANSAC *RANdom SAmple Consensus*

RBF *Radial Basis Function*

SSD *Sum Square Difference*

TPS *Thin Plate Splines*

VRAI *Virtual Reality and Active Interfaces*

RÉSUMÉ

Cette thèse propose de nouveaux paradigmes dans le domaine de la segmentation par modèles déformables.

Contrairement à l'approche classique qui consiste à faire évoluer itérativement un modèle déformable complexe de l'organe à segmenter dans l'espace des coupes scanners, notre approche simplifie et décompose le problème en trois phases : dans un premier temps, les caractéristiques anatomiques de l'organe à segmenter sont localisées automatiquement au moyen d'une technique de lancers de rayons. Ensuite, un modèle déformable simplifié est transformé pour correspondre à ces données. Il permet une localisation approximative du contour de l'organe. Ce positionnement brut peut être affiné par un recalage conventionnel. Enfin, le contour précis de l'organe est délimité par une méthode de segmentation par contours actifs (*snakes*). Nous utilisons une résolution basée sur la programmation dynamique afin d'éviter les problèmes des minima locaux inhérents à ce type de segmentation.

Les avantages principaux de cette nouvelle approche sont un très bon compromis entre la vitesse de segmentation, la qualité du résultat et la possibilité de revenir à tout moment à une segmentation interactive de l'organe.

La problématique de la segmentation du fémur (au niveau de l'articulation de la hanche) sur des séries CT chez des patients présentant de fortes pathologies est utilisée pour valider notre approche. Au vu des bons résultats obtenus, ce nouveau paradigme de segmentation pourrait, concrètement, être utilisé dans des logiciels de planification prothétiques, comme par exemple pour la pose d'implants fémoraux par resurfaçage.

0.1 Mots clefs

Segmentation fémorale, recalage non rigide, modèles déformables, programmation dynamique.

ABSTRACT

This Ph.D thesis presents new paradigms in the field of segmentation by mean of deformable models.

Classic model-based segmentation approach consists of iteratively deform complex models in the space of the scanner slices until the organ to segment is fitted. The proposed works tries to improve the process by decomposing the problem in three distinct phases :

1. First, anatomical landmarks of the organ to segment are automatically detected using a fast ray-casting technique ;
2. A basic deformable model is next initialized to fit these landmarks. It allows approximately fitting the organ contour. This raw approximation can optionally be refined by a conventional registration ;
3. Finally, an active contour method (snakes) allows to accurately delimiting the contour of the organ. Although contour delimitation by classic snake resolution usually converges to a local minimum, this problem is avoided by using dynamic programming.

A main advantage of our approach is to propose a good compromise between segmentation speed and quality of the result. It also allows coming back to an interactive segmentation of the organ anytime.

CT femoral segmentation of degenerative hips has been used to validate our approach. Good results were achieved on most of the patients. Only highly degenerated femurs with osteoporosis or arthritis were not fully automatically segmented. Thanks to the goods results achieved, this new segmentation paradigm could practically be used in prosthesis planning software like femur resurfacing.

0.2 Keywords

Femoral segmentation, non-rigid registration, deformable models, dynamic programming.

CHAPITRE I

AVANT-PROPOS

Dans son cahier des charges initial, ce travail consistait à identifier et implémenter des techniques de recalage temps-réel pour rapporter des informations de planification durant les chirurgies hépatiques, comme par exemple des plans de résection.

La première partie de ma thèse s'est déroulée en milieu clinique au CHU de Sherbrooke afin de comprendre et d'affiner ma compréhension de l'imagerie ultrasonique intra opératoire. J'ai ensuite développé, sur ce type d'images, des algorithmes originaux de segmentation du réseau hépatique avant de m'attaquer au coeur du problème, c'est-à-dire, le recalage non rigide d'un modèle préopératoire tridimensionnel sur un patient au cours d'une procédure opératoire. Pour démontrer l'approche de recalage que je comptais entreprendre, j'ai transposé cette problématique dans un domaine connexe : la segmentation basée sur des modèles déformables. C'est-à-dire qu'au lieu de déformer l'organe d'un même patient qui varie au cours du temps, il s'agit de déformer l'organe d'un patient ou d'une combinaison de patients (modèle statistique) pour s'adapter à l'anatomie du patient à segmenter.

Le choix des données utilisées pour valider cette nouvelle approche s'est porté sur la segmentation du fémur sur des images CT, problème qui, à l'heure actuelle, n'est pas complètement résolu malgré son apparente facilité. Il s'agit également d'un domaine de recherche dans lequel le groupe VRAI a développé de nombreux outils et possède une importante bibliothèque de données.

Le résultat de ce travail de thèse est de réaliser un système de recalage aussi rapide que possible ainsi que l'emploi d'un modèle déformable pouvant être transposable à la problématique hépatique.

1.1 Note sur l'état de l'art

L'état de l'art décrit dans cette thèse correspond à la situation au moment du début de sa réalisation, c'est-à-dire courant 2003. Dans l'intervalle, de nombreuses nouvelles techniques de segmentation et de recalage ont vu le jour et n'ont pas été intégrés dans ce mémoire.

1.2 Remerciements

Je tiens à remercier mon directeur de thèse, le Dr. Charles Baur, pour la confiance et l'indépendance qu'il m'a laissé durant ce travail.

Un grand merci à ma famille qui m'a soutenu tout au long de cette interminable rédaction.

Merci aussi à tout ceux, qu'ils soient professeurs, collègues, étudiants, famille ou amis qui ont laissé directement ou indirectement leur griffe dans ce travail.

CHAPITRE II

INTRODUCTION

"Les systèmes de Chirurgie Assistée par Ordinateur (CAO) ont été introduits à la fin des années 80 pour la chirurgie stéréotaxique en neurochirurgie dans le but d'assister le chirurgien afin qu'il puisse placer une sonde à l'intérieur de l'encéphale sans abord direct de la zone opérée. La grande idée des technologies de Chirurgie Assistée par Ordinateur (CAO) consiste à augmenter la précision de l'acte, à réduire la morbidité et le caractère invasif des gestes chirurgicaux, à réduire les doses de rayons X et à améliorer les protocoles chirurgicaux en permettant des études post-opératoires performantes, basées sur des enregistrements d'actes opératoires. Le principe de base des systèmes de CAO consiste à localiser en trois dimensions la position d'instruments chirurgicaux dans le champ opératoire et à faire apparaître ces mêmes instruments sur des images préopératoires (CT ou IRM) ou per-opératoires (images radiographiques)." [71]

Dans le domaine orthopédique, la CAO est principalement utilisée pour l'aide aux opérations du rachis et durant les arthroplasties du genou, de l'épaule ou de la hanche. Dans ce dernier domaine, les résultats des placements de prothèses modernes peuvent être considérés comme excellents si l'on envisage la fixation entre l'os et la prothèse. Cependant, les luxations de prothèses et l'usure prématurée de certains composants prothétiques sont les deux principales causes de complications rencontrées. Une assistance de navigation durant l'arthroplastie permet de mieux définir l'emplacement optimal des éléments prothétiques et, ainsi de garantir une bonne espérance de vie de l'implant.

La CAO ou navigation chirurgicale permet également d'autres approches de l'arthroplastie de hanche comme l'abord minimal invasif ou l'emploi d'une cupule fémorale (méthode de resurfacage) qui permet de conserver la tête fémorale. Il s'agit d'une alternative à la prothèse totale chez les patients jeunes et actifs. Cette méthode permet, en cas d'usure ou de descellement

après 10-15 ans, d'implanter une tige classique dans un fémur pratiquement intact. On peut donc s'attendre à une reprise moins conséquente et qui donne un résultat similaire à une prothèse classique.

L'arthroplastie de la hanche naviguée nécessite la localisation précise de différentes variables morphologiques du patient comme l'axe intramédullaire et l'axe du col du fémur. Ces paramètres peuvent être définis dans les images préopératoires du patient au moyen de méthodes de segmentation d'images.

La segmentation représente l'une des premières étapes du traitement de données radiologiques pour des applications cliniques aussi variées que le diagnostic radiologique, l'analyse cinématique, le monitoring, la radiothérapie, le planning chirurgical ou la navigation assistée par ordinateur. Elle permet la classification et la visualisation de tissus possédant des propriétés communes. Bien que la segmentation osseuse à partir d'images tomographiques (CT) a démontré son utilité dans de nombreuses applications médicales, son automatisation est une tâche difficile pour des raisons aussi variées que : la non uniformité des tissus osseux, la présence de vaisseaux sanguins, la diffusion des contours osseux, la faible distance entre les os, les artefacts liés aux imageurs ainsi que des effets de discrétisation (anisotropie).

La hanche est une articulation à rotule, composée du fémur emboîté dans le bassin. La partie proximale d'un fémur sain a une forme de boule : la tête fémorale, qui s'insère dans une cavité de la hanche : l'acétabulum. La tête fémorale est recouverte d'une couche de cartilage mou. Un examen CT de l'articulation de la hanche révèle habituellement un espace entre ces deux os dû à la faible densité du cartilage aux rayons X. Cet espace de quelques millimètres aide à visualiser plus facilement le contour des structures osseuses. Malheureusement des pathologies comme la coxarthrose, les déformations congénitales, ou les limitations techniques des imageurs comme les flous ou la discrétisation des données ne permettent pas toujours de délimiter correctement le contour osseux.

La segmentation des os dans les images CT est traditionnellement réalisée par un seuillage, suivi d'une technique de connectivité (ou classification) des régions résultantes. Théoriquement, ce processus est relativement efficace car l'intensité de l'os sain est plus élevée que les tissus environnants. Une des limites de cette méthode est, par exemple, atteinte lorsque le fémur et l'acétabulum sont physiquement adjacents (voir la figure 2.1). Pour parler à ce genre de problème, de très nombreuses techniques de segmentation sont disponibles dans la littérature. Malheureusement peu d'entre elles pré-

sentent des résultats avec un taux de succès élevé dont le temps de calcul soit raisonnable.

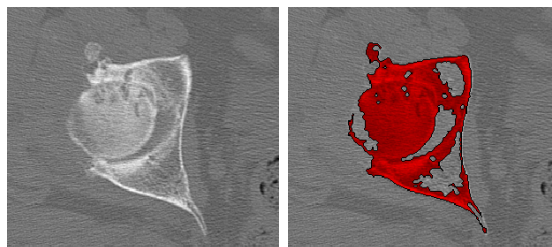


FIG. 2.1: Exemple de coupe CT où sont visibles la tête fémorale et l'acétabulum. A droite, image non segmentée, à gauche image segmentée par seuillage (en rouge), sur laquelle le problème est visible le problème de connectivité car les deux os sont adjasents.

Ce travail propose de définir des nouveaux paradigmes pour la segmentation de fémurs dans des coupes CT basé sur des modèles déformables (*model-based segmentation*), aussi appelés segmentation par approche frontière. D'après la définition de l'encyclopédie Wikipedia (http://fr.wikipedia.org/wiki/Segmentation_d'image), cette approche cherche à exploiter le fait qu'il existe, dans les données, (c'est-à-dire une image, une série temporelle ou un volume,) une transition détectable entre deux régions connexes. On fait intervenir pour cela des modèles déformables. Cette classe d'algorithmes consiste à poser sur les données une forme mathématique contrôlée, par exemple un polygone ou une spline dans un espace bidimensionnel ou une fonction à base radiale dans un espace tridimensionnel. On fait ensuite évoluer cette forme pour délimiter la zone à segmenter. Le processus de segmentation par modèles déformables se décompose classiquement en quatre phases :

1. Le choix de la classe de modèle déformable,
2. L'initialisation approximative du modèle sur les données à segmenter,
3. L'évolution itérative du modèle en fonction d'un ensemble de règles : souvent on minimise une énergie qui dépend à la fois des données (pour se rapprocher de l'objet à segmenter) et de la forme (pour régulariser celle-ci),
4. L'arrêt de l'évolution : le critère couramment utilisé est la stabilisation du modèle.

L'intérêt principal des méthodes de segmentation selon l'approche frontière est de minimiser le nombre d'opérations nécessaires en cas d'itération

du processus sur des séries de données peu différentes les unes des autres. En effet, une fois que les contours des régions ont été trouvés dans les données de référence, l'application du modèle déformable aux données à segmenter est plus efficace que de tout recalculer, si la différence entre ces jeux de données est peu importante. Bien que cette classe d'algorithmes présente, à notre avis, la méthode de segmentation la plus fiable pour l'imagerie médicale (car l'anatomie entre les patients varie faiblement), le processus classique souffre des limitations suivantes :

- L'initialisation du modèle déformable peut nécessiter une phase préalable de segmentation ou une localisation des points saillants (*salient points*),
- En règle générale, un modèle déformable complexe permet une grande variabilité de formes mais nécessite de considérables ressources de calcul. Inversement, un modèle trop simple ne représente que grossièrement de la forme à segmenter. Pratiquement le calcul d'un modèle simple est facilement mille fois plus rapide que le calcul d'un modèle complexe,
- Dans sa conception de base, un modèle déformable est un modèle itératif qui, par des techniques d'optimisation, converge vers l'anatomie du patient. Comme la fonction d'optimisation présente, dans la plupart des cas, de nombreux minima locaux, la qualité de la segmentation dépend fortement de son initialisation,
- L'aspect itératif, lié à la complexité du modèle nécessite des ressources de calcul considérables, qui ne sont pas adaptées à des approches temps-réel.

Ce travail 2.2 propose de lever tout ou partie des contraintes précitées en apportant des solutions originales à la segmentation basée sur des modèles déformables. Tout d'abord, la phase d'initialisation, que nous proposons, permet une localisation extrêmement rapide des caractéristiques anatomiques du fémur à segmenter. Ces caractéristiques sont obtenues uniquement par une exploration partielle du volume au moyen d'une technique de recherche par lancer de rayons. Un modèle déformable simple, tiré du domaine de l'animation de synthèse et basé sur une représentation par des cylindres généralisés, est ensuite recalé sur l'imagerie CT sans nécessiter d'itérations, c'est-à-dire, sans mécanisme d'optimisation. A l'issue de cette étape, le modèle n'identifie que grossièrement le contour du fémur. Il est finalement affiné par une technique de segmentation d'images originale basée sur les contours actifs. Contrairement aux approches classiques, cette dernière étape est non itéra-

tive et se base sur une représentation résolvable par programmation dynamique. En résumé, les points forts de ce travail sont de proposer un nouveau processus de segmentation par modèle rapide, non itératif mais néanmoins optimal.

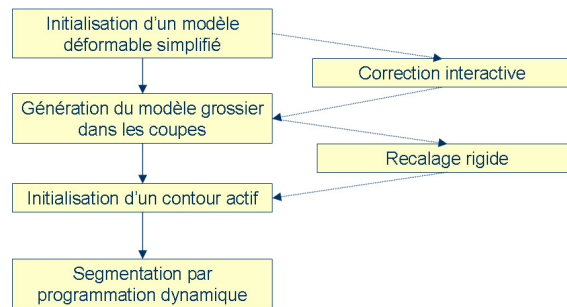


FIG. 2.2: Processus de segmentation proposé. La phase d'initialisation permet une localisation extrêmement rapide des caractéristiques anatomiques du fémur à segmenter. Un modèle déformable simple, tiré du domaine de l'animation de synthèse est ensuite recalé sur l'imagerie CT. Ce modèle sert de connaissance a priori pour la segmentation fine par contours actifs. Il est à noter que'en cas d'erreur, le praticien peut corriger la segmentation tout au long du processus.

Les chapitres 3 et 4 présentent différentes techniques de recalage rigide et non rigide ainsi que les modèles déformables paramétriques couramment utilisés. Le chapitre 6 introduit le modèle déformable employé dans ce travail. Celui-ci est basé sur les cylindres généralisés combiné aux modèles articulés. Les processus de segmentation, à proprement parlé, sont introduits au chapitre 5, pour la localisation des paramètres anatomiques, et au chapitre 7 pour la segmentation basée sur la programmation dynamique. Le dernier chapitre 8, présente une synthèse globale des résultats.

CHAPITRE III

RECALAGE RIGIDE

3.1 Introduction

Le recalage (*registration*) est le processus d'estimation d'une transformation optimale entre deux jeux de données. Le recalage est dit rigide si la transformation géométrique comprend une rotation et une translation. Le recalage est dit non-rigide si la transformation géométrique est d'ordre supérieur (affine, polynôme, splines...) ou si la transformation n'est pas paramétrique (voir chapitre 4).

Maintz ^[68] propose de classer les méthodes de recalage médical suivant neuf critères :

1. **La dimension** (spatio-temporelle) entre les données ;
2. **Le moyen de mise en correspondance** est soit intrinsèque, soit extrinsèque. Dans le cas intrinsèque, seul le contenu des données du patient sont utilisés lors du recalage. Dans le cas extrinsèque, des objets artificiels sont fixés sur le patient préalablement à l'imagerie ;
3. **La nature de la transformation.** Nombre de degrés de libertés (DDL) impliqués dans la transformation. Le cas le plus simple est la transformation rigide (rotation et translation) qui possède 4 DDL dans l'espace bidimensionnel et 6 DDL en 3D ;
4. **Le domaine de la transformation.** Celui-ci est soit global, soit local ;
5. **L'interactivité** définit le niveau d'automatisation de l'algorithme de recalage ;
6. **La méthode de recherche des paramètres de transformation.** Ceux-ci sont soit calculés au moyen d'une méthode directe (souvent linéaire), soit par optimisation ;

7. **Les modalités impliquées.** Lorsque les données proviennent du même imageur, il s'agit d'un recalage monomodal. Si les données sont issues de deux imageurs distincts, il s'agit de données multimodales. A noter que le patient "physique" peut également être considéré comme une modalité. Par exemple, lors d'une navigation chirurgicale, le recalage met en relation d'imagerie scanner avec le patient réel ;
8. **Le sujet.** Lorsque toutes les données impliquées lors du recalage proviennent du même sujet, il s'agit d'un recalage intra-sujet. Dans le cas de deux sujets distincts, il s'agit d'un recalage inter-sujets. Un des jeux de données peut également être construit sur la base de plusieurs patients. Il s'agit dans ce cas d'un recalage statistique ou par atlas ;
9. **L'objet** est la partie anatomique à recalier ;

Pour plus d'informations sur cette classification ainsi qu'un état de l'art complet, le lecteur est prié de se référer aux articles de Maintz ^[68] et van den Elsen ^[99].

Le tableau 3.1 présente le contenu de ce travail classé par rapport aux critères décrits précédemment.

| | | |
|------------------------------|---------------------------|---------------------|
| Dimension des données | 3D | |
| Mise en correspondance | intrinsèque | |
| Nature de la transformation | modèle déformable | chapitres 3, 4 et 6 |
| Domaine de la transformation | globale puis locale | chapitre 4 |
| Interactivité | (semi-)automatique | chapitre 5 |
| Recherche des paramètres | directe | chapitres 3, 4 et 6 |
| Modalités impliquées | monomodale (CT) | chapitres 5 et 7 |
| Sujet | inter-sujets | |
| Objet | partie proximale du fémur | annexe 3 |

TAB. 3.1: Classification de ce travail selon les critères de Maintz ^[68].

Ce chapitre présente le recalage rigide. Il s'agit du type de transformation le plus basique qui est le fondement de techniques plus complexes introduites au chapitre 4.

3.2 Recalage point à point

La forme la plus simple de recalage est la mise en correspondance de deux référentiels par des paires de points. La transformation résultante peut

être représentée par d'une transformation rigide composée d'une rotation (R) et d'une translation (\vec{t}). Historiquement, ce recalage est appelé, en photogrammétrie, estimation de pose (*pose estimation*). En 3D, la translation possède évidemment trois degrés de liberté (DDL) et la rotation trois autres (pour plus d'informations, voir l'annexe 2). Dans certains cas, l'utilisation du facteur d'échelle (s) peut rajouter un septième DDL. Trois paires de points non colinéaires connus dans chaque référentiel offrent neuf contraintes, ce qui est théoriquement suffisant pour déterminer les sept inconnues de cette transformation [98]. Pratiquement, il est judicieux d'utiliser plus de points durant le recalage pour compenser l'inexactitude des mesures. Dans ce cas, le problème n'a plus de solution exacte et une résolution par les moindres carrés est nécessaire pour calculer la transformation optimale.

Le problème de recalage de paires de points peut être posé comme suit : Supposons les coordonnées 3D des points \vec{d}_i et leur correspondants \vec{m}_i dans l'autre référentiel. Leur relation, sans prendre en compte le facteur d'échelle, est :

$$\vec{m}_i = R\vec{d}_i + \vec{t}$$

Le calcul de la détermination de la pose revient à déterminer R et \vec{t} à partir des paires d'appariement (\vec{m}_i, \vec{p}_i) . Avec au moins 3 points de référence non colinéaires, R et \vec{t} peuvent être obtenues par résolution du problème des moindres carrés suivant :

$$\min\left(\sum_i (\vec{m}_i - R\vec{d}_i + \vec{t})\right) \quad (3.1)$$

En ajoutant les contraintes d'orthogonalité de la matrice R , soit $R^t R = R R^t = R R^{-1} = I$, le problème devient insoluble avec les techniques directes d'algèbre linéaire. Historiquement, Horn [46] et Arun [6] sont les premiers en 1987 à proposer une formulation analytique (*closed-form solution*) et non-itérative à ce problème. Horn résout le problème en représentant la matrice de rotation par un quaternion et Arun par une décomposition en valeurs singulières (*Singular Value Decomposition - SVD*). Voici, dans les grandes lignes, l'algorithme basé sur la SVD.

Tout d'abord, il est nécessaire de translater les jeux de données par rapport à leurs centres de gravité. Les points sont ensuite décrits dans les

référentiels m' et d' .

$$\begin{aligned}\vec{m} &= \frac{1}{N_m} \sum_i \vec{m}_i & \vec{d} &= \frac{1}{N_d} \sum_i \vec{d}_i \\ \vec{m}'_i &= \vec{m}_i - \vec{m} & \vec{d}'_i &= \vec{d}_i - \vec{d}\end{aligned}$$

où N_m et N_d représentent le nombre de points des deux jeux de données respectifs. La matrice C d'inter-covariance entre les points \vec{d}'_i et \vec{m}'_i peut être définie comme :

$$C = [\vec{d}'_1 \vec{d}'_2 \dots][\vec{m}'_1 \vec{m}'_2 \dots]'$$

Horn ^[47] démontre que si R et \vec{t} minimisent l'équation 3.1, alors les équation suivante sont satisfaites :

$$\begin{cases} R = UV' \\ T = \vec{d} - R\vec{m} \end{cases} \quad (3.2)$$

où U et V sont les matrices obtenues par décomposition en valeurs singulières (SVD) tel que $C = USV'$.

Certaines implémentations proposent un facteur correcteur de rotation :

$$R = U \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & |UV'| \end{pmatrix} V'$$

Numériquement, ces deux derniers algorithmes ne présentent pas de différences significatives mais tout deux sont limités par le fait qu'un appariement exact entre les deux jeux de données est nécessaire. Il est possible de lever partiellement ce handicap en utilisant un algorithme de filtrage de points aberrants de type RANSAC ^[18] ou *softassign* ^[22]. Dans les deux cas, ces algorithmes sont inadaptés lorsque plusieurs centaines de points dont de nombreux aberrants sont présents dans les jeux de données. On préférera dans ce cas l'algorithme ICP présenté à la section suivante.

3.3 Algorithme du plus proche point itéré (ICP)

L'algorithme du plus proche point itéré ou ICP (*Iterative Closest Point*), que l'on devrait de préférence appeler algorithme du plus proche point itéré au moyen d'une correspondance itérative de points, est largement utilisé dans l'alignement géométrique de modèles tridimensionnels lorsqu'une approximation initiale de la pose est connue. Depuis son introduction au début des années 90 par Besl et McKay ^[10] ou une variante indépendante proposée par Chen et Medioni ^[20], l'algorithme ICP s'est décliné sous un grand nombre de formes. Nous ne présentons ici qu'un sous-ensemble nécessaire pour comprendre la suite de ce travail. Le lecteur pourra se reporter à l'état de l'art de Rusinkiewicz et Levoy ^[85] plus complet, qui présente également une bonne comparaison expérimentale des différentes approches.

L'algorithme ICP est largement utilisé pour le recalage de maillages tridimensionnels. Il débute avec deux nuages de points, généralement appelés modèle et données (ou scène), ainsi qu'une approximation initiale de leur transformation relative et va itérativement raffiner cette transformation en générant des paires de points correspondants entre les deux mailles par minimisation de l'erreur métrique.

L'alignement initial peut être réalisé au moyen d'un grand nombre de techniques différentes comme les approches extrinsèques (par exemple en connaissant la position relative de l'imageur 3D entre les acquisitions successives), l'identification et l'indexation des caractéristiques de surfaces ^[32], les signatures locales de surfaces par *spin-image* ^[55], par décomposition en axes principaux ^[29], par une recherche exhaustive des correspondances de paires de points ^[18] ou manuellement.

Synthétiquement, l'algorithme ICP nécessite la répétition alternative des étapes suivantes :

1. La **sélection** d'un sous-ensemble de points dans le modèle et/ou les données ;
2. L'**appariement** des points de données sélectionnés sur le modèle. Ce processus génère des paires de points ;
3. L'attribution de **poids** pour chaque paire appariée ;
4. **Le rejet** individuel de certaines paires (*outliers*) ou le rejet global de l'appariement ;

5. L'assignation d'une **erreur métrique** basée sur les paires de points retenus ;
6. La **minimisation** de l'erreur métrique par la mise à jour de la transformation.

3.3.1 Notation

Le modèle (resp. données) est noté M (resp. D) et constitué de n_M (resp. n_D) points ou primitives. Nous noterons les points du modèle \vec{m}_j et les points du jeu de données \vec{d}_i . Les transformations rigides seront notées $T = (R, \vec{t})$ (voir l'annexe 2) et appliquées au jeu de données (pour la ramener dans la même position que le modèle) de la façon suivante : $T \star \vec{d}_i = R\vec{d}_i + \vec{t}$. La véritable transformation qui sépare les deux ensembles est notée T_0 . La distance utilisée dans le critère sera notée Δ . Il s'agira, au début, de la distance euclidienne quadratique $\Delta^2(x, y) = \|x - y\|^2$.

3.3.2 Métrique entre les jeux de données

Le principe de fonctionnement de l'algorithme ICP est de minimiser la distance entre les données et le modèle. Cette distance se mesure en calculant pour chaque point d'un jeu de données, le point le plus proche dans l'autre jeu.

Initialement, la distance de mesure ou métrique proposée par Besl ^[10] utilise la distance quadratique du plus proche voisin, soit :

$$\Delta^2(T \star \vec{d}_i, M) = \min_j \Delta^2(T \star \vec{d}_i, \vec{m}_j)$$

Plus généralement, le critère de distance entre les deux jeux de données peut s'écrire :

$$\Delta^2(T \star D, M) = \sum_i \Delta^2(T \star \vec{d}_i, M) \quad (3.3)$$

Réciproquement, la distance entre le modèle et les données peut être calculé de manière similaire pour autant que la transformation T soit inversible.

Par souci de symétrie, le critère de distance entre deux ensembles de points peut s'écrire :

$$\Delta_{sym}^2(D, M) = \Delta^2(T \star D, M) + \Delta^2(M, T \star D) \quad (3.4)$$

Il est à signaler que d'autres mesures de distances (*kernel*) sont couramment utilisées et peuvent, dans certains cas, améliorer significativement la convergence de l'algorithme ICP. Le principe de ces mesures est d'atténuer l'importance des points aberrants qui, au cours de la convergence d'ICP, ont une tendance à avoir une distance de plus en plus grande par rapport aux points appariés. Parmi les mesures les plus répandues, on peut citer, la mesure linéaire Δ , ou encore la mesure de Lorentz :

$$\Delta_{Lorentz}(r) = \log\left(1 + \frac{r^2}{\sigma}\right); \quad (3.5)$$

où r est la distance euclidienne entre les points et σ est un paramètre de réglage. La mesure de Hubert (équation 3.6) utilise une distance quadratique lorsque les points sont rapprochés et linéaire lorsque ceux-ci sont éloignés :

$$\Delta_{Huber}(r) = \begin{cases} r^2 & r < \sigma \\ 2\sigma|r| - \sigma^2 & \text{autrement} \end{cases} \quad (3.6)$$

Les différents types de mesures de distance sont comparés à la figure 3.1. Pour plus d'informations concernant ces mesures, se référer à [37]. Parmi les autres types, citons la distance de Mahalanobis, privilégiée dans le travail de thèse de Granger [42], qui permet d'une part de prendre en compte le bruit sur les points (ce qui est particulièrement intéressant lorsque celui-ci est inhomogène ou anisotrope), et d'autre part de généraliser facilement l'algorithme à des types de primitives plus complexes. Ce dernier type de distance ne sera pas développé dans le cadre de ce travail.

3.3.3 Méthodes d'appariement

La mise en correspondance (ou appariement) de primitives, plus spécifiquement dans le cas du recalage non rigide, est un problème non trivial. La présence de bruit ou le fait que certaines primitives dans un jeu de données

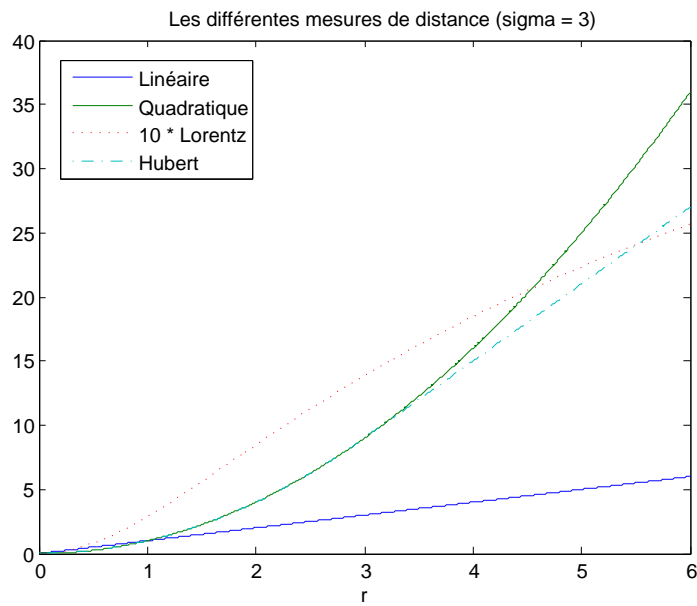


FIG. 3.1: Comparaison des différentes mesures de distances pour $\Delta(r)$ (linéaire), $\Delta^2(r)$ (quartique), $\Delta_{Lorentz}(r)$ (Lorentz) et $\Delta_{Hubert}(r)$ (Hubert) avec $\sigma = 3$. A noter que, pour plus de visibilité, la distance de Lorentz est multipliée sur le graphique par un facteur 10.

peuvent ne pas avoir de correspondance dans l'autre (système non homologue) impliquent que certains appariements doivent être pondérés ou rejetés. La pondération peut soit être gérée au niveau de la métrique (voir distances de Hubert ou Lorentz au chapitre 3.3.2), soit au moyen de la technique d'appariement.

La méthode d'appariement la plus simple est la méthode des plus proches voisins présentée dans le chapitre précédent. Cette technique présente de bons résultats lorsque les deux jeux de données sont homologues, c'est-à-dire que :

$$\forall d_i \in D \quad \exists! m_j \in M \quad | \quad T \star d_i = m_j \quad (3.7)$$

Cette hypothèse est évidemment trop restrictive en présence de points aberrants (*outliers*). Reprenons l'exemple précédent et représentons les appariements sous la forme d'une matrice de correspondance binaire ou matrice de permutation (Z) :

| Z | d_1 | d_2 |
|-------|-------|-------|
| m_1 | 1 | 0 |
| m_2 | 0 | 1 |

TAB. 3.2: Matrice de correspondance binaire homologue. Dans cet exemple $T \star d_1 = m_1$ et $T \star d_2 = m_2$.

Chui et Rangarajan ^[22] proposent une méthode combinée d'appariement et d'estimation de transformation au moyen de l'algorithme *softassign* et une technique de recuit simulé. Le recuit simulé est une méta-heuristique inspirée d'un processus utilisé en métallurgie. Ce processus alterne des cycles de refroidissement lent et de réchauffage (recuit) qui tendent à minimiser l'énergie du matériau. Elle est aujourd'hui utilisée en optimisation pour trouver les extrema d'une fonction. Les auteurs présentent différents exemples pour le recalage non rigide de sulcatures (*sulcal*) corticales.

Le principe de *softassign* (^[41] ou ^[81]) est de relaxer une matrice de correspondance binaire Z en une matrice doublement stochastique continue M , tout en renforçant itérativement les contraintes de lignes et de colonnes. La nature continue de la matrice de correspondance M permet de gérer de manière floue (*fuzzy*) le poids des appariements entre les jeux de données. Du point de vue de l'optimisation, l'aspect flou permet un meilleur comportement du recalage car l'appariement s'améliore graduellement et continuelle-

ment au cours de itérations, sans pour autant effectuer des sauts brusques (phénomène typique des matrices de permutations). Les contraintes de lignes et de colonne se renforcent au moyen d'une normalisation itérative des lignes et des colonnes.

La notion de correspondance floue permet d'introduire le concept de recuit simulé déterministe ([78] et [103]) par l'ajout d'un terme d'entropie et le paramètre de température. Le processus itératif du recalage commence par une température élevée, où le terme d'entropie force la correspondance entre les points à être floue. Au cours des itérations, la température est réduite et la matrice continue M tend à devenir une matrice de correspondance binaire Z . Cette technique permet de réduire les chances de tomber dans un minimum local durant l'optimisation.

3.3.3.1 Points aberrants

Les points aberrants impliquent l'invalidation de l'équation 3.7. Certains points des données n'ont plus de correspondance dans le modèle et réciproquement. Un moyen simple de gérer partiellement ces points est de supprimer le critère de symétrie présenté dans l'équation 3.4. Dans ce cas, l'appariement s'effectuera du jeu de données le moins complet au plus complet :

$$\Delta_{aberration}(D, M) = \begin{cases} \Delta(T \star D, M) & n_D > n_M \\ \Delta(M, T \star D) & n_M > n_D \end{cases}$$

L'algorithme de recuit simulé mentionné au chapitre précédent (3.3.3) implique l'utilisation de matrices carrées. Comme les jeux de données réel présentent un nombre différent de points ($n_D \neq n_M$), il est nécessaire d'introduire des *slack variables* ([23], [12]) pour compléter les lignes et colonnes manquantes de la matrice. Cette méthode permet de gérer de manière robuste les points aberrants.

L'annexe 1 présente exemple concret de l'utilisation de cette méthode d'appariement avec l'utilisation des *slack variables*.

3.3.4 Types de mise en correspondance

Les points sont souvent la forme la plus fondamentale des primitives utilisées lors de l'appariement. Ils peuvent également être la base de représentations plus complexes comme les surfaces ou triangles.

3.3.4.1 Correspondance point à point

La manière la plus simple de générer les paires de points, comme nous l'avons du reste fait jusqu'à présent, est l'appariement par le point le plus proche (voir équation 3.3). Pratiquement, les points de données et du modèle ne sont pas homologues, ce qui peut limiter cet algorithme. Par exemple, la figure 3.2 présente l'ambiguïté de l'appariement du point d_1 avec les points m_1 ou m_2 . Faut-il considérer l'appariement entre l'un ou l'autre des points du modèle ou une sorte d'interpolation entre ces deux points (comme la projection du point d_1 sur l'axe formé par m_1 et m_2) ? La résolution spatiale des jeux de données est un autre facteur réduisant linéairement la qualité de ce mode de correspondance. Finalement, le manque d'uniformité des points sur les surfaces à recaler biaise l'algorithme en favorisant les régions où la densité est grande.

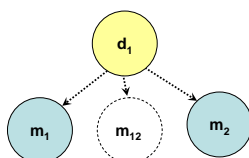


FIG. 3.2: Ambiguïté liée à la correspondance points à points. Est-ce que le point d_1 est apparié au point m_1 , m_2 ou à une interpolation entre les deux comme par exemple le point m_{12} ?

La mise correspondance peut être grandement améliorée soit en considérant un appariement flou (présenté au chapitre 3.3.3), méthode qui joue sur le poids des appariements pour créer une sorte d'interpolation, soit si d'autres caractéristiques liées au(x) jeu(x) de données peuvent être utilisées comme la surface. Celle-ci rajoute la dimension de voisinage et de courbure locale. Plus spécifiquement, une surface discrète peut être représentée par

un maillage (*mesh*) composé de points, reliés par des triangles orientés (c'est à dire que l'on connaît l'intérieur et l'extérieur d'une surface fermée). Cette représentation permet l'utilisation de méthodes de correspondance plus fines qui seront présentées dans les sections suivantes.

3.3.4.2 Correspondance point à plan

La correspondance point à plan a été introduite en 1992 par Chen [20]. Il s'agit d'assigner un plan (orienté dans le cas d'une surface fermée) à chaque point du jeu de données. Dans le cas discret, l'appariement utilisé est celui du plus proche voisin sauf que le calcul de distance n'est plus la distance euclidienne quadratique $\Delta^2(x, y)$ comme présenté au chapitre 3.3.1 mais la distance entre le point x et sa projection sur l'autre surface :

$$\Delta_{plan1}(x, y) = \vec{x}y \cdot \vec{n}(y) \quad (3.8)$$

où y est la correspondance point à point de x et $\vec{n}(y)$ est le vecteur unitaire, normal à la surface au point y . On dit aussi que y est un point orienté (muni d'un vecteur unitaire qui représente la normale à la surface).

Une variante de cet algorithme est, comme précédemment, de trouver y par une correspondance point à point. Ensuite, on ne considère que les triangles du maillage dont le point y fasse partie. Si $tri(K)$ est l'ensemble de ces triangles, alors cette distance peut s'exprimer comme :

$$\Delta_{plan2}(x, y) = \min_{k \in K} (\vec{x}y \cdot \vec{n}(tri(k))) \quad (3.9)$$

où $\vec{n}(tri(k))$ est le vecteur unitaire, normal au triangle k .

La figure 3.3 présente le résultat de cette deuxième méthode de correspondance point à plan avec l'exemple introduit à la figure 3.2.

Comparativement à l'approche point à point, la vitesse et la qualité de convergence de l'algorithme point à plan sont incomparables. L'excellent article de Rusinkiewicz sur les variantes efficaces d'ICP [85] présente une comparaison des deux approches. La figure 3.4 est une reproduction de ses résultats sur la vitesse de convergence en utilisant les mesures de distance

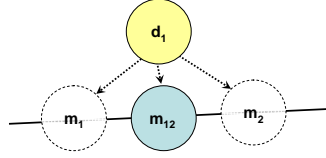


FIG. 3.3: Résultat de la correspondance point à plan de la figure 3.2. Dans cette seconde variante, $d_1 \vec{m}_{12} \perp m_1 \vec{m}_2$. La distance $\Delta_{plan2}(x, y)$ est la norme du vecteur $d_1 \vec{m}_{12}$.

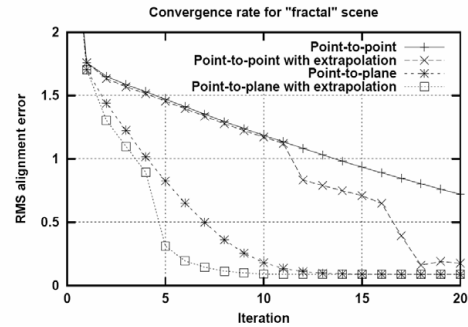
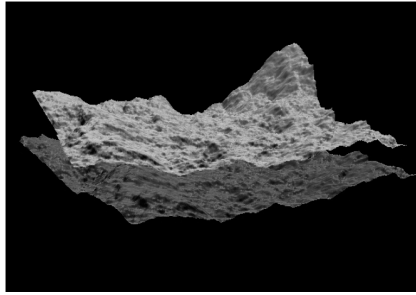


FIG. 3.4: Comparaison de convergence entre la méthode point à point et point à plan. A gauche, les deux jeux de données utilisés par Rusinkiewicz [85] et à droite, les résultats correspondants en fonction de la méthodes d'appariement. Expérimentalement, nous obtenons le même genre de courbes.

point à point et point à plan ($\Delta_{plan1}(x, y)$). On remarque une convergence quasi-linéaire pour la première et de forme $\frac{1}{x}$ pour la seconde !

Malgré la nette amélioration de la technique point à plan, celle-ci apporte quelques limitations dans les conditions de bord. Dans le cas où les jeux de données ne se superposent pas parfaitement, les correspondances impliquant les points en bordure peuvent introduire un biais systématique lors de l'appariement. Ce phénomène est illustré à la figure 3.5.

La correspondance point à triangles, introduit au chapitre suivant, tente de remédier à ce problème.

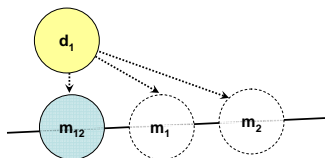


FIG. 3.5: Conditions de bord dans la méthode point à plan. Dans ce cas, la distance $\Delta_{plan2}(x, y) = \|d_1 \vec{m}_{12}\|$ introduit un biais. La distance correcte serait de considérer $\|d_1 \vec{m}_1\|$.

3.3.4.3 Correspondance point à triangle

La méthode point à triangle essaye de remédier aux conditions de bord introduites par la méthode précédente. Il s'agit d'une technique combinant la méthode point à point et point à plan. L'algorithme commence par un calcul d'appariement point à plan. Une fois le point y et les triangles connexes $tri(k)$ trouvés, un test supplémentaire est réalisé pour vérifier si la projection du point x est contenue à l'intérieur d'un triangle connexe. La distance point à plan est utilisée dans l'affirmative (figure 3.6), sinon c'est la distance point à point (figure 3.6).

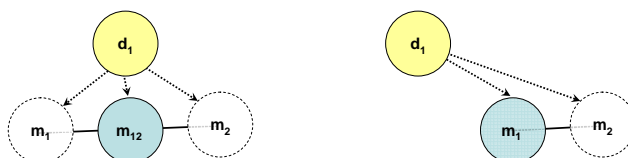


FIG. 3.6: Correspondance point à triangle. A gauche, cas où la distance point à plan est utilisée, à droite, cas où la distance point à point est utilisée.

Dans son article sur la reconnaissance des oreilles en 3D basée sur ICP, Yan ^[101] propose de comparer l'approche point à point et point à triangle. La principale conclusion de l'article est que le manque de performances de l'approche point à triangle (environ 5 fois moins rapide que l'autre méthode) est compensée par une meilleure stabilité dans la différence d'échantillonnage des deux jeux de données. L'approche point à triangle est donc beaucoup plus

adaptée dans le cas d'un appariement hiérarchique (voir la section 4.4).

3.3.4.4 Autres modèles pour la correspondance

Dans sa thèse, Delingette [26] propose une représentation particulière des surfaces discrètes connues sous le nom de maillages simplexes. Ceux-ci sont topologiquement duaux aux maillages triangulés. Un maillage simplexe présente une structure régulière dont chaque sommet possède exactement trois voisins (figure 3.7), des paramètres métriques $\epsilon_1, \epsilon_2, \epsilon_3$ et un angle simplexe ϕ_1 . Chaque face d'un maillage possède également une orientation. Cette représentation possède de nombreux opérateurs de régularisation topologique qui peuvent être implémentés de manière efficace. Dans sa thèse, Montagnat [74] utilise les maillages simplexes comme modèles déformables pour le recalage non rigide.

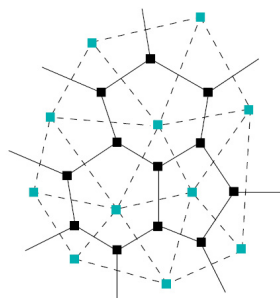


FIG. 3.7: Exemple de maillage simplexe, initialement proposé par Delingette [26], en noir est sa représentation duale en bleu (trait-tillé).

Il existe également de nombreuses autres techniques qui permettent de raffiner le modèle de correspondance point à triangle. Il s'agit la plupart du temps d'interpoler localement une surface explicite au maillage, comme une B-Spline, une fonction à base radiale (*RBF - Radial Basis Function*), une quadrique [80] ou une super quadrique. Ces modèles, excepté les RBF (chapitre 4), ne seront pas décrits dans cette thèse. Le lecteur intéressé trouvera plus de détails dans les travaux de Montagnat [74] ou Ilic [49].

Ce descriptif ne serait pas exhaustif sans l'introduction des *spin images* [55], [2]. Cette représentation ne requiert pas de connaissances a priori sur la transformation entre les jeux de données. Les *spin images* comprennent, pour chaque point orienté de la surface, une image décrivant localement

le maillage. La corrélation entre ces images permet d'établir les correspondances entre les différents points des jeux de données. La figure 3.8, tirée de [2] illustre cette représentation.

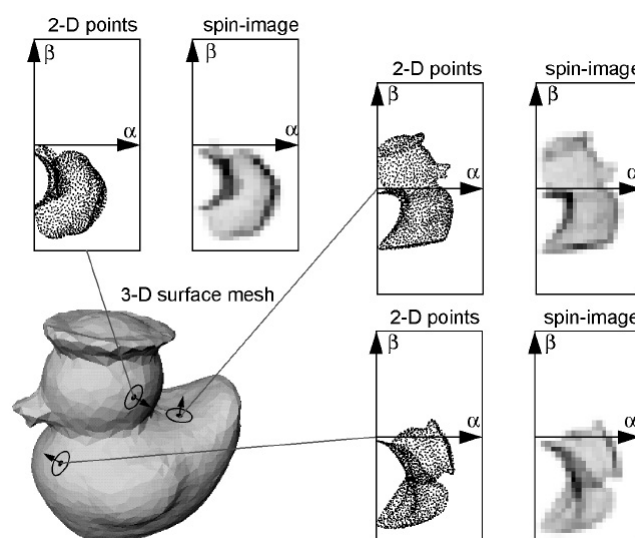


FIG. 3.8: Exemple de *spin images*. Pour chaque point orienté de la surface, les points du jeu de données sont localement projetés et représentés par une *spin-image*, sorte de *Maximum Intensity Projection*. Les points des deux jeux de données sont ensuite mis en correspondance par la corrélation des *spin images*.

3.3.5 Minimisation de l'erreur métrique

Nous avons vu au début du chapitre 3.3 que la minimisation de l'erreur métrique de l'algorithme ICP est réalisée par la mise à jour successive des correspondances et de la transformation. Dans la formulation initiale de Besl [10], la transformation est calculée à chaque itération par minimisation la fonction objective (équation 3.3) par une résolution algébrique. Cette optimisation s'effectue de la manière suivante :

Dans un premier temps, la matrice d'inter covariance C est calculée comme pour l'algorithme de recalage point à point de Arun (voir équation 3.2). Le calcul de la nouvelle rotation est donné par le calcul des valeurs propres de la matrice $4 * 4$ suivante :

$$Q(C) = \begin{pmatrix} \text{tr}(C) & \delta^T \\ \delta & C + C^T - \text{tr}(C)I_3 \end{pmatrix}$$

où I_3 est la matrice identité, $\delta = [A_{23}A_{31}A_{12}]^T$ avec $A_{ij} = (C - C^T)_{ij}$. Les valeurs propres de cette matrice $Q(C)$ représentent le nouveau quaternion de rotation optimal (voir annexe 2). La translation \vec{t} optimale peut être calculée directement au moyen de l'équation :

$$\vec{t} = \vec{m} - R\vec{d}$$

Besl présente, dans le même article, une preuve de la convergence monotone de cet algorithme vers un minimum local.

De par sa nature, l'algorithme ICP peut également être posé comme un problème d'optimisation non linéaire :

$$\min_T \{\Delta(T \star D, M)\} \quad (3.10)$$

où T est la transformation . Nous verrons plus loin que la transformation rigide peut être généralisée à des transformations possédant plus de degrés de liberté.

Fitzgibbon ^[37] propose de remplacer la minimisation de la fonction objective de Besl ^[10] par une optimisation fondée sur Levenberg-Marquardt (LM-ICP). Selon l' auteur, cette résolution apporte de nombreux avantages. Premièrement, le recalage rigide est plus rapide. Deuxièmement la convergence est bien meilleure (voir figure 3.9). Troisièmement, la transformation n'est pas limitée au recalage rigide. Ce point sera plus détaillé au chapitre 4.

Suivant le type de transformation utilisé, la contrainte de dérivabilité de la fonction objective liée à LM-ICP n'est plus forcément respectée. Il est, dans ce cas, plus approprié d'utiliser de la méthode d'optimisation des simplexes, inventée par George Dantzig en 1947.

Comme tout problème d'optimisation possède de nombreuses voies de résolution, il est également légitime de citer la méthode de Monte-Carlo et celle des algorithmes génétiques, explorée récemment par Silva ^[67]. D'après cette étude ainsi que quelques tests effectués par l'auteur, cette approche permet

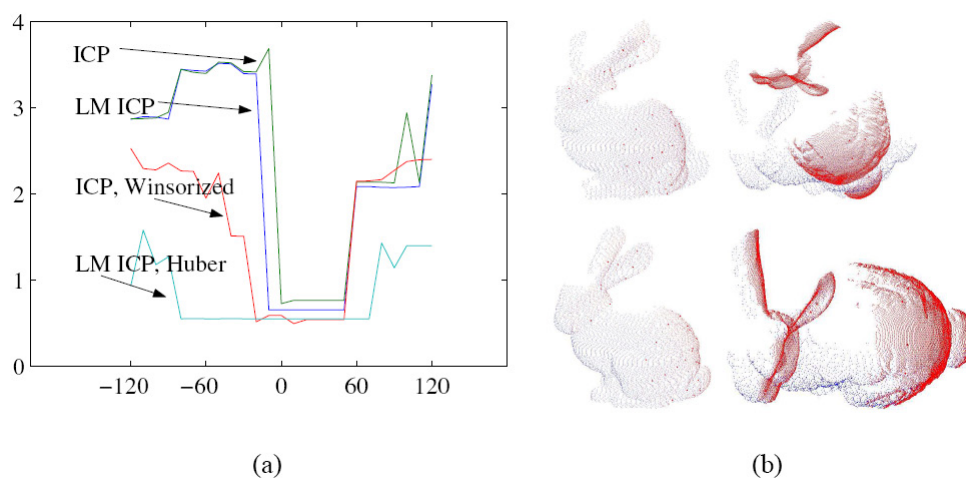


FIG. 3.9: Bassin de convergence de l’algorithm ICP, LM-ICP et LM-ICP avec la mesure de Hubert (Δ_{Huber}) sur les jeux de données (b) selon ^[37], les algorithmes sont initialisés à un degré d’intervalle entre $\pm 120^\circ$ de la solution correcte, le graphique (a), présente les minimums obtenus après ICP pour chacun des algorithmes en fonction de l’angle initial.

une meilleure gestion des valeurs aberrantes sans améliorer substantiellement la convergence d’ICP.

3.3.6 Limitations et améliorations d’ICP

Malgré son approche générique et universelle, l’algorithme ICP possède certaines limitations relatives à sa convergence erratique ainsi que son temps de calcul élevé lié à son aspect itératif. Les sections suivantes présentent ces principales limitations et proposent quelques pistes pour y remédier.

3.3.6.1 Minimum locaux

Comme nous l’avons vu au chapitre précédent, la méthode de minimisation de l’erreur métrique ainsi que de mise en correspondance définissent qualitativement la convergence de l’algorithme de recalage. Dans certain cas, celui-ci aboutit à un minimum local non-optimal (voir figure 3.10). Pour limiter ce phénomène, Simon ^[89] et Madsuda ^[70] proposent de modifier aléatoirement et partiellement la transformation avant (respectivement après) la

convergence. Cette astuce permet, dans certains cas, de quitter un minimum local. Conceptuellement, l'approche des algorithmes génétiques^[67] utilise les même "feintes" par l'emploi des mutations génomiques.

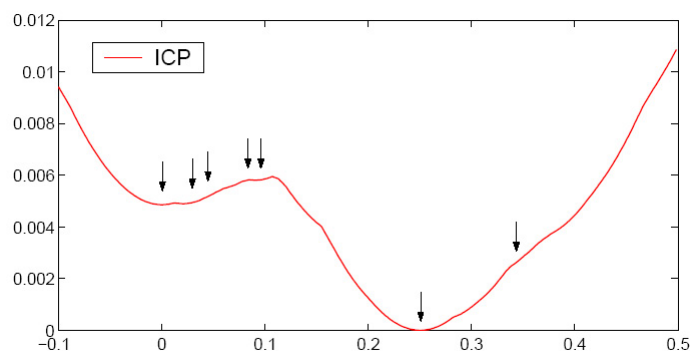


FIG. 3.10: Exemple de minimum local lors du recalage, illustration extraite du travail de Granger^[42], la courbe rouge représente l'erreur de recalage en fonction de la translation initiale sur l'axe de profondeur (Z) des données par rapport au modèle.

3.3.6.2 Bassins de convergence

Pour comprendre les propriétés globales de convergence, il est nécessaire de définir le bassin de convergence (*funnel of convergence*) d'un algorithme de recalage comme l'ensemble de poses initiales des données qui convergent vers le modèle sans, par exemple, être stoppées par un minimum local. A signaler que cette mesure ne s'intéresse qu'à la convergence globale et non à sa vitesse. Un large bassin, comme l'algorithme *LM-ICP Hubert* présenté à la figure 3.9 indique que l'algorithme gère avec succès un éventail de poses initiales. On dit qu'un algorithme possède un bassin stable si les zones de convergence sont regroupées en non pas distribuées arbitrairement. Cette stabilité est désirable pour un bon recalage, car elle permet de systématiser les conditions initiales. La figure 3.11, tiré du travail de Mitra^[73], illustre le bassin convergence très instable de l'algorithme point à point ICP.

3.3.6.3 Vitesse de convergence

Dans sa version la plus simple, l'algorithme de mise en correspondance point à point tel que décrit dans l'équation 3.3 calcul pour chaque point des

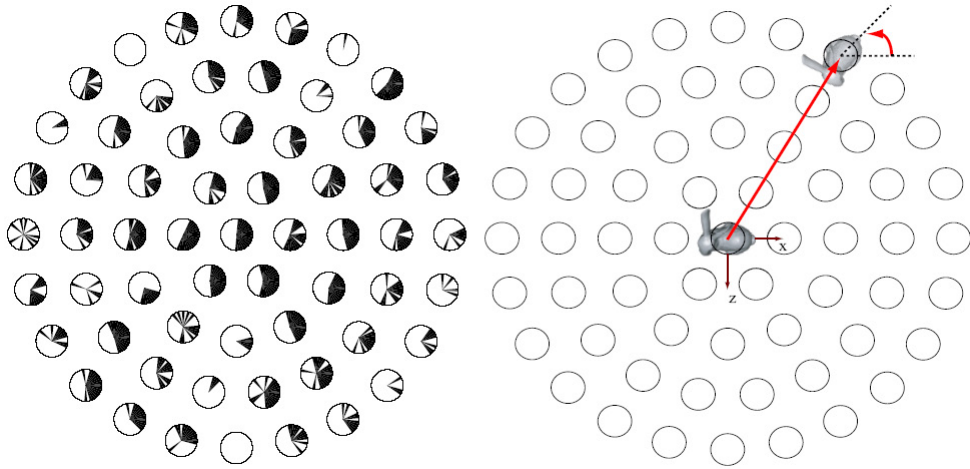


FIG. 3.11: Figure tirée du travail de Mitra ^[73] : Bassin de convergence planaire de l'algorithme de recalage point à point ICP sur le jeu de données *bunny*. À droite, exemple de position initiale avec une translation (x, y) et une rotation α , à gauche, les secteurs foncés des camemberts représentent les triplets (x, y, α) où l'algorithme converge, les secteurs blancs divergent vers un minimum local

données la distance avec tous les points du modèle. La complexité de cet algorithme est de l'ordre de $O(n^2)$ ou plus spécifiquement $O(n_M * n_D)$.

L'algorithme des *kd-trees* introduit par Friedman en 1977 ^[39] propose une autre méthode d'indexation des données afin d'améliorer la vitesse de calcul du point le plus proche. Cet algorithme segmente récursivement l'espace de recherche (comme par exemple le nuage de points du modèle) en sous-espaces délimités par des hyperplans. Ce processus peut être représenté par un arbre binaire dont chaque noeud représente un hyperplan (voir figure 3.12). De manière simplificatrice, la recherche du point d_i au point le plus proche du jeu de donnée M s'effectue en regardant récursivement à partir de la racine dans quel sous-espace se trouve le point d_i . Une description très détaillée de l'algorithme ainsi qu'un tutoriel sont décrits dans la thèse de Moore ^[75]. Le grand avantage de l'algorithme des *kd-trees* est de réduire la complexité de la mise en correspondance à $O(n_D * \log_2(n_M))$.

De par sa nature itérative, la vitesse de convergence de l'algorithme ICP dépend de nombreux facteurs. Parmi les principaux, on peut citer la méthode de mise en correspondance (voir figure 3.4) ou le type d'algorithme utilisé lors de l'optimisation. Bien que, selon Fitzgibbon ^[37], l'algorithme LM-ICP,

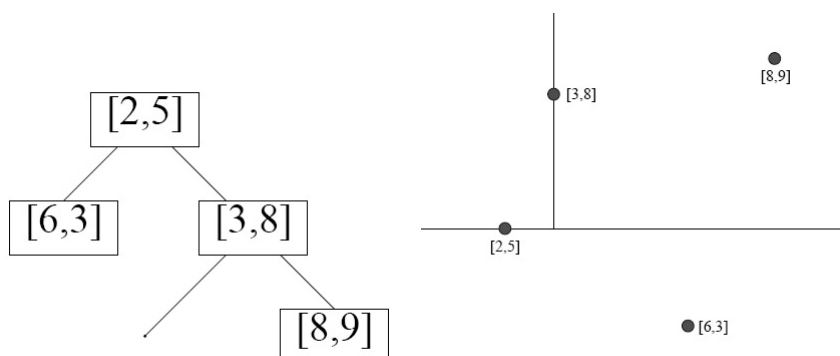


FIG. 3.12: Exemple de *kd-tree* tiré de [75], à gauche, arbre binaire et à droite, représentation des hyperplans correspondants.

ait une vitesse de convergence plus rapide que l'algorithme original dans le cas rigide, le nombre de calcul de correspondance par itération dépend du nombre de variables dans la transformation (x_t). En effet, le principe d'optimisation de Levenberg-Marquardt nécessite le calcul du Jacobien de la fonction objective, c'est-à-dire de $x_t + 1$ mises en correspondances, ce qui peut s'avérer très coûteux en temps de calcul lorsque le nombre de variables dans la transformation augmente. Des stratégies génériques d'amélioration de vitesse sont proposés dans la littérature comme l'approche multi-échelle [56], [42], [21], ou l'approche par champs de distance [37], [80], [73].

Le principe de l'approche multi-échelle est de réaliser les premières itérations en utilisant un jeu de donnée sous échantillonné et d'augmenter progressivement la résolution au cours des itérations, passant ainsi d'un recalage approximatif à un recalage fin. L'avantage principal de la multi-résolution est la réduction du temps de calcul global tout en réduisant le nombre d'itération car un calcul à basse résolution implique généralement une mise à jour plus importantes des facteurs de rotation et de translation. Pour plus d'informations concernant cette technique, se référer au travail de thèse de Jost [56].

L'approche des champs de distance consiste à coder pour tout point de l'espace la distance d'un objet à ce point (voir figure 3.13). La distance d'un point à un jeu de données peut ainsi directement être calculé sans nécessiter la phase de recherche de correspondances. Un tel algorithme nécessite une phase d'encodage de l'espace mais, en contrepartie, a une complexité de $O(n_D)$.

Généralement, l'implémentation de cet algorithme en 2D encode la distance de Hausdorff de chaque pixel par rapport au jeu de données [37]. Par extension, le même algorithme pourrait être envisagé en 3D par un encodage de voxels. Pratiquement, Mitra [73] et Pottmann [80] utilisent une méthode d'approximation par des quadriques. Quand à eux, Szeliski et Lavalée [94] utilisent des splines octree pour encoder l'espace entre les surfaces.

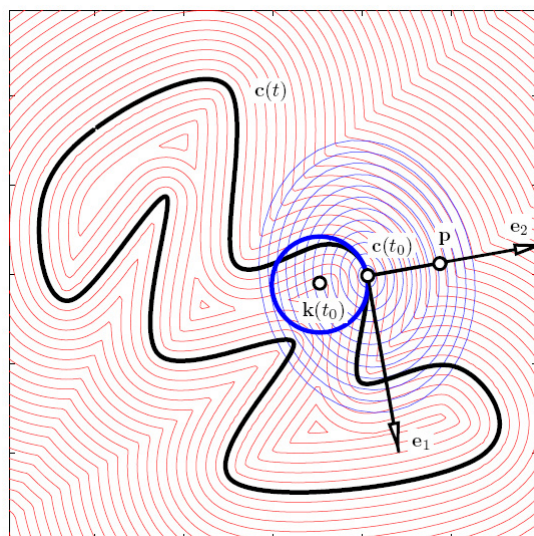


FIG. 3.13: Exemple de courbes de niveaux (en rouge) encodant la distance par rapport au jeu de données ($c(t)$), illustration tirée du travail de Mitra [73]

3.4 Conclusion

Dans ce chapitre, nous avons tout d'abord présenté le recalage point à point. La résolution du problème (formulation analytique) proposé par Horn et ensuite par Arun ont été testées extensivement par l'auteur, qui propose d'ailleurs une version open-source à télécharger (http://www.atracsys.com/_support/openSourceCorner.php). Ces deux algorithmes ne présentent pas de différences significatives mais tout deux sont limités par le fait qu'un appariement exact entre les deux jeux de données est nécessaire. Il est possible de lever partiellement ce handicap en utilisant un algorithme de filtrage de points aberrants de type RANSAC ou *softassign*. Dans les deux cas, ces algorithmes sont inadaptés lorsque plusieurs centaines de points dont de nombreux aberrants sont présents dans les jeux de données.

Nous avons ensuite introduit l'algorithme du plus proche point itéré (ICP) qui, par sa nature, permet de gérer le recalage de jeux de données complexes et bruités. De par nos tests, l'influence de la métrique (linéaire, quadratique, Lorentz, Huber, Mahalobis ou symétrique) dépend de nombreux facteurs comme le bruit dans les jeux de données, leur densité ou même leur transformation initiale. Nous avons testé intensément la plupart de ces algorithmes. Nous en concluons qu'en général, l'emploi du facteur quadratique est un bon compromis entre la qualité de convergence et la temps de calcul. Par contre, le type de mise en correspondance est un facteur beaucoup plus significatif. Parmi les trois présentés et implémentés, la correspondance point à triangle est de loin la plus efficace, permet une approche hiérarchique et compense la non-homogénéité des jeux de données. C'est le type de correspondance qui a été retenu pour la suite de ce travail. La correspondance point à plan est anecdotique et la correspondance point à point s'avère une bonne solution si les jeux de données sont denses (plusieurs milliers voir millions de points) ou si un maillage n'est pas disponible. La méthode de mise en correspondance par les *spin images* n'a pas été testée dans le cadre de ce travail car trop liée au recalage rigide. Une représentation des données par des maillages simplexes ne nous a pas semblé introduire un avantage significatif. Par sa nature, la formulation de l'algorithme ICP est un problème d'optimisation soluble par la méthode initialement présentée par Besl. La limitation de cette formulation est de ne prendre en compte que les transformations rigides. Comme dans la suite de ce travail nous allons aborder des transformations possédant plusieurs dizaines d'inconnues (DDL), une méthode d'optimisation non-linéaire plus générale comme Levenberg-Marquardt (LM) ou les simplexes est plus adaptée. Expérimentalement la différence entre les deux approches est difficilement quantifiable. On notera cependant l'intérêt de la méthode des simplexes lorsque le nombre d'inconnues ou de points dans les jeux de données devient trop élevé. Dans ce cas, le calcul de la matrice Jacobienne (dérivées partielles) de l'algorithme LM introduit un nombre trop élevé d'appariements réduisant ainsi la vitesse de convergence. Cette même vitesse est considérablement améliorée par l'utilisation des *kd-tree* et ceci quel que soit la métrique choisie. La méthode d'encodage de la distance de Hausdorff a été testée dans le cadre d'un autre projet. Nous la considérons comme la méthode la plus efficace d'appariement mais nous ne l'avons pas retenue parce qu'elle n'est pas adaptée à une représentation surfacique des jeux de données.

Bien que le recalage rigide présente la transformation la plus simple entre deux jeux de données, il s'agit d'un processus d'optimisation non-linéaire

complexe dont le résultat dépend énormément de la qualité des données, de leur type et des ressources disponibles pour leur résolution. Pratiquement, la convergence n'est jamais assurée et une phase d'initialisation manuelle grossière de la transformation entre les jeux de données est nécessaire pour éviter les pièges des minima locaux.

Le chapitre suivant présente les modèles de recalage non-rigides les plus utilisés ainsi que leurs méthodes de résolution.

CHAPITRE IV

RECALAGE NON RIGIDE

4.1 Introduction

Au chapitre 3 nous avons présenté une famille de méthodes de recalage rigide qui sont couramment utilisées dans le cadre de chirurgie assistée par ordinateur lorsqu'on doit, par exemple, mettre en correspondance une imagerie préopératoire avec le patient réel. Dans la majorité des cas, comme dans notre problématique, les applications médicales de recalage nécessitent plus de libertés dans les transformations afin de gérer soit les différences anatomiques entre les patients soit prendre en compte la déformabilité des tissus.

Le recalage non-rigide est la mise en correspondance de deux jeux de données qui contiennent des données similaires, mais non semblables, même à une transformation rigide près. L'application d'une déformation non-rigide des données la fait correspondre au modèle.

Il existe deux types de transformations non rigides : les globales qui appliquent une déformation au modèle complet et les locales qui ne transforment qu'une partie de celui-ci. Ce chapitre présente ces deux familles de transformations, du modèle le plus simple au modèle le plus complexe. Les transformations par modèles déformables seront introduites à part au chapitre 6. Cette dernière famille sera comparée avec les différentes approches introduites dans ce chapitre.

4.2 Déformations globales

4.2.1 Déformations linéaires

Les déformations linéaires constituent une extension directe de la transformation rigide. Pratiquement il s'agit d'une transformation matricielle entre les jeux de données. Mis à part les transformations localement affines, les déformations linéaires ont une portée globale. Elles se déclinent en deux grandes sous-familles suivant les contraintes apportées à la transformation.

4.2.1.1 Similitude

La forme la plus simple de déformation linéaire est la similitude. Il s'agit de la composition d'une transformation rigide et d'une homothétie. Les similitudes conservent les rapports de distance et possèdent sept degrés de liberté. La transformation d'un point $\vec{d}_i = (x_i, y_i, z_i)$ par une similitude $T_{similitude}$ de translation \vec{t} , de rotation R et de facteur d'échelle e est :

$$T_{similitude} = ER\vec{d}_i + \vec{t}$$

Où $E = eI$, I étant la matrice identité. La similitude peut être généralisée en imposant un facteur d'échelle différent pour chaque axe. Par exemple, en trois dimensions, le nombre de degrés de liberté (DDL) peut ainsi varier de 6 (transformation rigide ($e = 1$)), 7 (similitude) ou 8, 9 (similitude généralisée - anisotrope). Dans le cas de la similitude généralisée, la matrice E devient $E' = diag(e_1, e_2, \dots)$, où $diag$ est la matrice diagonale avec un facteur d'échelle e_x différent pour chaque axe.

4.2.1.2 Transformation affine

"C'est Euler, en 1748, qui est à l'origine du terme - transformation affine -, car dit-il, - deux courbes images l'une de l'autre par une telle transformation présentent entre elles une certaine affinité. De nos jours, une application affine est une application entre deux espaces qui préserve la structure affine, c'est-à-dire qui envoie les droites, plans, espaces, etc. en des droites, plans, espaces, etc. tout en préservant la notion de parallélisme." (tiré de Wikipedia sous la rubrique "application affine").

La transformation affine (T_{affine}) possède douze DDL et peut s'exprimer comme une matrice de changement de base $A_{3 \times 3}$ et d'une translation \vec{t} . En coordonnées homogènes :

$$T_{affine} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & t_1 \\ a_{21} & a_{22} & a_{23} & t_2 \\ a_{31} & a_{32} & a_{33} & t_3 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_i \\ y_i \\ z_i \\ 1 \end{pmatrix}$$

Une déformation globale affine n'est généralement pas suffisante pour mettre en correspondance les jeux de données de deux personnes différentes [33]. Elle peut par contre être utilisée pour le recalage des données du même sujet prise avec deux modalités différentes, comme le CT et l'IRM [92].

4.2.2 Déformations polynômiales

Les déformations globales polynômiales sont des homéomorphismes.

Cette famille de transformations a été historiquement utilisée pour le recalage non rigide simple de structures cérébrales par optimisation au moyen d'algorithmes génétiques. Actuellement, cette approche est moins en vogue car elle souffre d'un manque de flexibilité par rapport aux méthodes locales qui seront introduites à la section 4.3. Il est néanmoins judicieux de l'utiliser si l'espace de transformation du problème peut être représenté par une transformation trilinéaire ou quadratique.

4.2.2.1 Transformation trilinéaire

Une transformation trilinéaire T_{tri} possède 24 DDL. Elle permet une paramétrisation par la position des huit sommets d'un cube englobant les données. Cette déformation est, par exemple, employée dans les travaux de Jacq [51] ou Rouet [83].

$$T_{tri} = \begin{pmatrix} p_{11} & p_{12} & & p_{18} \\ p_{21} & p_{22} & \dots & p_{28} \\ p_{31} & p_{32} & & p_{38} \end{pmatrix} \begin{pmatrix} 1 & x_i & y_i & z_i & x_i y_i & y_i z_i & z_i x_i & x_i y_i z_i \end{pmatrix}^t$$

4.2.2.2 Transformation quadratique

On obtient de façon similaire les 30 paramètres de la famille des transformations quadratiques (T_{quad}).

$$T_{quad} = \begin{pmatrix} p_{11} & p_{12} & & p_{110} \\ p_{21} & p_{22} & \dots & p_{210} \\ p_{31} & p_{32} & & p_{310} \end{pmatrix} \begin{pmatrix} 1 & x_i & y_i & z_i & x_i y_i & y_i z_i & z_i x_i & x_i y_i z_i & x_i^2 & y_i^2 & z_i^2 \end{pmatrix}^t$$

Une version 2D de cet algorithme est par exemple utilisée par Stewart [91] dans le cadre de la détection de pathologies de la rétine par recalage d'images.

4.3 Déformations locales

Les déformations locales sont dotées d'un nombre élevé de DDL, ce qui leur permet de réaliser un domaine plus important de déformations que leurs homologues globales. Cette sous-famille de transformations est, dans la plupart des cas, définie par des points de contrôle.

4.3.1 Déformation localement affine

Au lieu de calculer une transformation affine globale, Feldmar [34] associe à chaque point de la surface une transformation affine locale (ou patch) telle que la variation de ces transformations sur la surface soit lisse.

4.3.2 Interpolation par des fonctions à base radiale

Les déformations locales peuvent être calculées au moyen d'une fonction polynômiale définie par morceaux, ou plus spécifiquement par les splines plaques-minces (*thin-plate splines - TPS*) de Harder et Desmarais [45]. Celles-ci ont été créées initialement pour la CFAO. Il s'agissait de calculer la déformation d'une plaque de métal mince comme celles qui recouvrent les ailes d'avions lorsqu'on contraint un certain nombre de points (x, y) à des déplacements z . Ces splines donnent le déplacement physique, de n'importe

quel point du plan 2D, qu'une plaque de métal mince de dimension infinie pourrait adopter si elle était contrainte à passer par les positions (x, y, z) spécifiées. Les TPS ont été formalisées et étendues par Duchon ^[31]. Elles sont une généralisation des splines cubiques et font partie d'un sous-ensemble de la famille des splines généralisées et des fonctions à base radiale (*Radial Basis Functions - RBF*).

L'interpolation par des RBF nécessite l'identification et la mise en correspondance de paires de points (appelés aussi points de contrôle) dans les jeux de données respectifs. Cette phase peut être réalisée manuellement ^[14], par exemple, par la saisie de points anatomiques facilement repérables dans des images radiologiques ou automatiquement au moyen d'un algorithme de recherche de points saillants (*salient points*), de corrélation croisée (*cross correlation*), de courbures extraites d'images 3D ^[96], etc. Une fois ces points de contrôle identifiés et appariés, ils sont utilisés pour déterminer la transformation de coordonnées (*mapping*) de n'importe quel point des données en un point correspondant du modèle. La particularité des RBF et de la plupart des autres splines est de réaliser une transformation exacte des points de contrôle. Pour les autres points interpolés, les RBF créent un champ de déformation régulier homéomorphique autour des points de contrôle grâce à une contrainte de régularisation. Typiquement, celle-ci est choisie dans l'espace de distribution de Beppo-Levi de \mathfrak{R}^3 avec des dérivées secondes doublement intégrables ^[14] ^[17].

Soit une série de p points de contrôle des données \vec{d}_i et leur points correspondants sur le modèle \vec{m}_i . La transformation par interpolation de fonctions à base radiale T_{RBF} de $\mathfrak{R}^n \rightarrow \mathfrak{R}^n$ est tel que :

$$T_{RBF}(\vec{d}_i) = \vec{m}_i, \forall i = 1..p \quad (4.1)$$

A noter que l'équation précédente peut être décomposée afin d'être résolue séparément pour chaque coordonnée. Dès lors, l'équation 4.1 devient :

$$\tau_k(\vec{d}_i) = m_{i,k}, \forall i = 1..p, \forall k = 1..n$$

où $\tau_k : \mathfrak{R}^n \rightarrow \mathfrak{R}$ est la décomposition de T_{RBF} relativement à la k^{me} coordonnée.

La fonction d'interpolation RBF est composée de deux parties :

$$\tau_k(\vec{x}) = \phi(\vec{x}) + R(\vec{x})$$

où $\phi(\vec{x})$ est un polynôme de type $\beta_0 + [\beta_1 \dots \beta_4] * \vec{x}$ et R est la somme des fonctions à base radiale :

$$R(\vec{x}) = \sum_{i=1}^p \alpha_i \rho(\|\vec{x} - \vec{d}_i\|)$$

où $\|\vec{x} - \vec{d}_i\| = r$ est la distance Euclidienne entre \vec{x} et le point de contrôle \vec{d}_i et $\rho(r)$ une fonction définie dans l'intervalle $[0, \infty[$. Cette fonction définit la sous-famille de la RBF. Le tableau 4.1 présente les plus utilisées.

| | |
|--------------------------|--|
| Splines plaques-minces | $\begin{cases} r^{4-w} \ln(r) & 4 - w \in 2\mathbb{N} \\ r^{4-w} & \text{autrement} \end{cases}$ |
| Multiquadriques | $(r^2 + c^2)^\mu, \mu \in \mathbb{R}_+$ |
| Multiquadriques inverses | $(r^2 + c^2)^{-\mu}, \mu \in \mathbb{R}_+$ |
| Gaussiennes | $e^{-\frac{r^2}{2\sigma^2}}$ |

TAB. 4.1: Exemple de sous-famille des RBF

Les deux premières fonctions augmentent (voir la figure 4.1) tandis que les deux dernières diminuent (voir la figure 4.2) avec l'augmentation de la variable r . De par sa nature, cette première famille a donc une influence globale sur la déformation tandis que la seconde une influence locale. Idéalement, les fonctions RBF n'ont pas de support compact ($f(r) \neq 0, \forall r$), pour éviter des effets indésirables lors des déformations.

Dans la suite de ce travail, nous nous intéresserons spécifiquement aux splines plaques-minces avec $w = 2$ pour les déformations globales et aux fonctions gaussiennes avec différentes valeurs de σ pour les déformations locales. Ce dernier paramètre permet le réglage de la localité des points de contrôle. Pour des références bibliographiques plus complètes concernant ces différentes fonctions, se référer aux travaux de Fornet et al. [38] ou Arad et Reisfeld [5].

Les paramètres α_i et β_i sont déterminés par la résolution du système linéaire suivant :

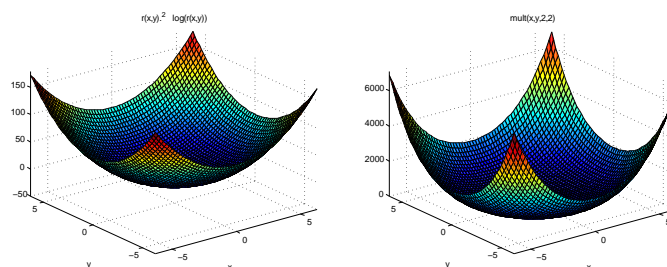


FIG. 4.1: Famille des fonctions à base radiale avec influence globale. A gauche fonction pour les splines plaques-minces avec $w = 2$, à droite fonction pour les multiquadriques avec $c = 2$ et $\mu = 2$. Bien que ces deux familles de fonctions présentent des courbes de même forme, on remarque un changement d'échelle considérable sur l'axe z . Au niveau de l'applicatif, les splines-plaques minces sont utilisées pour des recalages "doux" et ont la particularité que leur conditionnement est invariant d'une modification du facteur d'échelle. Les multiquadriques sont, quant à elles, utilisées dans des cas très particuliers comme le recalage de données topographiques.

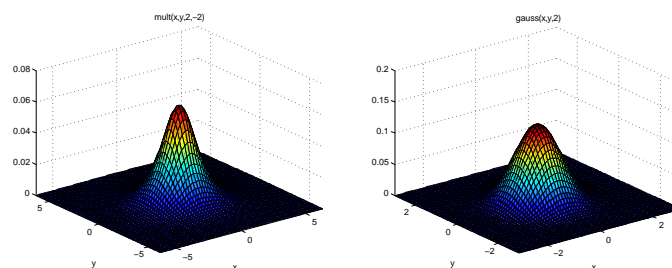


FIG. 4.2: Famille des fonctions à base radiale avec influence locale. A gauche fonction pour les multiquadriques inverses avec $c = 2$ et $\mu = 2$, à droite fonctions gaussiennes avec $\sigma = 2$. Ces deux derniers types de transformation nécessitent un facteur d'échelle (c ou σ) qui varie en fonction de la taille des jeux de données. Le comportement de ces fonctions sont très proches. Elles sont par exemple indifféremment utilisées dans les applications de réseaux de neurones.

$$\begin{pmatrix} K & P \\ P^t & 0 \end{pmatrix} \begin{pmatrix} \vec{\alpha} \\ \vec{\beta} \end{pmatrix} = \begin{pmatrix} \vec{m}_k \\ \vec{0} \end{pmatrix} \quad (4.2)$$

où K est la sous-matrice $n \times n$ donnée par $K_{i,j} = \rho(\|\vec{d}_i - \vec{d}_j\|)$ et P , la matrice $n \times 4$ définie par $P_{i,j} = \phi(\vec{d}_j)$, \vec{m}_k est le vecteur comprenant la k^{me} coordonnées des vecteurs \vec{m}_i .

L'intérêt des RBF a très longtemps été restreint à des problèmes de petite complexité possédant au maximum quelques centaines de points de contrôle. En effet, la résolution du système 4.2 par des méthodes standards d'analyse numérique, nécessite l'inversion d'une matrice $(n + 4) \times (n + 4)$ qui a une complexité de stockage de l'ordre de $O(N^2)$ et une complexité de résolution arithmétique de l'ordre de $O(N^3)$ [17]. Greengard et Rokhlin [43] ont levé cette limitation grâce à la méthode d'accélération par des pôles multiples (*Fast Multipole Method - FMM*). L'originalité de leur approche est de poser l'hypothèse que lorsque le calcul de la RBF est réalisé, une précision infinie n'est ni nécessaire ni souhaitable et qu'une approximation locale et hiérarchique peut être utilisée à la place. Dès lors, l'ordre de complexité de stockage devient $O(MN)$ et celle de calcul de $O(M + N \log(N))$, où M est le nombre de points locaux à prendre en considération. Plus ce nombre est important, plus la RBF est fidèle. L'exemple de la figure 4.3, présente une reconstruction surfacique réalisées par Carr et al. [17] utilisant cette technique d'optimisation. Bien évidemment, la même technique peut être utilisée dans le cas de recalage non rigide.

4.3.3 Déformation de forme libre

Les déformations de forme libre (*Free Form Deformations - FFD*) ont été introduites en 1986 par Sederberg et Parry [87]. Elles constituent une technique pour transformer librement des modèles géométriques. Il est ainsi possible de déformer des primitives surfaciques de n'importe quel type ou degré, par exemple des figures planes, des quadriques, des surfaces paramétriques ou des surfaces implicites. Par analogie, les FFD peuvent être comparées à un parallélépipède en gélatine transparente dans lequel sont incrustés un ou plusieurs objets à déformer. On les imagine flexibles afin qu'ils puissent se déformer en même temps que la gélatine qui les entoure. La figure 4.4 présente l'exemple original décrit par Sederberg avant et après déformation par les FFD.

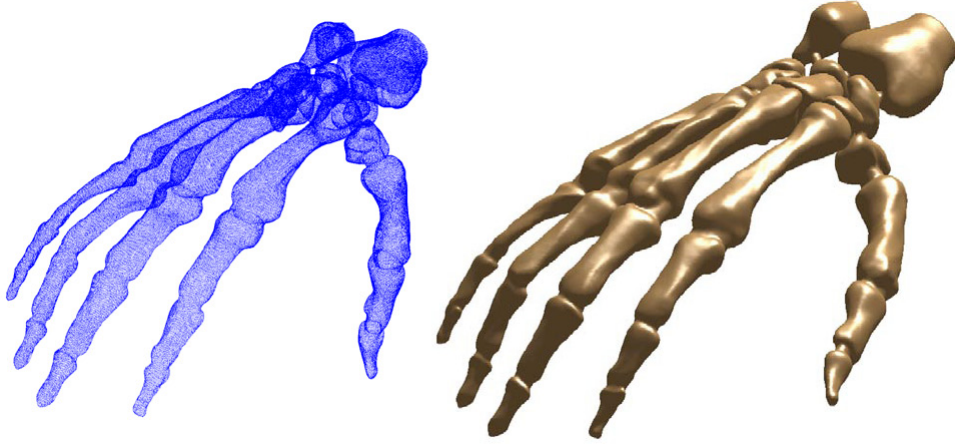


FIG. 4.3: Exemple tiré du travail de Carr et al. ^[17] présentant une possibilité d'accélération du calcul des RBF grâce à la méthode d'optimisation par pôles multiples (FMM). Dans cet exemple, les RBF sont utilisées pour la reconstructions d'une surface à partir d'un nuage de plusieurs dizaines de milliers de points.

Signalons qu'actuellement le terme de FFD s'est généralisé et regroupe, par extension, tous les types de déformation basés sur des points de contrôle, ce qui peut impliquer des confusions sur les méthodes employées.

Mathématiquement, si on impose un référentiel local $(\vec{S}, \vec{T}, \vec{U})$ sur un parallélépipède avec l'origine \vec{X}_0 sur une des arêtes, un point \vec{X} possède les coordonnées (s, t, u) dans ce système définies comme :

$$\vec{X} = \vec{X}_0 + s\vec{S} + t\vec{T} + u\vec{U}$$

Si on impose aussi une grille de points de contrôle $P_{ijk}, \forall i = 0..l, \forall j = 0..m, \forall k = 0..n$ sur le parallélépipède avec les points de contrôle positionnés tel que :

$$\vec{X} = \vec{X}_0 + \frac{i}{l}\vec{S} + \frac{j}{m}\vec{T} + \frac{k}{n}\vec{U}$$

la nouvelle position \vec{X}' (après déformation) d'un point \vec{X} se calcule ainsi :

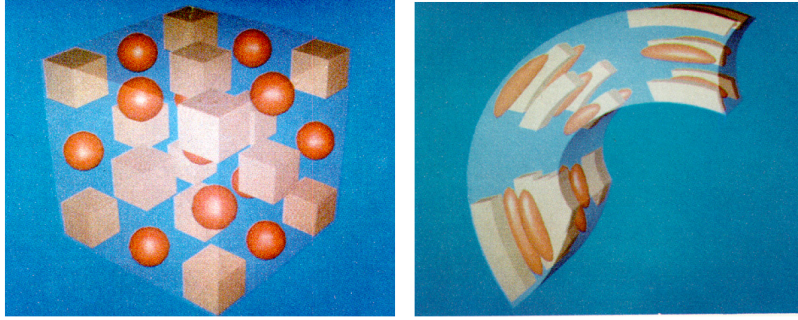


FIG. 4.4: Exemple de déformation de forme libre (FFD) tiré de l'article de Sederber ^[87]. Objets avant (gauche) et après (droite) la déformation. Ce type de transformation est uniquement caractérisé par le déplacement de points de contrôle. La position des autres points est interpolée au moyen de polynômes de Bernstein. Par analogie, les FFD peuvent être comparées à un parallélépipède en gélatine transparente dans lequel sont incrustés un ou plusieurs objets à déformer. On les imagine flexibles afin qu'ils puissent se déformer en même temps que la gélatine qui les entoure.

$$\vec{X}' = \sum_{i=0}^l \sum_{j=0}^m \sum_{k=0}^n P_{ijk} B_i^l(s) B_j^m(t) B_k^n(u)$$

où $B_i^l(s)$, $B_j^m(t)$ et $B_k^n(u)$ sont les polynômes de Bernstein définis comme :

$$B_i^l(s) = \frac{l!}{i!(l-i)!} (1-s)^{l-i} s^i$$

$$B_j^m(t) = \frac{m!}{j!(m-j)!} (1-t)^{m-j} t^j$$

$$B_k^n(u) = \frac{n!}{k!(n-k)!} (1-u)^{n-k} u^k$$

Les FFD de base ont été étendus à plusieurs reprises. Dans le domaine médical, les polynômes de Bernstein sont souvent remplacés par des B-Splines. Une autre extension couramment utilisée (en modélisation) est l'insertion d'un facteur de pondération des points de contrôle rajoutant ainsi un degré de liberté supplémentaire lors de la manipulation des déformations.

Szeliski et Lavallée ^[94] sont les pionniers pour l'emploi des déformations

de forme libre dans le domaine médical. Bardinet et al. [8] les ont ensuite servies pour le suivi et la segmentation du ventricule droite. Rückert et al. [84] utilisent quant à eux les FFD pour le recalage monomodal de mammographies du même patient par imageur IRM. Ces derniers proposent une approche hiérarchique 4.4, où le déplacement global de la poitrine est modélisé par une transformation affine (voir section 4.2.1.2), tandis que les déformations locales sont gérées par des FFD. A noter que dans leur implémentation les polynômes de Bernstein sont remplacés par des fonctions B-Splines et une fonction de régulation est utilisée pour lisser la déformation. Les FFD sont un cas spécifique de fonctions à base radiale - RBF (voir chapitre 4.3.2), où les points de contrôle sont, de préférence, uniformément répartis. Le tableau 4.2 présente les résultats de Rückert et al. Pour comparer les séries d'images IRM, le critère utilisé est la somme des différences d'intensités (*Sum of Intensity Differences - SSD*).

| Type de recalage | SSD (mean) | SSD (variance) |
|------------------------|------------|----------------|
| Pas de recalage | 38,52 | 53,90 |
| T_{rigide} | 23,63 | 33,38 |
| T_{affine} | 21,38 | 29,84 |
| $T_{affine} + T_{FFD}$ | 13,28 | 20,91 |

TAB. 4.2: Approche hiérarchique de Rückert et al. [84] utilisant des transformations affines combinées à des FFD. Les résultats présentent l'erreur de recalage moyen en mm. Notons que leur approche par FFD utilise environ $10 \times 10 \times 10 = 1000$ points de contrôle sur une grille, où les points sont espacés de 15mm. Leurs résultats montrent que, pour le recalage de mammographies, les FFD améliorent la méthode affine d'environ 30%.

Il serait long et fastidieux de lister tous les travaux basés sur les FFD. Signalons que la majorité des auteurs utilisent une approche hiérarchique [15], où ils combinent une transformation linéaire (voir le chapitre 3 et la section 4.2.1) pour s'approcher suffisamment de la solution finale avec les déformations de forme libre.

4.4 Approche hiérarchique

Appliquées directement sur les jeux de données, les méthodes de recalage local donnent d'assez bons résultats mais présentent souvent un état de surface comme si on l'avait emboutie (voir par exemple l'effet "balle de golf" de la figure 4.12). Il est possible d'améliorer ce rendu ainsi que la robustesse du recalage vis-à-vis du bruit des données en utilisant préalablement

des classes de transformations possédant moins de DDL. Montagnat ^[74] propose, dans son travail de thèse, une approche hiérarchique pour réaliser un recalage précis des données tout en conservant la qualité géométrique de la surface. Ce processus commence par l'application d'une transformation globale jusqu'à atteindre une position d'équilibre. Ensuite, le nombre de degrés de libertés (DDL) est progressivement accru pour laisser plus de variabilité à la surface.

Il reste à déterminer quand le changement de classe de transformation est possible. La figure 4.5 présente l'erreur RMS à chaque étape de l'algorithme du plus proche point itéré (ICP) en utilisant différents types de transformation. De manière générale, cette courbe a un profil de forme $\frac{1}{x} + c$ et, au bout de quelques itérations, le modèle ne se déforme plus de manière significative. Montagnat ^[74] propose d'augmenter automatiquement le nombre de DDL lorsque "la quantité de déplacement du modèle devient inférieure à 0,1% de la quantité de déplacement initial".

4.5 Approche progressive

Aux vues des considérations de la section précédente, il serait intéressant de concevoir une technique de transformation permettant d'augmenter progressivement et automatiquement le nombre de DDL. Le recalage pourrait ainsi être affiné jusqu'à la précision escomptée.

Dans le domaine du recalage médical, la presque totalité des approches qui utilisent les déformations de forme libre (voir section 4.3.3) utilisent des fonctions de base de type B-spline. Cette famille de fonctions nécessite la construction d'un maillage régulier autour de l'image ou de la surface à déformer (^[94], ^[8]) de telle sorte qu'un nombre important, typiquement plusieurs milliers, de DDL sont nécessaires pour obtenir une déformation correcte. Des travaux plus récents comme ceux de Rückert et al. ^[84], Xie et al. ^[100] ou Houhou et al. ^[48] utilisent une approche hiérarchique de maillage comme celui présenté à la figure 4.8.

Similairement au principe des *octrees*, Xie et al. proposent une technique hiérarchique localisée. Ainsi, une maille plus compacte est générée dans les régions du maillage présentant une grande erreur de recalage. Ce processus est réalisé itérativement pour tous les niveaux hiérarchiques.

Quelles que soient les astuces d'optimisation, l'utilisation d'une maille de

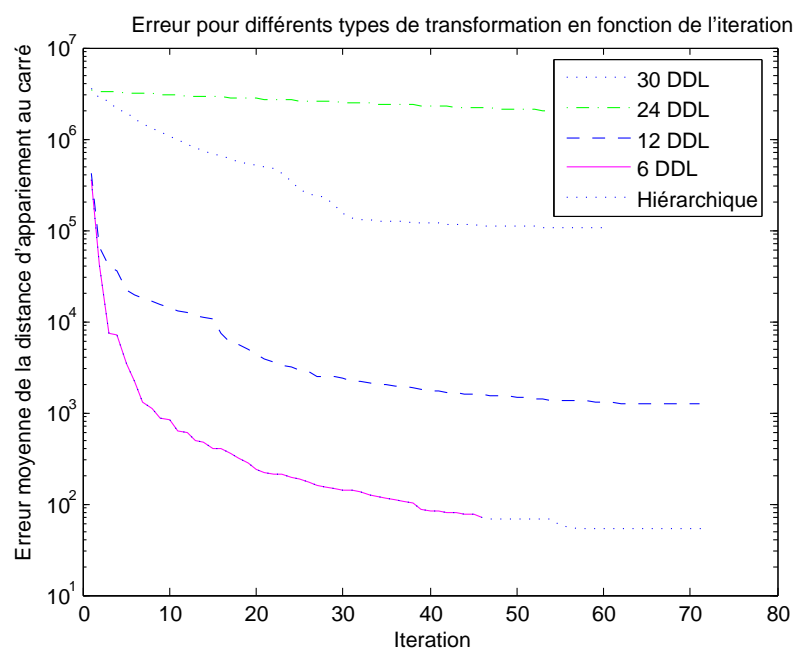


FIG. 4.5: Exemple de recalage par la méthode LM-ICP (voir section 3.3.5) avec la méthode d'appariement point à triangles (voir section 3.3.4.3) en utilisant différents types de transformations (rigide - 6 DDL, affine - 12 DDL, trilinéaire - 24 DDL, quadratique - 30 DDL) ainsi que l'approche hiérarchique. L'axe des erreurs présente la moyenne de la distance au carré de l'appariement. A noter que l'initialisation de la transformation avec trop de degrés de liberté (typiquement 24 ou 30 - voir figure 4.7) conduit à une divergence de l'algorithme. Dans cet exemple, force est de constater que l'approche hiérarchique n'apporte pas d'amélioration substantielle par rapport à une transformation rigide simple. Cet exemple est illustré en figure 4.6.

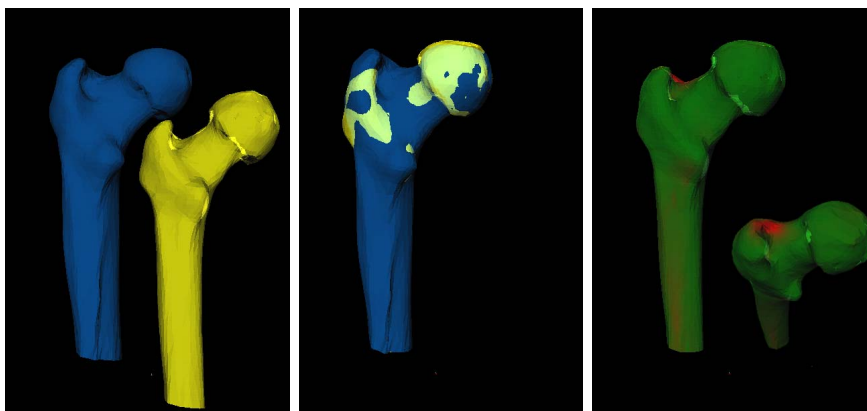


FIG. 4.6: Résultat du recalage rigide de l'exemple de la figure 4.5 entre deux fémurs provenant de deux patients différents. A gauche, représentation des deux os dans leur position initiale après segmentation et reconstruction 3D, au milieu, recalage avec la méthode LM-ICP après environ 50 itérations, à droite, deux vues de l'erreur d'appariement projetée sur la surface d'un des deux fémur. L'erreur est encodée du vert (peu d'erreur) au rouge (erreur importante). Dans ce cas précis, une erreur importante est localisée au niveau de la fossette digitale (voir annexe 3) du fait de la grande variabilité de l'anatomie des sujets à cet endroit.

points de contrôle est fondamentalement lente. Pour preuve, les différents auteurs cités présentent des temps de calcul variants de dizaines de minutes à plusieurs heures pour des déformations de petits volumes (typiquement $256 \times 256 \times 128$).

Au lieu d'utiliser une maille possédant des points de contrôle espacés uniformément dans l'espace, la complexité de la déformation peut être considérablement réduite, pour peu que l'on puisse précisément localiser les régions de forte variation entre les données et le modèle. L'utilisation d'une transformation par RBF, où l'on rajoute progressivement et judicieusement des points de contrôle, permet une convergence rapide de la déformation avec un nombre limité de DDL.

4.6 Contribution

Afin de palier aux limitations des RBF basés sur des mailles, je propose l'algorithme suivant appelé "recalage non rigide progressif" car son nombre de DDL augmente graduellement. A chaque nouvelle itération, les données

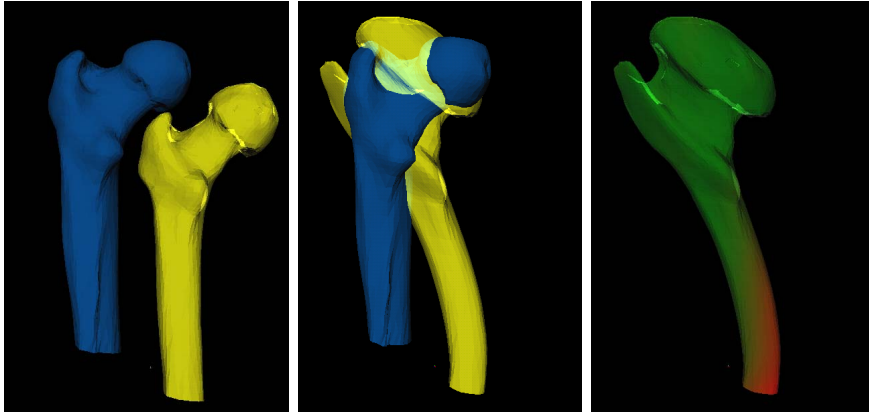


FIG. 4.7: Résultat du recalage non-rigide (30 DDL) de l'exemple de la figure 4.5 entre deux fémurs provenant de deux patients différents. A gauche, représentation des deux fémurs dans leur position initiale après segmentation et reconstruction 3D, au milieu, recalage avec la méthode LM-ICP après environ 50 itérations, à droite, erreur d'appariement projetée sur la surface d'un des deux fémur. L'erreur est encodée du vert (peu d'erreur) au rouge (erreur importante). Ce cas illustre un exemple de divergence lorsque la surface présente prématurément trop de degrés de variabilité. Clairement, l'optimisation a privilégié une transformation conduisant rapidement à un appariement de la tête fémorale. Elle n'a ensuite pas réussi à faire converger son corps. Une approche hiérarchique aurait typiquement évité ce phénomène.

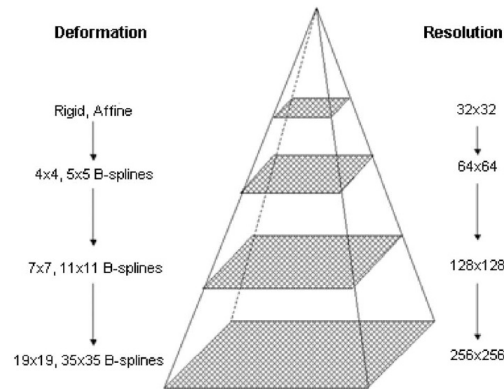


FIG. 4.8: Exemple de recalage tiré des travaux de Xie et al. ^[100] présentant une approche hiérarchique. Une transformation rigide est suivie d'une transformations FFD (indiqué B-Spline) avec un maillage de taille 32 x 32. Après convergence, on double la résolution et le processus est recommencé jusqu'à obtenir une maille de points de contrôle de taille 256 x 256. Les avantages d'une approche hiérarchique sont de permettre une convergence plus sûre et surtout plus rapide.

sont déformées par rapport aux points de contrôle de l'itération précédente. Lors de l'étape d'appariement, de nouveaux points de contrôle sont sélectionnés dans les régions de fortes variations parmi les paires de points présentant les distances les plus importantes. De manière abstraite, on peut voir cette approche comme une maille hiérarchique dont on rajoute un point de contrôle à chaque changement de résolution. Cette approche a les avantages de réaliser des déformations avec un nombre restreint de points de contrôle et permet de stopper le processus d'ajout de points lorsqu'une précision escomptée est atteinte.

Prenons par exemple le résultat du recalage rigide de la figure 4.6 comme point de départ de l'algorithme "progressif". On constate très clairement que la région possédant l'erreur la plus importante se situe dans la zone rouge, au niveau de la fossette digitale. En effet, cette caractéristique anatomique varie fortement d'un patient à l'autre. Donc, la première paire de points appariés sélectionnée par l'algorithme se trouve dans ce périmètre et définit le premier point de contrôle pour la transformation RBF ($T_{RBF}(0)$). Les points des données sont ensuite déformés relativement à la transformation courante et une seconde itération de l'algorithme est entreprise. Ce processus est répété jusqu'à atteindre la précision voulue.

Intuitivement, l'algorithme est optimal lorsque toutes les paires appariées sont sélectionnées comme points de contrôle. En pratique, on se heurte au fait que ces points génèrent un champ de déformation régulier homéomorphe non ponctuel. Aussi le type de fonction utilisé et son influence géographique (locale ou globale) modifie radicalement la convergence de l'algorithme. On serait tenté de penser qu'une fonction de base comme une gaussienne ou une multiquadrique inverse serait la famille de fonctions la plus adaptée car le rajout d'un point de contrôle n'influence que localement la déformation. En pratique, on remarque que les splines plaques minces (TPS) sont plus adéquates car elles ont un comportement global, c'est à dire qu'elles possèdent une propriété d'invariance au facteur d'échelle. Les fonctions locales doivent être utilisées avec un facteur d'échelle adéquat, sinon, elles présentent différents types d'instabilités. Par exemple, la figure 4.9 met en évidence l'instabilité si le facteur d'échelle est trop grand. L'utilisation d'une fonction de base radiale Gaussienne avec $\sigma = 0,5$ (courbe "RBF-Gauss Progressif (0.5)") a un rayon d'influence trop important. Dans un premier temps, le rajout de points de contrôle conduit à une réduction rapide de l'erreur (dans l'exemple précité jusqu'à environ 200 DDL). Ensuite, le rayon d'influence (trop élevé), surcompense et contribue à augmenter l'erreur moyenne d'appariement. Un

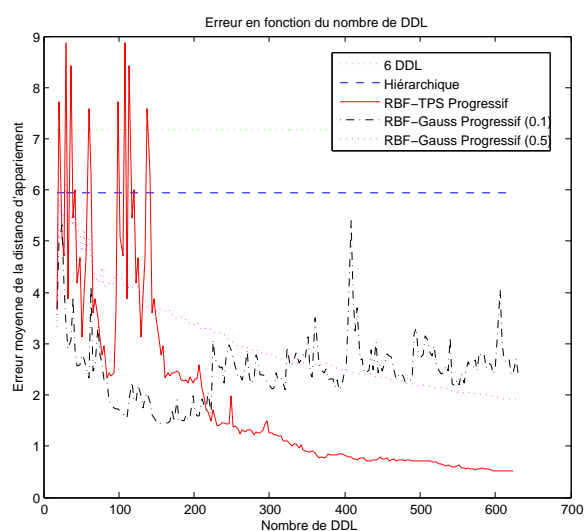


FIG. 4.9: Exemple de recalage non rigide progressif après recalage rigide de l'exemple de la figure 4.6. Les deux droites verticales présentent l'erreur d'appariement pour les cas rigides (haut) et hiérarchique (bas) de la figure 4.5. Les autres courbes sont trois exemples de recalage progressif RBF avec différentes fonctions de base radiales. Avec peu de points, les techniques de RBF progressives montrent une forte tendance à l'instabilité. On constate que les techniques locales ne présentent pas de convergence spectaculaire, sont très sensible au paramètre d'échelle et dans le meilleur des cas, convergent vers une surface présentant des défauts de type "boule de golf" (voir la figure 4.12). La fonction globale de TPS présente, quant à elle, les meilleurs résultats (voir la figure 4.10) vu que cette famille de fonction est invariante au facteur d'échelle.

facteur d'échelle trop petit $\sigma = 0,1$ (courbe "RBF-Gauss Progressif (0.1)") permet une décroissance quasi linéaire de l'erreur et tend vers zéro lorsque le nombre de points de contrôle atteint le nombre de paires d'appariement. Ce type de déformation génère cependant des effets indésirables. On remarque par exemple sur la figure 4.12 l'aspect "boule de golf" de la tête fémorale causée par l'influence trop locale de la transformation : le point de contrôle se retrouve l'unique point déformé. Afin de tirer parti des fonctions gaussiennes, la réduction du paramètre σ au cours des itérations est envisageable pour rendre la déformation progressivement de plus en plus locale. Les différents tests effectués par l'auteur ont montré des résultats non concluants, d'autant plus que le paramètre σ doit spécifiquement être *tuné* pour chaque recalage. Contrairement à l'intuition, les fonctions TPS (courbe "TPS Progressif" ainsi que les exemples des figures 4.10 et 4.11) ont un comportement bien plus intéressant que les fonctions gaussiennes (exemple de la figure 4.12). De par leur portée globale, ces fonctions ont une convergence chaotique à faible DDL car l'ajout d'un point de contrôle perturbe fortement les points non contraints situés à grande distance. Dans un second temps, lorsque les points de contrôle sont suffisamment denses (environ 50 - 150 DDL - dans l'exemple précité), les TPS convergent régulièrement car l'influence globale de déformation diminue avec l'augmentation des points de contrôle.

4.6.0.1 Initialisation

L'algorithme progressif présenté emploie les fonctions à base radiales de la section 4.3.2. Il nécessite l'emploi d'un minimum de 6 points de contrôle initiaux. Ces points ne peuvent pas être détectés par la méthode itérative car aucune déformation ne s'effectue après la sélection d'un seul point de contrôle. Aussi, l'algorithme va au cours des itérations toujours choisir ce même appariement et le recalage ne pourra jamais démarrer. Une méthode d'initialisation peut être le choix arbitraire du minimum de points de contrôle requis comme une sélection à la "Monte Carlo" (c'est-à-dire en choisissant 6 points aléatoirement sur la surface). Cette approche ne prévient cependant pas d'une déformation initiale aberrante. Une des constatations de la section précédente est l'optimalité de l'algorithme progressif TPS lorsque les points de contrôle sont répartis de manière uniforme sur les surfaces des jeux de données.

Dans mon implémentation, nous utilisons la sphère circonscrite aux données, sur laquelle nous choisissons un sextuplet de points cardinaux (haut,

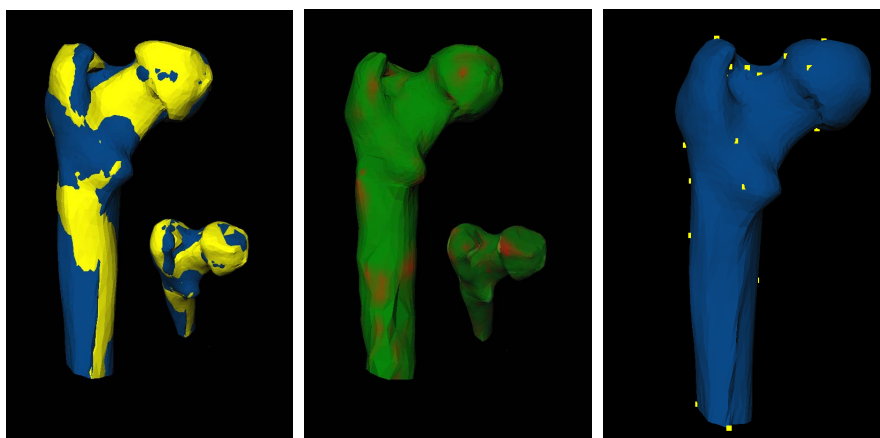


FIG. 4.10: Résultat du recalage non rigide progressif par fonctions à base radiale TPS vers 25 itérations (75 DDL). A gauche, superposition des deux jeux de données, au milieu, encodage de l'erreur d'appariement du vert (proche) au rouge (éloigné), à droite, emplacement des points de contrôle. On constate que l'erreur moyenne d'appariement est mieux répartie sur la surface que les méthodes décrites précédemment, bien qu'elle soit tout de même comprise entre 2,5 et 9 mm, c'est-à-dire potentiellement plus élevée qu'un recalage rigide (7,2 mm).

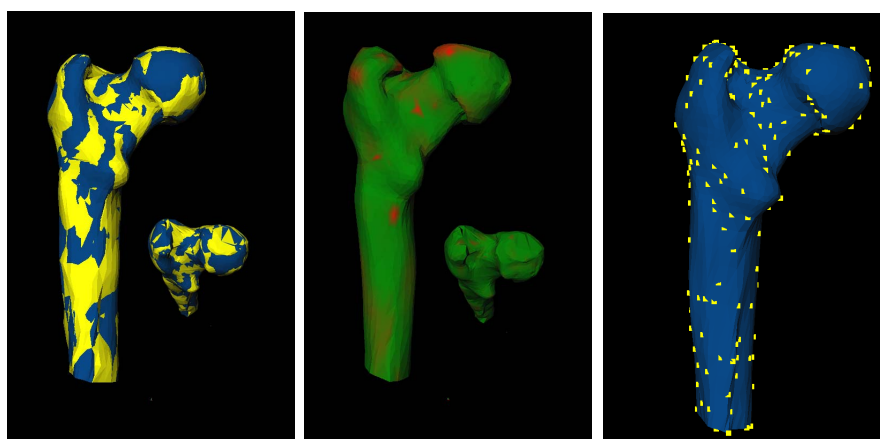


FIG. 4.11: Résultat du recalage non rigide progressif par fonctions à base radiale TPS après 200 itérations. A gauche, superposition des deux jeux de données, au milieu, encodage de l'erreur d'appariement du vert (proche) au rouge (éloigné), à droite, emplacement des points de contrôle. Notons qu'à cette itération, l'erreur d'appariement moyenne est inférieure à 1 mm, c'est-à-dire six fois moins élevée qu'un recalage rigide (6 mm). Les points de contrôle sont répartis de manière presque uniforme sur la surface.

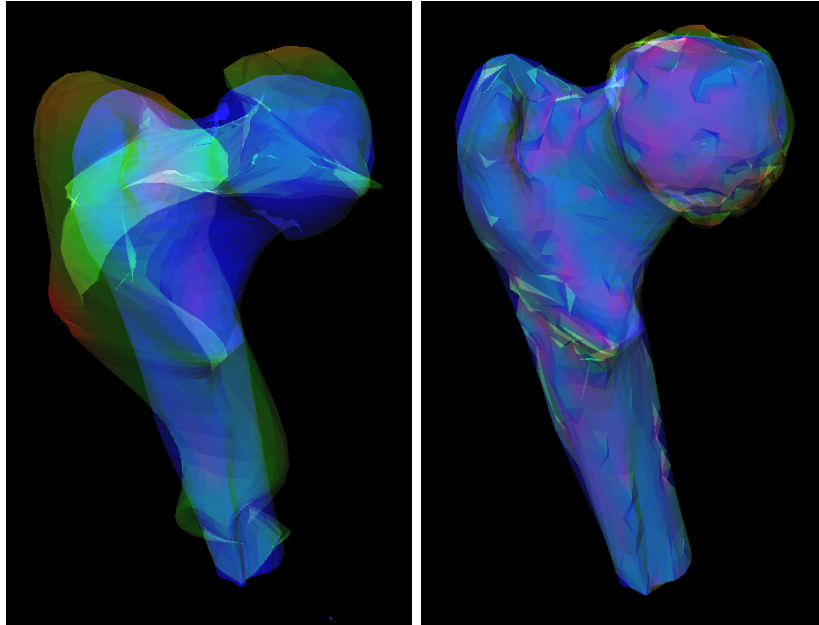


FIG. 4.12: Résultat du recalage non rigide progressif par fonctions à base radiale gaussienne après 25 (gauche) et 200 (droite) itérations. Les fonctions locales doivent être utilisées avec un facteur d'échelle adéquat, sinon, elles peuvent présenter différents types d'instabilités. Cet exemple illustre un facteur d'échelle trop petit permettant, en théorie, une décroissance quasi linéaire de l'erreur. On constate cependant un résultat erratique lorsque le nombre de points de contrôle est insuffisant (gauche). Le recalage converge lorsque le nombre de points de contrôle atteint le nombre de paires d'appariement. Pratiquement, on remarque un défaut de type "boule de golf" (droite), créant des irrégularités sur la surface. Ce phénomène est causé par une influence trop locale du paramètre σ de la fonction.

bas, gauche, droite, devant et derrière) qui sont appariés sur les données. Cette approche permet une distribution quasi uniforme des points de contrôle initiaux sur une surface convexe ou pseudo-convexe, comme par exemple le fémur.

4.7 Variations autour du recalage progressif

Parmi toutes les techniques de recalage non rigides présentées jusqu'à présent, les déformations de fonctions à base radiale TPS présentent les meilleurs résultats. De plus, en combinant ces fonctions à l'algorithme progressif, la convergence est rapide et permet de choisir un seuil d'arrêt qualitatif. L'approche progressive permet également une réduction importante du nombre de DDL par rapport aux FFD classiques et, avec un peu d'imagination, ouvre la porte à des débouchés dans de nombreux domaines connexes comme la compression de maillages.

4.8 Conclusion

Dans ce chapitre, nous avons tout d'abord vu certaines extensions globales du recalage rigide. La complexité de la transformation est passée de 6 DDL à 7, 8 et 9 DDL avec l'adjonction du facteur de similitude puis 12 DDL pour les transformations affines. Plus exotiques, les déformations polynômiales avec 24 DDL ou 30 DDL ont été brièvement introduites. Toutes ces méthodes ont été testées et validées sur nos jeux de données. Les résultats sont présentés à la section 4.4. Les meilleures convergences sont obtenues en effectuant une approche hiérarchique, c'est-à-dire qu'une approche rigide est suivie d'une approche non rigide globale, dans laquelle les DDL sont progressivement ajoutés après convergence de l'étape précédente. Au final une déformation de 30 DDL n'apporte pas, qualitativement, de résultats fondamentalement meilleurs.

La seconde partie de ce chapitre est consacrée à présenter deux sous-familles de déformations locales : les FFD et les RBF. Dans leur version classique, ces deux approches nécessitent un nombre important de points de contrôles (typiquement plusieurs milliers) uniformément répartis dans l'espace. Ces méthodes nécessitent un calcul de plusieurs minutes (voir heures) et gèrent des déformations de milliers de DDL. Bien que présentant qualitativement d'excellents résultats et s'appliquant à n'importe quels types de

déformations, elles ne sont pas adaptées à une approche interactive du recalage. Pour pallier à ces différents problèmes, nous proposons une approche originale dite "progressive" possédant les qualités précitées, dans laquelle les DDL sont, au fil des itérations, rajoutés, ce qui a pour effet de rendre le recalage de plus en plus précis jusqu'à la qualité escomptée. Pratiquement, l'emploi de 400 DDL avec une fonction TPS dans la problématique fémorale permet d'obtenir un recalage six fois plus précis que le recalage rigide. Malheureusement, même le recalage progressif nécessite quelques minutes de calcul sur un ordinateur moderne avant le début de la convergence, ce qui rend cette approche inutilisable dans un contexte temps-réel.

Nous verrons dans les chapitres suivants que l'introduction de contraintes supplémentaires, comme par exemple une connaissance de la topologie de la forme, permet une amélioration considérable quantitativement et qualitativement du recalage. L'emploi de modèles déformables permet même, si des points de contrôle sont soigneusement localisés dans les images scanners, de supprimer partiellement ou intégralement la composante itérative du recalage. Le chapitre suivant présente une méthode de localisation automatique de ces points dans des coupes CT fémorales.

CHAPITRE V

LOCALISATION AUTOMATIQUE DE PARTIES ANATOMIQUES DU FÉMUR

5.1 Introduction

La recherche dans le domaine de la segmentation automatique d'images médicales se focalise, dans la plupart des cas, dans la délimitation fine des contours d'organes. La plupart des techniques de segmentation classiques comme les *snakes* ou celle présentée au chapitre 7 nécessitent une bonne approximation du contour à segmenter durant l'initialisation pour garantir une délimitation précise du pourtour de l'organe. Aussi, la difficulté de la tâche de ces algorithmes requiert, dans la majorité des cas, une interaction manuelle de l'utilisateur et, dans tous les cas, une validation du résultat. L'emploi de systèmes automatisés ne se justifie que si le temps (respectivement le coût) du processus automatique n'excède pas le temps (respectivement le coût) pour la même procédure réalisée manuellement.

Plutôt que de délimiter directement les contours d'organes par des algorithmes de segmentation aléatoires, notre approche consiste à trouver grossièrement et rapidement certains paramètres anatomiques facilement localisables dans les images et d'utiliser ensuite ces informations pour recalibrer un modèle déformable (voir le chapitre 6) avant d'effectuer une segmentation fine (voir le chapitre 7).

Dans ce chapitre, nous présentons la localisation automatique de parties anatomiques du fémur au moyen de rayons de recherche (*Search Rays*) initialement introduite par Christian Lorenz ^[64] en 1997. Nous présentons ensuite l'adaptation de cette technique à la problématique du fémur.

Pour comprendre les différentes terminologies médicales utilisées dans la suite de ce texte, il est fortement conseillé de se référer à l'annexe 3.

5.2 Etat de l'art

Les procédures existantes de localisation d'organes dans des images scanner (CT/IRM) sont souvent basées sur des techniques de traitement d'images, comme par exemple au moyen d'images échantillonnées à différentes échelles [62], où après une combinaison de filtres de seuillage et des opérateurs morphologiques [88], par modèles déformables (voir le chapitre 6). Une liste des méthodes de segmentation médicale est détaillée dans l'état de l'art de Nicholas Ayache [7]. De par leur approche, ces méthodes sont consommatrices de temps de calcul, spécialement lorsqu'elles sont appliquées à de larges jeux de données. La méthode des rayons de recherche est une approche novatrice et originale basée sur une analyse partielle dans les volumes de données et se basant exclusivement sur la géométrie et l'apparence des objets à détecter.

5.3 Localisation par rayons de recherche

La méthode des rayons de recherche est peu utilisée dans le domaine de la segmentation d'images scanner. Les travaux précurseurs de Christian Lorenz et al. ont validé cette approche pour la détection de vertèbres [64], ou de l'aorte [66] dans des images CT, du myocarde et de la ventricule gauche dans des images CT angiographiques [65]. Plus récemment, le travail de diplôme de Jäne [52] du Fraunhofer Institute propose une méthode plus complète pour la segmentation du coeur dans des images CT basée, entre autres, sur les rayons de recherche. Pour être exhaustif, signalons le travail de semestre de Dörig [30] dans le cadre de la segmentation de la partie proximale du fémur.

Le principe de cette technique est de générer des rayons dans le volume formé par les images. Le profil d'intensité de chaque rayon est ensuite analysé pour voir si l'objet à rechercher a été traversé. Basé sur certains indices du profil, le rayon résultant est soit classé comme négatif ou comme potentiellement positif. Dans ce dernier cas, on l'appelle un candidat. Pour chaque candidat, une validation est réalisée par un envoi récursif de nouveaux rayons. Selon Lorenz, le processus de recherche par lancement de rayons *Ray Casting* est constitué des éléments suivants :

- Un générateur (*Ray Generator*) définit la méthode d'envoi de rayons, par exemple, par de droites parallèles, ou par des faisceaux de droites. Dans ce dernier cas, il faut distinguer les faisceaux coplanaires, sphériques ou cylindriques (voir la figure 5.1) ;

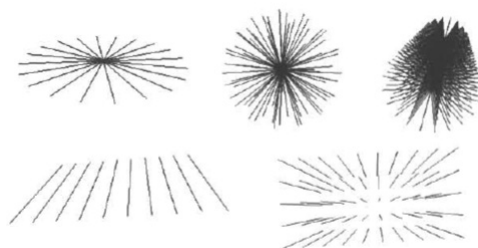


FIG. 5.1: Exemple de générateurs de rayons, tirés du travail de Lorenz ^[65]. En haut, de gauche à droite : faisceau de droites coplanaires, sphériques et cylindriques. En bas, droites parallèles coplanaires en non-coplanaires.

- Un échantillonneur de rayons (*Ray Sampler*), permet d'extraire le profil d'intensité le long du rayon (qui n'est pas forcément dans une coupe orthogonale) ;
- Un analyseur de rayon (*Ray Analyser*) qui, à partir d'un profil, va déterminer la probabilité de la présence d'un objet $p(obj)$ et en cas de succès ($p(obj) > k$) va générer un candidat ;
- Un candidat (*Ray Result*) est un rayon retenu par l'analyseur. Il comprend les informations sur sa position et les voxels traversés (signal 1D) ainsi qu'une probabilité d'appartenance à l'objet recherché ($p(obj)$).

Bien entendu, chaque résultat peut être utilisé pour initier une validation supplémentaire, ou être le point de départ pour la détection d'un nouvel objet. C'est pour cette raison que la technique est appelée récursive.

Lorenz propose par exemple de rechercher un objet allongé de la manière suivante. Initialement, un générateur de rayons parallèles non-coplanaires est utilisé pour localiser l'objet dans le volume. Dans un deuxième temps, on utilise un générateur de faisceaux sphériques sur les candidats potentiels pour identifier l'axe de l'objet recherché. La validation finale se faisant en utilisant un générateur cylindrique appliqué le long de l'axe.

5.4 Classification des données scanner

L'excellente étude de Zoroofi et al. ^[107] de l'université d'Osaka, propose de classer les images CT de la hanche en quatre catégories en fonction de leur difficulté de segmentation :

- Groupe 1 : L'acétabulum et la tête fémorale sont distinctement séparées

l'une de l'autre, la distribution des intensités osseuses de la hanche est uniforme et le niveau de gris de l'os est considérablement plus élevé que les tissus environnants (voir figure 5.2 a).

- Groupe 2 : L'acetabulum et la tête fémorale sont proches l'une de l'autre, la forme de la tête fémorale est approximativement une ellipse 3D et, autant l'os cortical que l'os spongieux ont une intensité plus grande que les tissus environnants (voir figure 5.2 b).
- Groupe 3 : L'acetabulum et la tête fémorale sont physiquement adjacents, les pathologies et malformations du bassin et du fémur causent une dégénérescence de la tête fémorale. Celle-ci n'a plus la forme d'une ellipse 3D. Les tissus osseux dans différentes parties de l'acetabulum et de la tête fémorale ne sont pas uniformément distribués (voir figure 5.2 c).
- Groupe 4 : L'espace articulaire cartilagineux a complètement disparu à cause de maladies osseuses très prononcées (coxarthrose). La distribution de l'intensité osseuse est très chaotique et plusieurs trous apparaissent dans les os spongieux (voir figure 5.2 d).

Les tests effectués dans le cadre de ce travail ont été réalisés sur la base de 7 séries d'images CT provenant de patients adultes ayant diverses pathologies de la hanche (voir annexe 3). Ces séries tombent dans les groupes précités suivants :

| Groupe | Nombre de séries |
|--------|------------------|
| 1 | 0 |
| 2 | 1 |
| 3 | 3 |
| 4 | 3 |

TAB. 5.1: Classification des séries CT utilisées lors de ce travail suivant les groupes proposés par Zoroofi et al. ^[107]

Les patients, dont nous avons utilisé les données, ont tous nécessité une arthroplastie.

Par souci de simplification, notre implémentation des rayons de recherches est intimement liée au scanner utilisé ainsi qu'au protocole radiologique. Une modification de ces paramètres nécessite des adaptations mineures de la procédure décrite ci-après.

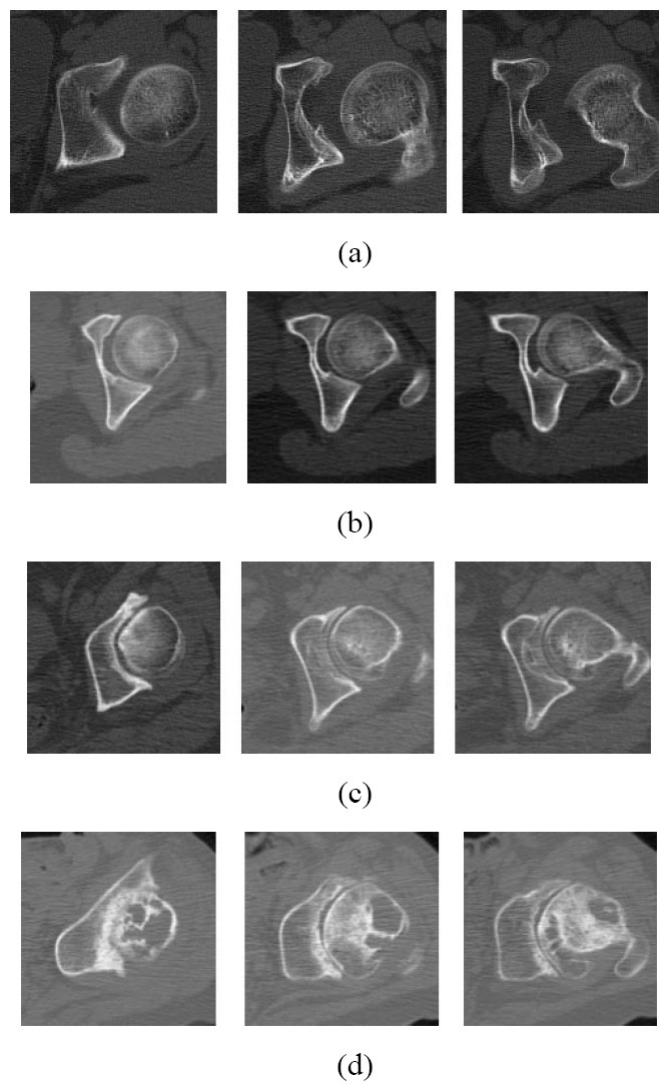


FIG. 5.2: Figure tirée du travail de Zoroofi et al. ^[107] présentant la classification des pathologies en 4 groupes distincts (voir texte). Ceux-ci sont classés dans l'ordre croissant de dégénérescence : (a) type 1, (b) type 2, (c) type 4 et (d) type 4.



FIG. 5.3: Vue radio d'un clou fémoral intra-médullaire inséré dans la diaphyse.

5.4.1 Recherche de la diaphyse

La diaphyse est la partie allongée du fémur (voir annexe 3). Elle est rectiligne et évidée. Elle contient une partie importante de la moelle du corps. En cas de fracture, on peut facilement y glisser un clou intra-médullaire sur toute sa longueur (voir figure 5.3). Dans sa partie médiane, on peut l'approximer par un cylindre de rayon interne et externe constant. Nous posons l'hypothèse que pour un patient adulte, le diamètre interne est d'environ 15 mm et le diamètre externe est compris entre 35 et 50 mm. Cette constante permet de calculer l'écartement minimum entre les rayons de recherche (dans une configuration de droites parallèles coplanaires - voir la figure 5.1) pour assurer une détection par au moins un rayon. Selon le théorème d'échantillonnage de Shannon, cet écartement est d'au plus 17,5 mm. Dans notre implémentation, nous utilisons un générateur de rayons parallèles séparés de 10 mm (voir la figure 5.4).

Etant donné que les coupes de notre protocole radiologique sont transverses, elles sectionnent le fémur au niveau de la cuisse suivant une forme circulaire (voir figure 5.5). Un rayon de recherche traversant la diaphyse dans sa région centrale possédera deux maxima sur son profil d'intensité (voir figures 5.4 et 5.5). L'intérêt de commencer la recherche dans ces coupes est que le fémur est la seule partie osseuse, ce qui réduit considérablement les erreurs de détection.

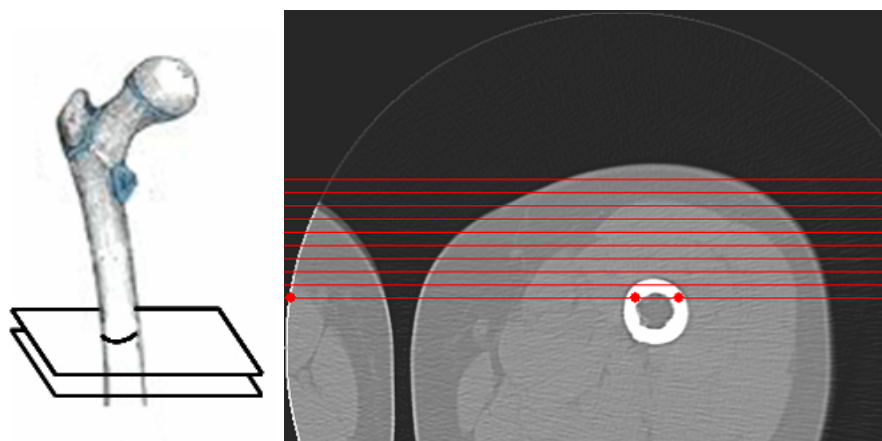


FIG. 5.4: A gauche, position de la coupe scanner, à droite, rayons tirés pour la localisation de la diaphyse. Les points représentent les régions où potentiellement se trouvent les tissus osseux.

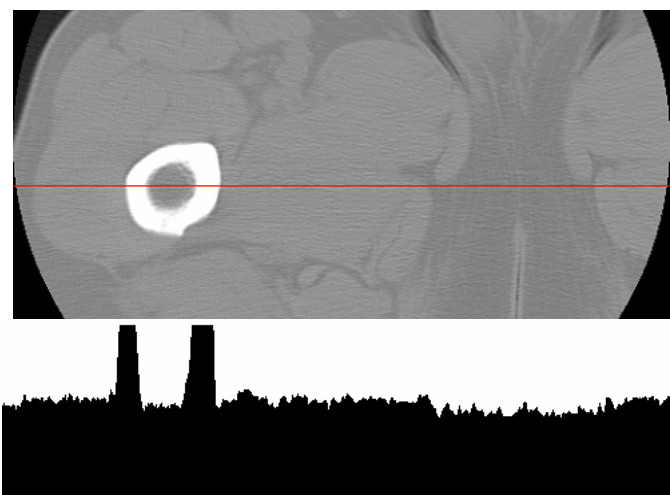


FIG. 5.5: En haut, rayon coupant la diaphyse en son centre, en bas, profil d'intensité correspondant.

5.4.2 Analyse des profils diaphysaires

Les imageurs médicaux sont théoriquement calibrés pour respecter l'échelle absolue définie par l'inventeur du scanner médical, soit Sir Godfrey Newbold Hounsfield. Cette échelle, définie en unités HF, va de l'air (-1000 HF), en passant par l'eau (0 HF), jusqu'à l'os (+1000 HF). On remarque cependant de grandes libertés d'implémentation d'un fabricant de scanner à l'autre. Aussi, la détection initiale de la diaphyse dans les profils d'intensité ne peut pas se faire directement par seuillage, d'autant plus que l'intensité de l'os dépend de nombreux autres facteurs comme sa corticalité qui peut varier en fonction de pathologies comme l'ostéoporose, ou divers artefacts présents dans l'image, comme par exemple, la proximité d'un objet métallique.

Les deux pics diaphysaires peuvent cependant être localisés au moyen l'intensité absolue du signal, complété par sa dérivée. On prendra garde, qu'appliquée directement au profil, cette méthode générera une quantité importante de bruit. Aussi, il est nécessaire d'utiliser un filtre passe bas pour améliorer la qualité de la détection. Se basant sur la théorie des espaces multi-échelles (*Scale Space*) [82], un filtre passe-bas peut être vu comme la convolution d'une Gaussienne au signal (voir figure 5.6) :

$$\begin{cases} S * G \\ G(X) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \end{cases}$$

où $*$ est le symbole de convolution, μ la moyenne et σ la variance de la Gaussienne. Une augmentation de cette dernière valeur rendra le signal de plus en plus flou. Cette théorie nous apprend que la convolution de la dérivée d'ordre $n \in \mathbb{N}^+$ d'une fonction Gaussienne avec un signal est distributive, c'est-à-dire que :

$$\Delta^n(S * G) = S * \Delta^n G \quad (5.1)$$

où Δ^n est l'opérateur de dérivation d'ordre n . Cette propriété permet d'accélérer la vitesse de calcul du filtre car il ne s'agit plus d'effectuer deux opérations successives gourmandes en temps de calcul, c'est-à-dire la corrélation du signal avec la Gaussienne suivi d'une dérivation, mais d'une seule opération qui est la corrélation du signal avec la dérivée de la Gaussienne. Cette dernière fonction pouvant être précalculée et stockée dans une table.

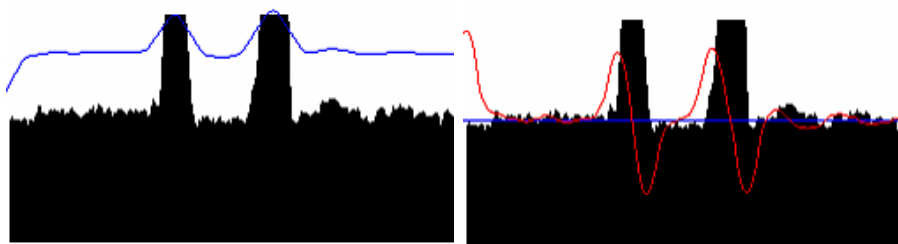


FIG. 5.6: A gauche, convolution du profil avec une Gaussienne, à droite, sa dérivée première. On remarque bien les passages à zéro aux maxima.

Nous avons vu précédemment que le critère de détection de la diaphyse dépend de l'identification de deux pics écartés d'une distance théorique de d_{dia} . Nous pouvons donc écrire la fonction de probabilité de détection suivante :

$$p_{dia}(r) = e^{-\delta \frac{k}{m}} \quad (5.2)$$

où r est le rayon mesuré, δ est défini comme :

$$\delta = \|x_{max1} - x_{max2} - d_{dia}\|$$

où x_{max1} et x_{max2} sont deux maxima d'intensité du signal, tel que $x_{max1} \geq x_{max2}$. À noter que toutes les combinaisons sont testées si plus de deux maxima sont présents dans le profil. La fonction δ indique l'écart Euclidien entre le diamètre extérieur théorique de la diaphyse et sa mesure. Celui-ci peut facilement fluctuer car le rayon ne passe pas forcément par le centre diaphysaire. La variable m de l'équation 5.2 est définie comme étant la différence entre l'ordonnée du second maximum y_{max2} et l'intensité moyenne du signal. Il s'agit donc d'un facteur de pondération qui octroie plus d'importance à une grande différence d'intensité entre l'os et les autres tissus traversés. Finalement, le facteur $k \in \mathbb{R}^+$ permet de pondérer le tout. Pour simplifier, l'équation 5.2 génère de manière non-linéaire une probabilité proche de 1 si les deux maxima s'approchent du diamètre théorique fixé.

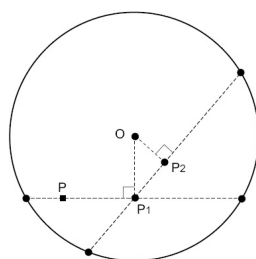


FIG. 5.7: Tiré du travail de Jiang ^[54]. Méthode de construction géométrique permettant la convergence des rayons vers le centre du cercle. Voir le texte pour plus d'explications.

5.4.3 Recherche de l'axe diaphysaire

Dans le domaine de la calibration de caméras, Guang Jiang et Long Quan ^[54] proposent une méthode intéressante de détection de centres de cercles. Cette technique itérative, basée sur l'analyse successive de segments, peut très facilement être adaptée à nos rayons de recherche. Soit n'importe quelle droite passant par un point P inscrit dans un cercle (voir figure 5.7). Cette droite coupe le cercle en deux points dont P_1 est le centre. Si maintenant le point P_1 devient le point P et que l'opération est répétée, les points successifs (P_2, P_3, \dots) vont converger vers le centre O du cercle. A noter que si le nouveau segment est perpendiculaire au précédent, la convergence se réalise en une seule itération. Pratiquement, l'utilisation d'un angle plus faible augmente la robustesse de l'algorithme. Jiang propose une démonstration intuitive dans son article. Pour une démonstration et une analyse géométrique plus formelle, se référer à l'annexe 4.

Dans notre implémentation, l'algorithme de Guang Jiang et Long Quan est répété 20 fois pour chaque rayon candidat avec un angle de 50° entre les segments successifs (voir figure 5.8). A chaque itération, nous calculons la distance entre le centre et celui calculé précédemment. Si le processus s'effectue dans une forme circulaire, cette mesure doit nécessairement diminuer. Nous estimons qu'il y a convergence vers l'axe diaphysaire si la distance entre les cinq derniers points est inférieure à un seuil prédéfini. Lorsqu'un centre est trouvé, celui-ci sert de point de départ pour la recherche des centres situés dans les coupes adjacentes. Les longueurs des différents segments permettent également d'estimer les rayons interne et externe de la diaphyse. Cet algorithme de recherche est propagé sur toutes les coupes jusqu'à divergence du calcul du centre qui intervient la plupart du temps dans les coupes situées

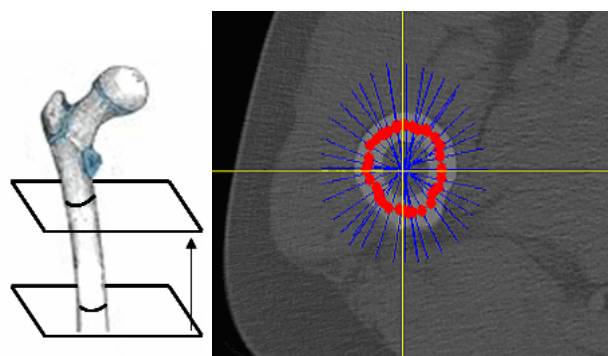


FIG. 5.8: A gauche, position de la coupe scanner, à droite, différents segments de l'algorithme de recherche de l'axe diaphysaire.

au-dessus du petit trochanter. Les différents centres détectés servent à estimer l'axe diaphysaire par interpolation d'une droite au moyen des moindres carrés.

5.4.4 Localisation du grand trochanter

L'axe diaphysaire calculé durant l'étrape précédente va couper le fémur dans sa partie proximale au niveau de la fossette digitale sur le col du fémur. Comme le trochanter se trouve relativement proche de cet axe, on le recherche par analyse de l'histogramme des pixels situés dans la prolongation du cylindre du corps fémoral. L'algorithme scanne les coupes de haut en bas en démarrant à la coupe la plus haute. Nous estimons que la coupe contenant l'extrémité proximale du grand trochanter est trouvée lorsqu'une proportion suffisante de pixels osseux est détectée. Cet algorithme sera développé dans le paragraphe suivant. Une méthode alternative basée uniquement sur les rayons de recherche est disponible dans le travail de semestre de D. Dörig [30].

5.4.4.1 Classification des pixels osseux

L'extraction des différents tissus d'une image CT s'effectue couramment par une analyse statistique de l'histogramme [79]. Celui-ci peut être représentés par une superposition de distributions de Laplace-Gauss (*Gaussian Mixture*) dont les tissus peuvent être extraits par des méthodes de classification comme par *l'expectation-maximisation* [106].

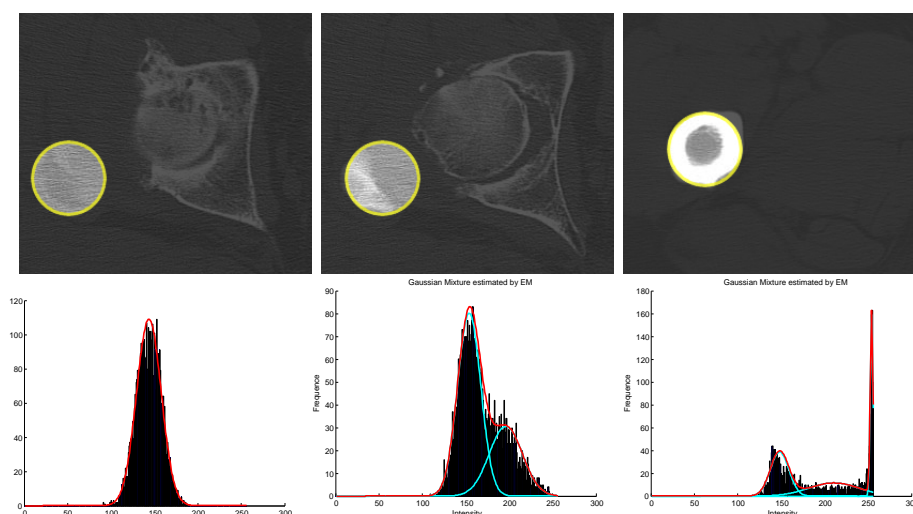


FIG. 5.9: En haut, position du cylindre diaphysaire dans différentes coupes. En bas, histogramme correspondant et interpolation par des courbes gaussiennes. De gauche à droite : exemple de coupe de type 1, 2 et 3 (voir texte).

Dans notre cas de détection du grand trochanter dans le cylindre diaphysaire, nous pouvons distinguer trois régions distinctes :

1. Dans les coupes situées au dessus du grand trochanter, seul les tissus graisseux et musculaires sont visibles. Comme leur intensité est très proche dans les images CT, l'histogramme de ces coupes peut aisément être interpolé par une seule gaussienne (voir figure 5.9 - gauche).
2. Dans la première coupe qui comprend le grand trochanter apparaissent des pixels osseux dont l'intensité est atténuée par l'épaisseur de la coupe. Aussi, deux distributions gaussiennes avec une moyenne assez proches sont nécessaires pour interpoler l'histogramme (voir figure 5.9 - milieu).
3. Les coupes inférieures présentent un os cortical de plus grande intensité. Dans ce cas, l'histogramme peut être représenté par la superposition de trois distributions normales (voir figure 5.9 - droite).

Bien que d'une grande efficacité, l'utilisation de cet algorithme statistique pour déterminer quelle est la première coupe comprenant le grand trochanter est une procédure lourde en calcul. Sous Matlab, cette analyse prend plusieurs secondes sur un ordinateur récent. De plus, une gaussienne est, de par sa nature, décomposable en une combinaison de courbes du même type [82].

La table 5.2 corrobore cette hypothèse et montre qu’une détermination uniquement basée sur la corrélation entre l’histogramme et le mélange de gaussiennes est difficile. Par contre, l’utilisation de la variation de l’écart-type par approximation d’une seule gaussienne permet une localisation robuste et instantanée. Dans cet exemple précis, le passage de l’écart-type en dessus de 15 indique le début du grand trochanter.

| type de coupe | 3 gauss. | 2 gauss. | 1 gauss. | μ | σ |
|---------------|----------|----------|----------|----------|----------|
| 1 | 0,9913 | 0,9918 | 0,9912 | 143,8376 | 14,4838 |
| 1 | 0,9908 | 0,9912 | 0,9899 | 143,6265 | 14,2230 |
| 2 | 0,9882 | 0,9876 | 0,9012 | 168,8054 | 25,4869 |
| 3 | 0,9864 | 0,9835 | 0,9376 | 156,9936 | 25,8845 |
| 3 | 0,9272 | 0,8400 | 0,2795 | 196,1231 | 46,8706 |

TAB. 5.2: Résultats de corrélation entre l’histogramme et le mélange de gaussiennes lorsque trois, deux ou une courbes sont utilisées pour l’approximation. Dans le dernier cas, la moyenne et l’écart-type de la gaussienne sont également représentés.

5.4.5 Recherche de la tête fémorale

Anatomiquement, la tête fémorale d’une personne saine a la forme de $2/3$ de sphère (voir annexe 3). Nous savons également que la coupe comprenant le grand trochanter divise grossièrement la tête en son centre et que celui-ci se situe approximativement sur la droite définie par le trochanter et l’axe diaphysaire. Comme, à ce stade, la position du trochanter n’est pas connue, sa direction est estimée par l’axe du cylindre et la moyenne pondérée des pixels localisés dans le tube diaphysaire (voir figure 5.10). Dans le cas où la moyenne pondérée est trop proche du centre du cylindre (voir figure 5.10 centre), la direction de l’axe est, dès lors, mal conditionnée. Cette faiblesse peut être levée par l’application du même algorithme sur la coupe directement inférieure ou, préférablement, par l’emploi de la technique développée par D. Dörig ^[30]. Dix rayons parallèles à cette direction, centrés sur cette droite et distants de 2 mm sont ensuite tirés. Un premier point de la tête est sélectionné à l’endroit où au moins deux de ces rayons coupent un os cortical. Le point milieu de ces deux candidats définit un premier point sur la surface de la tête fémorale. On définit approximativement la position du centre se situant à une distance de $\frac{3}{2}r_{diaphyse}$ de ce point.

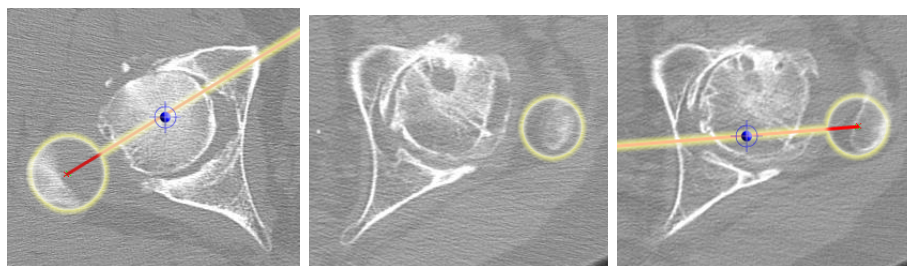


FIG. 5.10: Méthode de recherche du centre de la tête fémorale. Le cercle jaune représente la prolongation du cylindre diaphysaire. En rouge, approximation de la droite passant par le grand trochanter et par l'axe diaphysaire (voir texte) et en bleu, estimation du centre de la tête fémorale. A milieu, exemple où la détermination de la droite par l'algorithme présenté est impossible car la position du trochanter est trop proche de l'axe diaphysaire. L'image de droite représente la coupe directement inférieure de la même personne sur laquelle une recherche de la tête fémorale a été effectuée.

5.4.5.1 Approximation de la tête fémorale par un cercle

La première étape pour l'approximation de la tête fémorale par un cercle consiste à déterminer son contour extérieur. Elle est réalisée par l'envoi de rayons dans huit directions centrés sur l'approximation du paragraphe précédent. Chaque générateur est composé de sept rayons parallèles, centrés, distants de 1 mm et de longueur $2r_{diaphyse}$ (voir figure 5.11 gauche). L'analyse du profil de ces rayons est complexe car la densité de l'os varie énormément en fonction de la pathologie et le contour de la tête est souvent flou. Aussi, au lieu d'utiliser le seuillage pour déterminer le bord du fémur, nous utilisons trois conditions sur les maxima locaux :

1. Tous les rayons du même générateur possèdent un maximum situé approximativement à la même distance du centre ;
2. Ce maximum est séparé d'une distance minimale des autres maxima du même rayon, ce qui permet d'éviter des régions avec de fortes variations de contraste ;
3. Les candidats de chaque faisceau doivent se trouver à la même distance du centre ;

Nous cherchons ensuite à approximer le cercle centré sur ces huit candidats par la technique des moindres carrés.

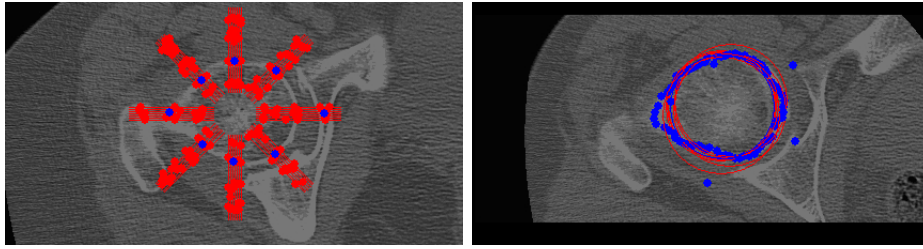


FIG. 5.11: A gauche, calcul des candidats pour l'approximation d'un cercle. A droite, les neuf cercles possibles et leurs candidats (voir texte).

Comme nous le constatons à la figure 5.11, ces conditions ne sont pas suffisantes pour éviter la détection de faux positifs localisés sur l'acetabulum ou sur le tubercule. Pour palier à ce problème, le processus décrit dans cette section est répété huit fois en incrémentant l'angle initial de 5° à chaque itération. Au final, nous avons neuf approximations de cercles (voir figure 5.11 droite). Nous conservons celui dont l'erreur de projection de tous les candidats est minimale. Ce processus est très proche de l'algorithme RANSAC [36] et propose des résultats plus fiables qu'une approximation directe à partir de tous les candidats.

5.4.5.2 Approximation de la tête fémorale par une sphere

D'après notre expérience, la modélisation de la tête fémorale par une boule plutôt que par une ellipse est judicieuse et ceci quel que soit le degré de dégénération pathologique. En effet, l'approximation par ce dernier type de conique permet, en théorie, une interpolation plus fidèle grâce aux degrés de libertés supplémentaires mais, pratiquement, nous constatons qu'elle réduit considérablement la robustesse de l'approximation.

Le processus d'approximation de la tête fémorale commence par la recherche des cercles situés dans les coupes connexes jusqu'à divergence de l'algorithme. L'interpolation de la sphère se base sur la sélection des quatre cercles dont les candidats ont la variance minimale sur l'erreur de projection. Uniquement les candidats ayant l'erreur la plus faible sont conservés. Le problème devient l'interpolation d'une sphère par des points et peut être résolu au moyen des moindres carrés (pour un développement mathématique complet, voir l'annexe I du travail de D. Dörig [30]).

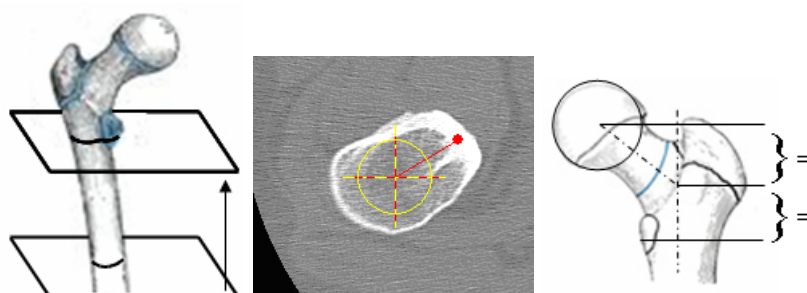


FIG. 5.12: A gauche, position de la coupe scanner dans le jeu de données, au centre, coupe avec rayon de recherche du petit trochanter (trait rouge) et, à droite, hypothèse anatomique sur l'angle de l'axe diaphysaire (voir texte).

5.4.6 Recherche de l'axe du col du fémur

Pour obtenir un modèle complet du fémur, la localisation de l'axe de son col est encore nécessaire. Anatomiquement, celui-ci se situe dans le plan formé par le centre de la diaphyse et passe par le centre de sa tête. Reste encore à calculer sa direction. Nous suggérons qu'il coupe le centre de la diaphyse dans la coupe située à équidistance entre le centre de la tête fémorale et le petit trochanter (voir figure 5.12).

Pratiquement, la direction de recherche du petit trochanter s'effectue depuis la coupe distale et monte jusqu'à celle contenant le centre de la tête fémorale. Dans chaque coupe, un rayon est tiré depuis l'axe diaphysaire dans la direction (projetée) de la tête fémorale. Le critère d'arrêt du rayon est le maximum correspondant à l'os cortical. On exprime ensuite la distance entre ces deux points sur toutes les coupes analysées. Celle comprenant le petit trochanter correspond au maximum local le plus important de la fonction. Si cet extremum n'est pas relevant, le critère utilisé est la première coupe ayant une distance de deux fois le rayon moyen de l'axe diaphysaire.

5.5 Résultats

Nous sommes les premiers à valider l'approche de la localisation des points anatomiques par lancers de rayons de recherche sur des tissus osseux. Cette méthode permet de localiser précisément l'axe de la diaphyse, son diamètre, la position des trochanters, le centre de la tête fémorale, son diamètre ainsi que l'axe du col. Parmi les sept séries CT de patients présen-

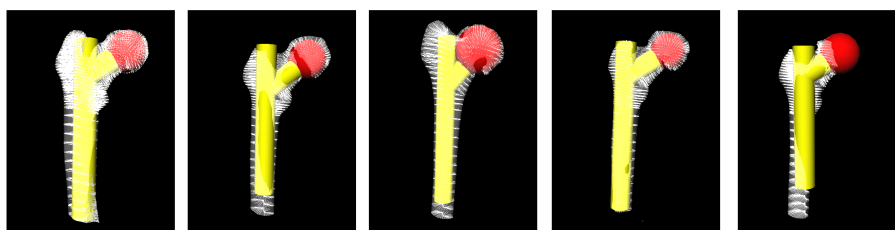


FIG. 5.13: Détection automatique des paramètres anatomiques du fémur de cinq patients par lancés de rayons. Ces paramètres permettent de construire un modèle composé d'une sphère rouge (tête fémorale) et de deux cylindres jaunes (diaphyse et col du fémur). La surface grise transparente (extérieure du fémur) a été segmentée manuellement puis reconstruite. Les traits blancs sont les projections des points de la surface sur le modèle.

tant des degrés divers de dégénération du fémur testés dans le cadre de cette thèse, la présente technique a permis de définir précisément et rapidement les paramètres de cinq d'entre elles, c'est-à-dire avec plus de 70% de succès dans les 4 groupes ou 100% dans les groupes 1 à 3 (voir figure 5.13). Dans les deux cas restants provenant du groupe 4, le centre de la tête fémorale n'a pas été détectée correctement. En regardant les images de l'une d'elle plus attentivement (voir figure 5.14) on constate que cette tâche n'est même pas facile à réaliser manuellement tellement la dégénérescence liée à la coxarthrose et à l'ostéoporose est importante.

Pour valider quantitativement cette méthode de recherche des caractéristiques anatomiques, nous avons tenté de la comparer à une estimation manuelle des même paramètres. Pour pouvoir comparer ces résultats, nous avons ensuite réalisé un recalage rigide entre le modèle et sa reconstruction tridimensionnelle. Le tableau suivant présente les résultats ainsi obtenus :

| | série 1 | série 2 | série 3 | série 4 | série 5 | moyenne |
|---------------------|---------|---------|---------|---------|---------|-------------|
| Moyenne (auto) | 4,11 | 3,67 | 3,79 | 3,76 | 3,88 | 3,84 |
| Ecart type (auto) | 2,88 | 2,47 | 2,91 | 2,61 | 2,75 | 2,73 |
| Moyenne (manuelle) | 3,79 | 3,51 | 3,66 | 3,37 | 3,25 | 3,52 |
| Ecart type (manuel) | 2,68 | 2,27 | 2,78 | 2,29 | 2,42 | 2,50 |

TAB. 5.3: Comparaison de la recherche automatique des caractéristiques par l'algorithme des rayons de recherche (auto) et de leur estimation manuelle (manuel) sur les cinq séries testées. Les chiffres indiquent l'erreur Euclidienne (moyenne et écart type) de recalage rigide entre le modèle et la surface tridimensionnelle extérieure correspondante.

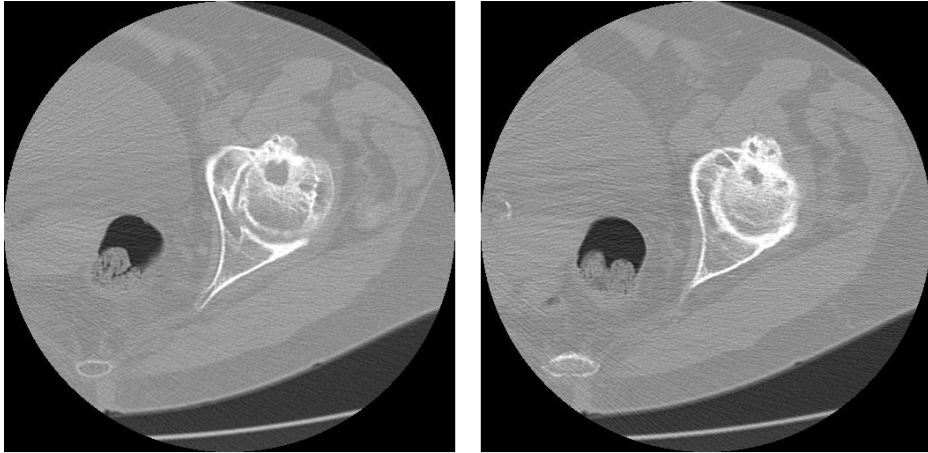


FIG. 5.14: Deux coupes d'une série présentant des problèmes lors de la localisation de la tête fémorale. On constate que le bord de la tête est flou (coxarthrose) et que l'os comporte des trous (ostéoporose).

L'avantage de notre approche est de permettre également une localisation interactive des caractéristiques anatomiques. Une estimation manuelle des paramètres présente des résultats légèrement plus précis que l'approche automatique (de l'ordre de 10%). L'erreur principale de l'algorithme automatique semble se situer dans la détermination précise du grand trochanter. Nous verrons par la suite que, quel que soit l'approche utilisée, l'estimation des paramètres est suffisante pour initialiser un modèle déformable.

5.6 Conclusion

Ce chapitre présente une méthode originale de localisation et détermination de paramètres anatomique de la partie sagittale du fémur par une technique de rayons de recherche. Il s'agit de la première contribution théorique de travail. L'algorithme présenté est extrêmement rapide (de l'ordre de quelques secondes) et fiable. Il a été testé sur huit séries de patients atteints de différentes pathologies fémorales. Son taux de réussite approchant 70% sur des séries CT comportant de fortes dégénérescences osseuses (1 cas). Les problèmes survenus dans 30% des cas sont principalement liés à la détermination de la tête fémorale, les autres paramètres étant estimés correctement. Soulignons que dans le cas d'un échec (partiel) de l'algorithme, il est trivial pour un médecin de placer ou corriger interactivement les paramètres

anatomiques non-détectés.

L'approche par rayons de recherche permet la construction d'un modèle simplifié du fémur comportant deux cylindres et une sphère. Quantitativement, la localisation des paramètres anatomiques automatique est environ 10% moins précise que l'estimation manuelle, spécifiquement pour la détermination du grand trochanter. Nous verrons au chapitre suivant que l'estimation automatique présentée dans ce chapitre est suffisante pour l'initialisation d'un modèle déformable.

CHAPITRE VI

RECALAGE NON RIGIDE PAR MODÈLE DÉFORMABLE

Au chapitre 4 nous avons vu les principales méthodes utilisées dans le domaine médical pour le recalage non rigide global et local. Nous verrons que l'introduction d'autres contraintes géométriques, liées à des problèmes particuliers, permet des transformations plus robustes. Contrairement aux méthodes précitées, il est ainsi possible de réduire considérablement le nombre de degrés de liberté lors de la transformation. Ce chapitre présente un modèle déformable original intégrant des contraintes de forme particulières et, dont l'avantage est de pouvoir s'initialiser simplement à partir des données morphologiques, comme par exemple, celles obtenues au chapitre 5. Le résultat de ce recalage de forme est loin d'être parfait mais il permet, entre autre, l'initialisation rapide d'un algorithme de segmentation tel que celui décrit au chapitre suivant.

6.1 Introduction

Ayache ^[7] propose la définition suivante des modèles déformables :

"Les modèles déformables sont des courbes ou des surfaces qui évoluent dans un espace 3D pour délimiter au mieux une structure anatomique ou pathologique. L'évolution du modèle est guidée par l'optimisation simultanée de deux critères. Le premier mesure la régularité géométrique du modèle, en s'appuyant par exemple sur le calcul de sa courbure locale. Le second mesure une propriété de l'intensité lumineuse le long des points de l'image traversés par le modèle déformable, par exemple la norme du gradient de l'intensité lumineuse pour favoriser l'attraction du modèle vers des points de fort contraste. Ces méthodes sont très efficaces lorsque l'utilisateur peut initialiser le modèle de manière approximative autour d'une région d'intérêt. Le modèle déformable évolue alors depuis cette solution initiale grossière, pour

délimiter avec précision et de manière automatique la frontière de la région recherchée. Cette approche est généralement beaucoup plus robuste que le simple seuillage de l'intensité lumineuse. Ces modèles sont également très populaires dans l'analyse des séquences temporelles d'images : en effet, la solution à un instant donné sert d'initialisation dans l'image acquise à l'instant suivant. Sur l'utilisation des modèles déformables en analyse d'images médicales, un excellent état de l'art a été proposé par McInerney and Terzopoulos [50]."

Montagnat [74] propose de formaliser la définition des modèles déformables d'Ayache par la relation suivante :

$$f(\vec{p}_i) = \alpha(i)f_{int}(\vec{p}_i) + \beta(i)f_{ext}(\vec{p}_i), \forall i \quad (6.1)$$

où \vec{p}_i sont les points d'un maillage. Les coefficients $\alpha(i)$ et $\beta(i)$ sont contraints d'appartenir aux intervalles $[0; \frac{1}{2}]$ et $[0; 1]$ respectivement pour assurer la stabilité du schéma numérique. Les forces internes (f_{int}) et externes (f_{ext}) peuvent être d'amplitude très différente. Les forces internes sont très localisées puisque calculées en fonction du voisinage de \vec{p}_i géométriquement très proches. Les forces externes en revanche sont homogènes à un déplacement vers des points de données qui peuvent être très éloignés. L'équation 6.1 se résout de manière itérative, où à chaque pas, le modèle est calculé au moyen des contraintes de régulation internes et externes du modèle de l'itération précédente. Il est possible de rajouter des contraintes spécifiques à un problème particulier pour réduire l'espace de déformation du modèle. On peut, par exemple, ajouter une contrainte axiale à des modèles présentant une symétrie de révolution.

6.2 Modèle articulé

Jusqu'à présent, nous avons vu des modèles déformables d'une grande complexité (comme les fonctions à base radiale - RBF), qui ne sont pas spécifiques à un espace de déformation particulier. Pratiquement, ceux-ci ont souvent une convergence lente et chaotique vu leur nombre élevé de degrés de libertés (DDL). L'introduction de contraintes supplémentaires est une première étape pour l'amélioration de la convergence du recalage.

On peut toutefois pousser la simplification plus loin en limitant l'espace

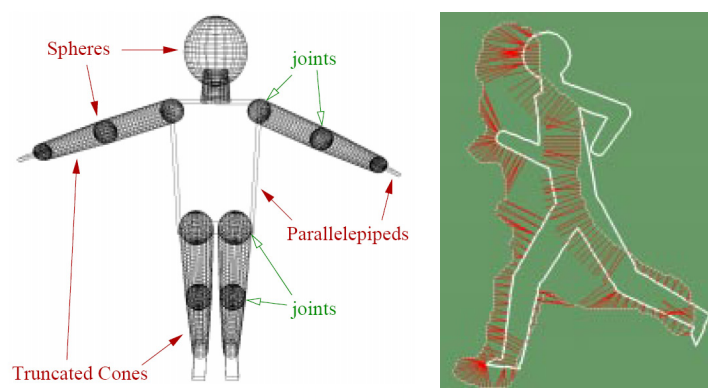


FIG. 6.1: A gauche, modèle articulé simplifié d'un être humain proposé par Delamarre et Faugeras ^[25], à droite, forces de mise à jour (rouge) entre la silhouette d'une personne réelle et une position du modèle articulé.

de déformation à des modèles 3D articulés comme, par exemple, celui proposé par Delamarre et Faugeras ^[25] dans le domaine de suivi de personnes sur des séquences vidéo. Leur travail présente une méthode pour mettre à jour un modèle 3D d'un être humain au moyen de la silhouette d'une personne en mouvement. Ces forces sont appliquées à chaque partie rigide du modèle. L'intérêt de leur étude est de démontrer que l'utilisation d'un modèle approximatif (voir la figure 6.1 gauche) est suffisant pour retrouver la posture de la personne. Leur modèle est composé des primitives suivantes : des cônes tronqués pour représenter les bras et les jambes, des sphères pour le cou, les articulations et la tête et des parallélépipèdes rectangles pour les mains, les pieds et le corps. Des liens sériels permettent de connecter ces différentes primitives. Le modèle complet comporte ainsi 20 DDL. Le modèle 3D est projeté dans les plans images des caméras pour comparer la pose du modèle articulé avec la posture de la personne réelle (voir la figure 6.1 droite). L'algorithme de mise à jour du modèle se base sur la différence entre les deux formes dans le plan image pour calculer une force de correction qui est basée sur les démons de Maxwell ^[97]. L'emploi de plusieurs points de vue, par exemple au moyen de caméras stéréoscopiques, permet de résoudre la plupart des ambiguïtés liés à la profondeur et aux occlusions. Bien que cette méthode présente de bons résultats sur des séquences continues d'image - la convergence prend environ 0,1 seconde sur des ordinateurs de l'époque - cet algorithme nécessite une initialisation manuelle, ce qui limite pratiquement son utilisation.

Notons qu'au chapitre précédent, nous avons implicitement utilisé cette

méthode pour présenter les résultats de la localisation automatique de parties anatomiques du fémur (voir la figure 5.13). Le modèle fémoral est, dans notre cas, composé de cylindres pour représenter la diaphyse et le col du fémur et d'une sphère pour la tête. Nous justifions l'emploi d'un modèle articulé pour représenter les différentes morphologies de l'os. Ce modèle simplifié à l'extrême, qui comporte uniquement 13 DDL, permet une identification grossière du contour de l'os mais est insuffisant pour un recalage médical précis.

6.3 Du squelette à la peau

Dans le modèle articulé de Delamarre et Faugeras^[25], présenté à la section précédente, les primitives utilisées possèdent deux rôles distincts. Les sphères aux articulations sont des éléments de jonctions permettant d'obtenir une silhouette continue et n'ont pas de réels DDL propres, tandis que les autres primitives possèdent véritablement les informations de pose.

Dans leur travail sur les simulateurs laparoscopiques, Conti et al.^[24] ont développé un modèle dynamique, sorte de compromis entre les systèmes masse-ressorts et les modèles à éléments finis (*FEM*). A partir d'un maillage fermé, la construction de leur modèle s'effectue en trois étapes :

- Le calcul du squelette du maillage au moyen d'un axe médian^[13] (voir la figure 6.2 (a)),
- L'utilisation d'un modèle volumique, composé de sphères pleines placées le long de cet axe et interconnectées via des liens élastiques tridimensionnels (voir la figure 6.2 (b)). Ces sphères sont des sortes d'éléments finis dont les forces internes sont transmises via ces liens,
- Une enveloppe connectée au modèle volumique au moyen de liens élastiques individuels (voir la figure 6.2 (c)). Cette connexion s'effectue en cherchant pour chaque point du maillage le lien ou la sphère la plus proche. Lorsqu'une sphère ou un lien bouge, les points qui lui sont rattachés suivent automatiquement ce mouvement. La section suivante présente plus en détail la topologie spécifique de ces connexions.

Dans le même ordre d'idée, citons les travaux de Capell et al.^[16] dans le domaine de l'animation de caractères. Comme pour Conti et al.^[24], leur modèle se base sur un squelette. L'enveloppe n'est, par contre, pas connectée à des sphères et des liens mais à un maillage de contrôle simplifié. La résolution

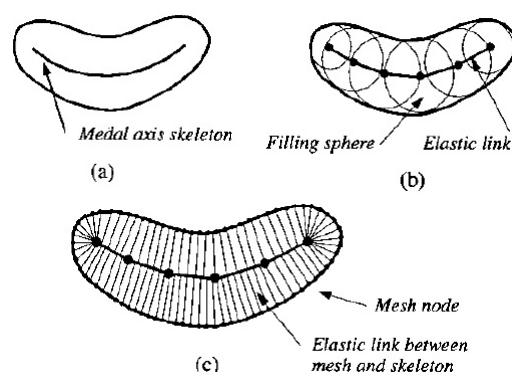


FIG. 6.2: Illustrations tirées du travail de Conti et al. [24]. (a) exemple de squelette d'un maillage, (b) modèle volumique composé de sphères inscrites dans le volume et centrées sur le maillage, (c) connexion de l'enveloppe avec le squelette.

numérique de leur système est rapide et se base sur une FEM.

Pris dans un contexte statique, l'avantage des deux dernières méthodes présentées par rapport au modèle de Delamarre et Faugeras [25] est de pouvoir dissocier les éléments constitutifs du modèle (les sphères - les liens ou maillage de contrôle) par rapport à son enveloppe.

Dans le but de réduire encore plus les degrés de libertés, nous allons, dans la suite de ce travail, augmenter les contraintes de forme en simplifiant les objets à déformer par des (assemblages de) tubes.

6.4 Cylindres généralisés

Les cylindres généralisés (GC) sont un sous-ensemble de modèles déformables paramétriques utilisés pour représenter des objets cylindriques. C'est en 1971 que Binford [11] introduit les GC dans les applications de vision. Les cylindres généralisés consistent en une courbe ou axe combiné à une fonction de coupe définie sur cet axe. La fonction de coupe étant généralement une ellipse mais elle peut également être basée sur une formulation de Frenet-Serret [105] ou des cylindres généralisés extrudés [77]. Les objets tubulaires sont définis par un élément de coupe qui est balayé le long de son axe.

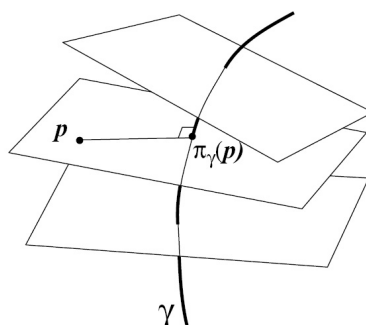


FIG. 6.3: Tiré du travail de Lazarus ^[59], exemple de système de coordonnées cylindriques généralisées. Le point p est projeté orthogonalement sur γ .

6.4.1 Système de coordonnées cylindriques

Lazarus développe dans sa thèse ^[59] un système de coordonnées pour les cylindres généralisés. Cette section reprend une partie de son travail.

Le système de coordonnées cylindrique est défini dans un voisinage de \mathbb{R}^3 à partir d'une courbe. Etant donnée une courbe plongée dans \mathbb{R}^3 , il est possible de définir sur un voisinage tubulaire V_γ de cette courbe un découpage en tranches. Elles sont constituées de morceaux de plans normaux à la courbe γ . Ce découpage constitue une fibration de V_γ de base γ et dont la projection π_γ est orthogonale sur la courbe (voir la figure 6.3). En munissant chaque fibre d'un repère et en choisissant une paramétrisation de la courbe, on peut définir un système de coordonnées :

$$\begin{aligned} V_\gamma &\rightarrow \mathbb{R}^3 \\ p &\mapsto (s, x, y) \end{aligned}$$

où s est le paramètre du point d'attache $\pi_\gamma(p)$ de p sur γ et (x, y) les coordonnées cartésiennes ou, de préférence, polaires de p dans le repère associé à sa fibre.

Notons que le choix d'un repère dans chaque fibre revient au choix d'un repère mobile pour γ .

Dans ses travaux antérieurs, Lazarus ^[60] ^[61] définit un outil de déformation pour l'image de synthèse fondé sur des systèmes de coordonnées associés à deux courbes γ_0 et γ_1 . L'idée est de fixer un objet par rapport à la première courbe (γ_0) et de lui faire subir la déformation induite par γ_1 . Un point p_0

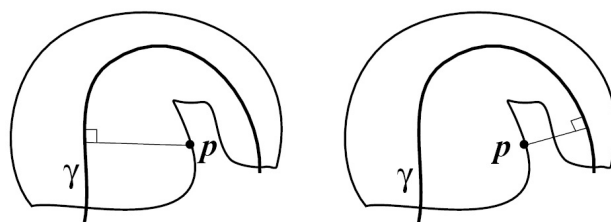


FIG. 6.4: Tiré du travail de Lazarus ^[59], exemple d'ambiguïté d'attachement du point p sur la courbe γ .

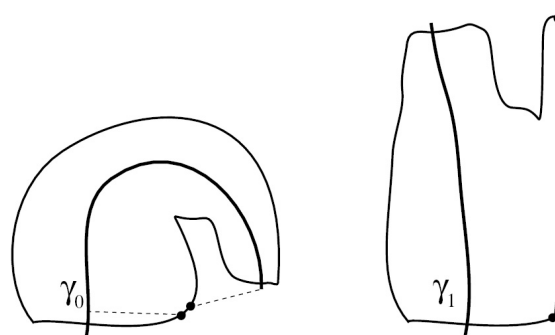


FIG. 6.5: Tiré du travail de Lazarus ^[59], exemple où la fonction d'attachement est discontinue au niveau des points noirs (figure de gauche). Après déformation de l'axe (figure de droite), le résultat est erroné.

est ainsi déformé en un point p_1 de sorte que les coordonnées cylindriques de p_1 relativement à γ_1 soient identiques à celles de p_0 relativement à γ_0 .

Afin de déformer correctement un objet avec cette technique, il faut donc s'assurer que celui-ci est contenu dans un voisinage convenable de γ_0 puisque nous devons calculer les coordonnées de ses points relativement à γ_0 . En fait, la forme de ce voisinage dépend conjointement de la courbe et de l'objet comme on peut le voir sur la figure 6.4. Dans cet exemple, on remarque qu'il n'est pas possible, à l'aide d'une information purement locale, de dire à quelle position de l'axe un point de l'objet doit être rattaché.

Dans ^[60], Lazarus propose d'attacher chaque point de l'objet au point le plus proche de l'axe. Ce procédé ne permet donc pas de lever l'ambiguïté de la figure 6.4 et peut générer des incohérences lors de déformations axiales comme l'exemple de la figure 6.5.

Son travail de thèse ^[59] propose différentes méthodes plus complexes pour

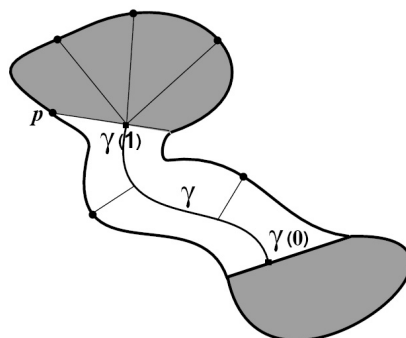


FIG. 6.6: Tiré du travail de Lazarus ^[59], les parties grisées de l'objet représentent les calottes. Les points de l'enveloppe situés sur les calottes sont codés en coordonnées sphériques avec les repères mobiles situés aux points γ_0 et γ_1 .

lever cette ambiguïté. Une méthode simple et intuitive consiste à calculer l'intersection du maillage 3D avec le plan contenant le point potentiel p . Ce point n'est accepté que si il se trouve sur une courbe fermée et si le point $\pi_\gamma(p)$ se situe à l'intérieur de la surface délimitée par cette courbe.

Dans la mesure où les objets ne sont pas entièrement attachés à l'axe mais où des parties peuvent également se situer aux extrémités) (appelées aussi les calottes), il est indispensable de s'assurer que le système de coordonnées le long de l'axe se prolonge convenablement dans les demi-espaces au dessous du premier plan et au dessus du dernier plan. Pour ce faire, nous utilisons un système de coordonnées sphériques associé au premier et au dernier repère mobile (voir figure 6.6).

6.4.2 Modèle du fémur

Le modèle déformable utilisé dans le cadre de cette thèse est un mélange entre le modèle articulé de Delamarre et Faugeras (voir section 6.2) et celui des cylindres généralisés de Lazarus présenté à la section précédente. Contrairement au modèle de Delamarre et Faugeras, notre modèle a l'avantage de conserver un aspect hiérarchique tout en dissociant partiellement l'enveloppe du squelette grâce à une représentation basée sur les cylindres généralisés. Bien que des modèles similaires peuvent être trouvés dans le domaine de la segmentation de vaisseaux sanguins dans des images angiographiques, comme par exemple le travail de Yim et al ^[102], l'originalité de notre approche est de

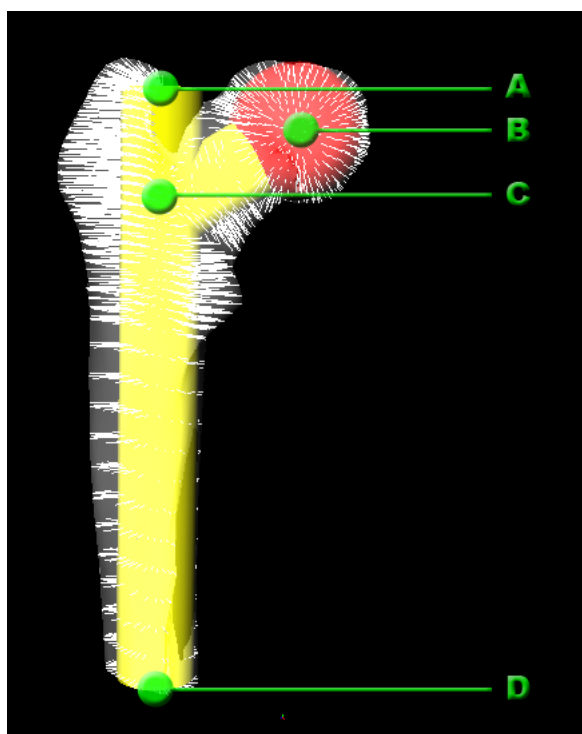


FIG. 6.7: Modèle déformable du fémur composé de 13 DDL (voir texte). Les liens de connexion de l'enveloppe avec le squelette sont représentés par les traits blancs.

proposer un paramètre supplémentaire de déformation qui est l'homothétie de l'enveloppe par rapport aux différentes primitives (voir section 6.4.3.2).

Notre modèle a partiellement été introduit au chapitre 5 pour la localisation automatique des parties anatomiques du fémur. En résumé, il est composé d'un cylindre représentant l'axe diaphysaire, un autre l'axe du col du fémur et une sphère pour la tête fémorale. Afin de limiter au maximum le nombre total de DDL, les courbe γ du squelette sont simplifiées par des axes rectilignes. Les points composant le maillage de l'enveloppe sont représentés dans un système de coordonnées cylindrique généralisé décrit à la section 6.4.1. Au total, le modèle possède 13 DDL (voir figure 6.7) que l'on peut exprimer au moyen des paramètres suivants :

- L'axe diaphysaire est défini par le point antérieur de l'axe (point D sur la figure 6.7) et, pour la partie postérieure, par le tubercule (point A). Ces deux premiers points représentent 6 DDL ;
- Le centre de la tête fémorale est approximé par une sphère (ou par une

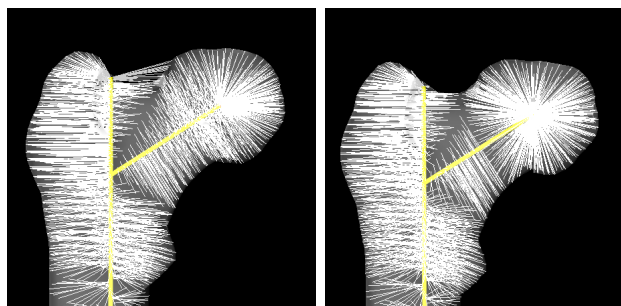


FIG. 6.8: Association des points de l'enveloppe au niveau de la tête fémorale dans le cas, à gauche, des cylindres généralisés et, à droite, dans le cas d'un rayon d'attraction autour du centre de la tête fémorale. Sur la figure de droite, on remarque une nette amélioration de la connexion de l'enveloppe au niveau de la fossette digitale.

calotte selon la nomenclature de Lazarus ^[59] - point B sur la figure).

Ce point supplémentaire augmente les DDL à 9 ;

- L'axe du col du fémur est, d'un côté, centré sur la tête fémorale et, de l'autre coupe l'axe diaphysaire au point C. Aussi, il ne rajoute qu'un seul DDL supplémentaire ;
- Les trois derniers DDL sont les diamètres des trois coniques (rayon de la tête fémorale, de la diaphyse et du col du fémur).

L'attachement des points de l'enveloppe au squelette se fait, dans un premier temps, en associant chaque point à la conique la plus proche. Pour prendre en compte le mieux possible la sphéricité de tête fémorale, nous utilisons une règle d'association particulière quant-à cette primitive. Nous définissons une région sphérique autour de la tête qui lui associe automatiquement les points situés à l'intérieur. Pratiquement le rayon de cette région est : $r_{region} = \frac{3}{2}r_{tete\ femorale}$. La figure 6.8 présente l'association avec et sans cette règle particulière.

Dans un deuxième temps, le point d'attache est, dans le cas d'une sphère, son centre et, dans le cas d'un cylindre, il se base sur la représentation en coordonnées cylindriques décrite précédemment. L'attachement des points de l'enveloppe sur des coniques (sphères, cylindres) plutôt que directement sur le squelette évite de nombreuses erreurs de connexion, principalement dans les régions proches des jointures des coniques.

6.4.3 Opérations de déformation

Le modèle comporte deux types de déformations possibles : le déplacement d'un point de contrôle et l'homothétie d'une conique (modification de son rayon).

6.4.3.1 Translation d'un point de contrôle

Dans le cas du cylindre, le déplacement d'un point de contrôle peut modifier sa position et sa longueur sans, pour autant, changer les coordonnées des points de l'enveloppe. Il s'agit d'une propriété des coordonnées cylindriques généralisées introduite à la section 6.4.1. Pour une sphère, la translation de son centre génère un déplacement et une rotation des points de l'enveloppe. La rotation a lieu si la sphère (ou calotte) est connectée au référentiel mobile d'un cylindre. Par exemple, dans le cas du fémur, la tête est connectée au col. Comme le système est articulé, un déplacement du centre de la tête fémorale va modifier l'axe du col (voir la figure 6.9 centrale). Au même titre que pour les cylindres, les coordonnées de l'enveloppe restent inchangées en coordonnées sphériques.

6.4.3.2 Homothétie de la conique

L'homothétie de la conique revient à modifier son rayon. Il en résulte une modification de la distance des points de l'enveloppe par rapport à leur point d'attache (voir la figure 6.9 droite). Contrairement à la translation, cette transformation modifie le paramètre d'amplitude des points dans leurs systèmes de coordonnées respectifs.

L'avantage de ce paramètre est, par exemple, de pouvoir changer le rayon de la tête fémorale sans modifier le reste du fémur. Chose qui ne serait pas possible d'effectuer avec seulement un paramètre d'homothétie global.

6.5 Validation du modèle

Nous avons adapté notre modèle déformable pour que sa transformation puisse être optimisée au moyen de l'algorithme du simplex-ICP (voir le chapitre 4). Nous avons implémenté deux variantes de déformations : la première

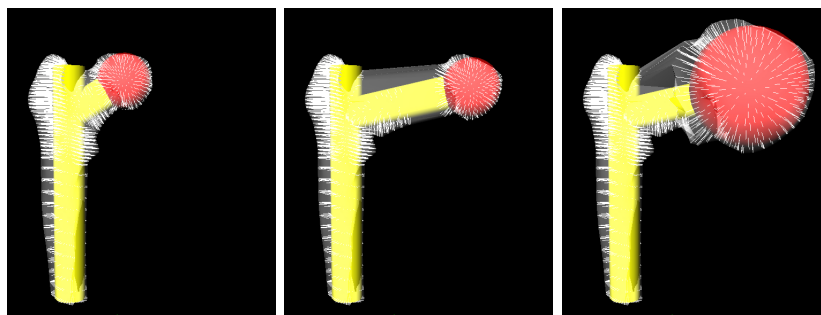


FIG. 6.9: A gauche, modèle proximal du fémur. Au milieu, exemple de déplacement du centre de la tête fémorale, à droite, modification du rayon de la sphère. Sur ces deux dernières figures, on remarque l'inconsistance du maillage au niveau du col. Il s'agit d'un artefact dû au fait que le nombre de points à cet endroit est sous-échantillonné.

comporte 10 DDL et optimise uniquement la position des points de contrôle du modèle, c'est-à-dire le centre de la tête fémorale, le point proximal et distal de l'axe de la diaphyse ainsi que la position du point de jonction du col du fémur. La seconde variante complète les paramètres de la première en y ajoutant l'optimisation des rayons des trois coniques, pour un total de 13 DDL (voir la figure 6.10). La figure 6.11 reprend l'exemple introduit à la figure 4.8 en utilisant les transformations suivantes : transformation rigide, modèle déformable avec 10 DDL et 13 DDL, approche hiérarchique pour les deux variantes précédentes et méthode quadratique hiérarchique avec au maximum 30 DDL (voir la section 4.2.2.2 pour plus d'informations relatives à ce modèle).

Ces résultats confirment qu'une approche hiérarchique, c'est-à-dire en ajoutant progressivement les DDL durant l'optimisation, est indispensable pour une convergence des jeux de données même lors d'une approche par modèle déformable. Cette dernière approche dans sa variante hiérarchique (max. 13 DDL) présente des résultats environ trois fois plus efficaces que la méthode de recalage global quadratique hiérarchique (max. 30 DDL). Aussi, l'ajout de contraintes spécifiques à un modèle de déformation particulier permet d'améliorer l'appariement tout en réduisant le nombre de DDL. A titre de comparaison, pour obtenir les mêmes résultats de recalage avec une transformation de type RBF-TPS "progressif" (voir la section 4.3.2), il est nécessaire de générer un modèle comportant au moins 220 DDL, mettant ainsi en évidence l'inefficacité de cette approche dans le cadre de notre problématique.

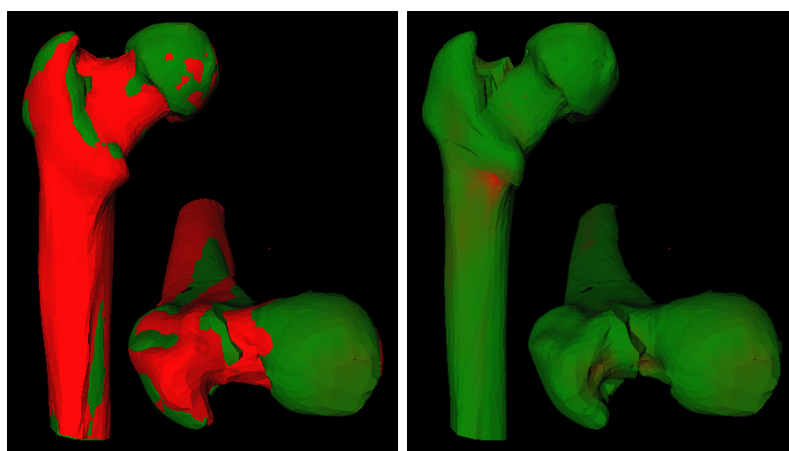


FIG. 6.10: Résultat d'appariement après convergence de l'algorithme simplex-ICP. A gauche, les deux jeux de données mélangés. A droite, codage de la distance du point le plus proche du vert (proche) au rouge (plus éloigné). On remarque une zone rouge sous le petit trochanter car le modèle déformable utilisé est trop simplifié pour prendre en compte précisément la position de cette partie anatomique.

6.6 Recalage non-itératif

Jusqu'à présent, nous avons décrit des variantes de l'algorithme ICP et constaté la complexité du processus d'optimisation ainsi que son manque de fiabilité lorsqu'on utilise une solution initiale inadéquate. Cette phase nécessite d'effectuer un alignement ou une pré-registation, par exemple au moyen d'un recalage point à point de Horn (décrit à la section 3.2). En pratique, le chirurgien définit trois appariements, c'est-à-dire 9 DDL. Pourquoi donc ne pas introduire les paramètres de notre modèle déformable à cette étape, d'autant plus que, comme nous l'avons vu au chapitre 5, ceux-ci peuvent être trouvés automatiquement par des rayons de recherche. Le tableau 6.1 présente l'appariement, par la méthode de correspondance point à triangle (voir la section 3.3.4.3), de notre modèle déformable (13 DDL) créé à partir du jeu de données du patient 2 et recalé dans les autres jeux de données grâce aux paramètres anatomiques extraits par rayons de recherche.

Si on se réfère à la figure 6.11, la précision du recalage par modèle déformable non itératif entre le patient 2 et le patient 1 est comparable à un recalage rigide par la méthode du simplexe ICP après 25 itérations. Notre méthode est par contre 10 fois moins précise que l'approche quadratique hiérarchique à 30 DDL après 130 itérations.

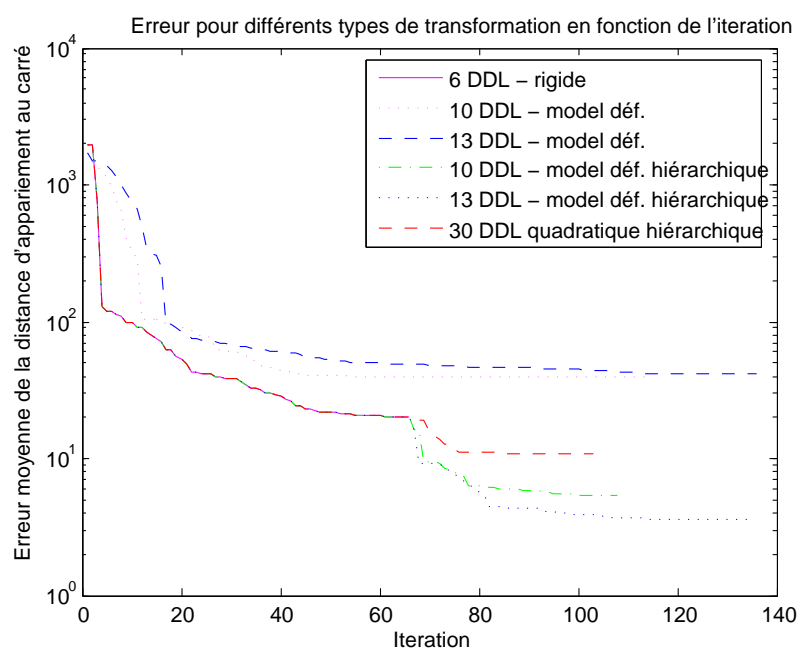


FIG. 6.11: Exemple de recalage non rigide de type simplex-ICP avec un appariement point à triangles en utilisant différents types de transformations (rigide - 6 DDL, modèle déformable avec 10 DDL et 13 DDL, approche hiérarchique pour les deux variantes précédentes et, méthode quadratique hiérarchique - 30 DDL). L'axe des erreurs présente la moyenne quadratique de la distance d'appariement.

| Jeu de données | 1 | 3 | 4 | 7 | moyenne |
|-----------------|-------|-------|------|------|-------------|
| Moyenne [mm] | 5,33 | 4,59 | 3,45 | 3,74 | 4,28 |
| Ecart type [mm] | 12,57 | 10,68 | 6,11 | 8,10 | 9,37 |

TAB. 6.1: Moyenne et écart type de l'appariement (par la méthode de correspondance point à triangle) de notre modèle déformable créé à partir du jeu de données du patient 2 et recalé pour correspondre aux autres jeux de données

6.7 Conclusion

Dans ce chapitre, nous avons tout d'abord présenté un modèle déformable de faible complexité : les modèles articulés qui sont principalement utilisés pour la reconnaissance de pose de personnes dans les séquences vidéo. Ceux-ci présentent une base intéressante de simplification d'un modèle déformable mais restreignent considérablement le niveau de précision du contour de l'objet. Ensuite, différentes méthodes basées sur les squelettes et les cylindres généralisés (GC) ont été décrites. Bien que beaucoup plus simple que la majorité des modèles déformables utilisés dans les applications médicales, ils présentent un nombre considérable de DDL comparativement aux modèles articulés.

En prenant en compte ces considérations, nous avons introduit un modèle déformable proche du modèle articulé mais présentant certaines particularités des GC comme la fonction d'attachement à l'enveloppe et la technique de déformation par rapport au squelette. Nous avons ensuite étendu la méthode des GC pour présenter une méthode de conique généralisée dont nous avons adjoint la possibilité de réaliser une mise à l'échelle globale des points attachés à la même primitive. Notre modèle final du fémur comporte 10 DDL pour la représentation la plus simple ou 13 DDL pour le modèle complet. L'utilisation de ce modèle simplifié à l'extrême et proche de ceux utilisés dans le domaine de l'animation de synthèse constitue la contribution principale de ce travail.

Ce modèle a été ensuite comparé aux différentes techniques de transformation exposées jusqu'à présent au moyen de l'algorithme simplexe-ICP, l'expérience consistant à déformer un modèle de fémur pour le faire correspondre à l'un des huit jeux de données segmentés manuellement. Conformément à nos attentes, l'ajout de contraintes liées à la forme permet de réduire considérablement le nombre de DDL par rapport à un modèle de transformation générique. A titre de comparaison, notre modèle déformable présente des ré-

sultats 3 fois supérieurs à l'approche quadratique hiérarchique (30 DDL) qui était la technique de recalage la plus précise décrite jusqu'à présent. De plus, notre modèle peut être comparé à une transformation RBF-TPS "progressif" possédant 220 DDL.

Bien que ces résultats soient excellents, la force de notre modèle déformable n'est pas exclusivement liée à la suppression du processus itératif présent dans l'approche ICP. La nature de notre modèle permet des transformations guidées par le placement de paramètres anatomiques "humainement compréhensibles" et automatisables, par exemple, par la méthode des lancers de rayons présentée au chapitre précédent. Aussi, contrairement à ICP qui nécessite une phase d'alignement initial, notre méthode de recalage par modèles déformable est effectuée instantanément en une seule passe (pour autant que les paramètres anatomiques du patient soient connus). A titre d'exemple, dans le cas du fémur, notre méthode est comparable à une transformation rigide optimisée par ICP après 25 itérations, c'est-à-dire, qu'elle présente une erreur moyenne de 4,3 mm avec un écart type de 9,4 mm et ceci, sans analyse du contraste des images scanner durant le recalage !

Fondamentalement, notre méthode de recalage par modèle déformable n'est ni un recalage grossier, ni un recalage fin mais une sorte d'entre-deux. On peut imaginer que pour certaines applications, la précision obtenue durant cette étape soit suffisante mais, dans la majorité des tâches de segmentation, elle est insuffisante. Une phase supplémentaire de raffinement et donc nécessaire. Une telle méthode est présentée au prochain chapitre.

CHAPITRE VII

SEGMENTATION FINE PAR PROGRAMMATION DYNAMIQUE

Nous avons vu au chapitre 5 une méthode de localisation précise de paramètres anatomiques du fémur. Le chapitre 6 présente un modèle déformable pouvant se recalibrer grossièrement grâce à ces informations. Dans ce chapitre, nous abordons la possibilité de déterminer précisément le bord du fémur dans les coupes CT au moyen d'une méthode de segmentation de contours. La particularité de cette technique est de ne nécessiter aucun processus d'optimisation itératif, contrairement aux techniques usuellement employées comme les *snakes* ou les rayons actifs.

7.1 Introduction

La segmentation osseuse à partir d'images CT ^[63] est utilisée dans de nombreux domaines comme l'analyse cinématique des articulations ^[86], la planification préopératoire et la navigation pour la prothèse totale de hanche ^[95], ^[104] et celle du genou ^[93] ou d'autres applications cliniques ^[72] ^[35].

La segmentation osseuse implique dans un premier temps la séparation des images en régions, dont les contours représentent les frontières entre les os et les autres tissus. Dans une seconde étape, il s'agit de déterminer si l'os fait partie de la hanche ou du fémur. Lorsqu'on essaye d'automatiser ce processus surgit un problème de classification difficile à résoudre avec des causes aussi variées que ^[107] :

- le manque d'uniformité osseuse avec des contours flous et diffus (ostéoporose),
- la présence de vaisseaux sanguins calcifiés ou contenant du produit de contraste radiologique (CTA),

- des régions articulaires sans cartilage (arthrose) présentant, dans certains cas, la tête fémorale et l’acetabulum soudés,
- des artefacts liés à l’imageur, dus par exemple à la proximité de métal (prothèses, matériel d’ostéosynthèse, piercings). On constate néanmoins une amélioration logicielle des scanners pour corriger ces distortions et,
- des effets indésirables générés par un mauvais échantillonnage des coupes.

Les paragraphes suivants présentent les techniques de segmentation osseuses les plus couramment employées avant de présenter notre approche.

7.2 Segmentation par seuillage

Traditionnellement, la segmentation osseuse dans des images CT est réalisée par seuillage (*thresholding*) [76] [53], suivie par une procédure d’assemblage ou de connexion entre les régions segmentées [4]. La segmentation par seuillage est une technique qui classe tous les pixels compris dans un intervalle d’intensité comme faisant partie du tissu recherché, les autres pixels étant rejetés (*outliers*). Cette méthode est adaptée pour segmenter des tissus osseux car leur densité, c’est-à-dire l’intensité dans les images CT, est plus élevée que les tissus environnants. La segmentation par seuillage peut présenter des dysfonctionnements dans les régions articulaires, d’autant plus importants lorsque le patient présente des pathologies cartilagineuses comme une coxarthrose.

7.3 Segmentation par snakes

Une sous-famille des modèles déformables appelés *snakes* [50] [57] sont des courbes ou des surfaces qui évoluent dans un espace 2-D ou 3-D pour délimiter au mieux une structure anatomique ou pathologique. L’évolution du modèle est guidée par la minimisation de deux facteurs : l’énergie interne et l’énergie externe. Le premier définit la régularité géométrique du modèle, en s’appuyant par exemple sur le calcul de sa courbure locale et, le second, mesure une propriété de l’intensité lumineuse le long des points de l’image traversés par le modèle déformable (comme par exemple la norme du gradient de l’intensité lumineuse) pour favoriser l’attraction du modèle vers des points de fort contraste. Ces méthodes sont très efficaces lorsque l’utilisateur peut initialiser le modèle de manière approximative autour de

la région d'intérêt. Le modèle déformable évolue alors depuis cette solution initiale grossière, pour itérativement, au cours d'un processus d'optimisation, délimiter avec précision et de manière automatique la frontière de la région recherchée. Cette approche est généralement beaucoup plus robuste que le simple seuillage de l'intensité lumineuse mais, par la nature de la minimisation de l'énergie (descente de gradient), elle est sujette aux minima locaux.

Les rayons actifs (*active rays*), initialement proposés par Denzler et Niemann [27] [28] sont une représentation particulière des *snakes*. Au lieu de définir les contours dans un espace Euclidien, ils sont représentés de manière radiale/polaire par une succession de rayons, ce qui permet de surmonter certaines limitations inhérentes aux modèles déformables. Cette astuce de représentation interdit les croisements des éléments du contour durant leur évolution. Un autre avantage est de réduire le problème de minimisation de l'énergie dans un espace de recherche unidimensionnel. Finalement, comparée aux *snakes* conventionnels, cette approche est beaucoup moins sensible à l'initialisation du modèle.

Inspirés par la méthode des rayons actifs, Chen et al [19] ont développé, dans le domaine du suivi du visage dans des séquences vidéos, une nouvelle représentation de l'énergie interne et externe des contours actifs. Celle-ci encode la régularité du contour et contrôle sa forme et sa position par une formulation discrète de programmation dynamique (PD). La PD est une méthode d'optimisation basée sur le principe d'optimalité. Si les solutions d'un problème sont une combinaison de solutions de sous-problèmes alors on peut s'attendre à ce que la solution optimale du problème soit une combinaison des meilleures solutions des sous-problèmes respectifs. Contrairement aux *snakes* conventionnels, la résolution du problème des contours actifs par la PD converge vers l'optimum en un nombre fini d'itérations car sa fonction de cout décroît de manière monotone [3] [4].

Dans ce chapitre, nous présentons une adaptation de l'approche de Chen et al. [19] à la problématique de la segmentation fine de la tête fémorale dans les images CT. Au lieu d'utiliser une ellipse comme forme initiale pour le contour, nous utilisons soit la coupe adjacente que nous venons de segmenter, soit une coupe du modèle déformable calculé au chapitre précédent. L'énergie externe est, dès lors, représentée par cet à-priori ainsi que la direction du gradient de l'image. L'énergie interne, qui impose des contraintes de régularité et de continuité sur le contour, est très similaire à celle employée par Chen et al. La solution optimale, qui minimise les énergies globales, est

calculée au moyen de la PD.

Contrairement à la méthode de seuillage, cette approche n'utilise pas d'informations absolues d'intensité pour la segmentation et permet d'éviter les problèmes typiques de fusion ou fragmentation de régions. Comparée à l'approche de Chen et al., la représentation de l'énergie externe par la direction du gradient que nous utilisons s'avère plus efficace.

7.4 Rayons actifs

Denzler et Niemann [27] [28] proposent la description suivante pour les rayons actifs (voir figure 7.1 a) : Un rayon actif $\delta_m(\phi, \lambda)$ est défini sur une image plane $(x, y)^t$ comme une fonction monodimensionnelle dépendant les valeurs d'intensité $f(x, y)$ de l'image, pouvant être écrit comme un segment avec comme extrémité le point $m = (x_m, y_m)^t$ et une direction ϕ , soit par l'équation :

$$\delta_m(\phi, \lambda) = f(x_m + \lambda \cos(\phi), y_m + \lambda \sin(\phi)), 0 \leq \lambda \leq n_\phi$$

, où n_ϕ est la longueur maximale du rayon, valeur donnée par la taille de l'image. On représente un contour radial discret C_m par un point m situé à l'intérieur du contour et une fonction monodimensionnelle $r_m(\phi)$ ou $r_m^j(\phi)$, où j représente l'index de la coupe (voir figure 7.1 b). Cette représentation n'est valide que si tous les rayons de C_m coupent une seule fois le contour original (C). Il s'agit de la définition de la convexité, c'est-à-dire :

$$\forall m, m \text{ inside } C | \exists r_m(\phi)$$

Dans [69], nous proposons la définition suivante : C_m est dit pseudo-convexe si $r_m(\phi)$ existe pour au moins un point dans C :

$$\exists m, m \text{ inside } C | \exists r_m(\phi)$$

La représentation par rayons actifs n'est valide que si C_m est convexe ou pseudo-convexe. Exprimé différemment, si on donne une représentation discrète radiale d'un contour et un point m à l'intérieur de celui-ci, tous les

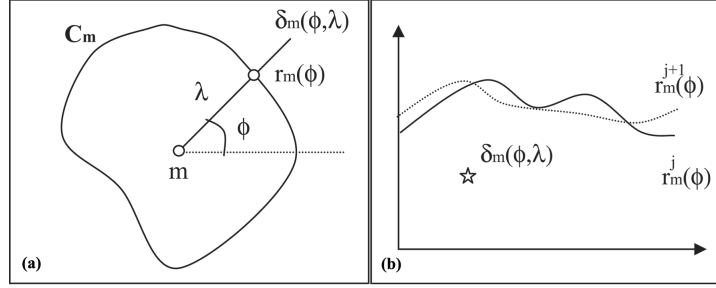


FIG. 7.1: (a) Représentation d'un rayon actif $\delta_m(\phi, \lambda)$, défini sur l'image $(x, y)^t$ comme une fonction unidimensionnelle dépendant des valeurs d'intensité $f(x, y)$ de l'image. $m = (x_m, y_m)^t$ est un point arbitraire situé dans le contour radial discret C_m . N'importe quel contour pseudo-convexe peut être représenté par une fonction discrète $r_m(\phi)$. (b) Représentation radiale d'un contour de la coupe j : $r_m^j(\phi)$ (ligne pleine) et de la coupe adjacente $j \pm 1$: $r_m^{j \pm 1}(\phi)$ (ligne trait-tillée) par rapport au même point m . L'intersection d'une ligne verticale avec les deux courbes représente pour un angle donné ϕ la distance entre les contours.

rayons vont couper le contour par exactement un point.

7.5 Représentation de l'énergie

Dans la représentation traditionnelle des *snakes*, l'énergie interne et externe doit être définie pour chaque point du contour. Dans notre cas de segmentation, l'énergie interne définit la contrainte de régularité (*smoothness*) du contour. Ce facteur permet de pénaliser les contours présentant des rayons de courbure trop faibles. Chen et al. proposent d'approximer cette énergie en utilisant la distance quadratique de deux rayons successifs. Leur formulation de l'énergie interne est la suivante :

$$E_{int}(r_m(\phi_i)) = \alpha \|r_m(\phi_i) - r_m(\phi_{i-1})\|^2$$

Où α est un paramètre de pondération. L'énergie externe comprend l'influence du gradient d'intensité de la coupe ainsi que l'approximation initiale du contour :

$$E_{ext}(r_m(\phi_i)) = E_{init}(r_m(\phi_i)) + E_{grad}(r_m(\phi_i))$$

Si le point m est situé proche du centre du contour, alors l'énergie du gradient peut être représentée par :

$$E_{grad}(r_m(\phi_i)) = \beta \cdot g\left(\frac{d}{d\lambda} \delta_m(\sigma, \lambda)\right)$$

Où β est un paramètre de pondération et $g()$ une fonction monotonique décroissante.

Lors du calcul de l'énergie initiale du contour (E_{init}), on constate que, plus la distance entre deux coupes consécutives est petite, plus la distance de Hausdorff entre ces deux contours respectifs est faible (voir la figure 7.1). Cette observation peut être approximée par :

$$E_{init}(r_m(\phi_i)) = \gamma \|r_m^j(\phi_i) - r_m^{j\pm 1}(\phi_i)\|$$

Où γ est un paramètre de pondération et j le numéro de la coupe (voir la figure 7.1 b). De manière similaire, la distance de Hausdorff, peut être utilisée entre l'approximation du contour par modèle déformable et le contour courant.

L'énergie totale du contour est dès lors donnée par l'expression :

$$E_{tot}(r_m(\phi_i)) = \int_0^{2\pi} E_{int}(r_m(\phi_i)) + E_{ext}(r_m(\phi_i)) d\phi$$

Le meilleur contour C_m est celui qui minimise l'énergie totale. Grâce à cette reformulation, le problème présente une solution par programmation dynamique.

7.6 Programmation dynamique

Cette section est partiellement tirée du cours donné par le professeur Rebaïne de l'université du Québec à Chicoutimi.

La programmation dynamique est un paradigme de conception qu'il est possible de voir comme une amélioration ou une adaptation de la méthode

"diviser et régner". Ce concept a été introduit par Bellman ^[9], dans la fin des années 50, pour résoudre typiquement des problèmes d'optimisation.

Pour la petite histoire, Bellman a choisi le terme programmation dynamique dans un souci de communication : son supérieur ne supportait ni le mot "recherche" ni celui de "mathématique". Alors il lui a semblé que les termes "programmation" et "dynamique" donnaient une apparence qui plairait à son supérieur. En réalité, le terme programmation signifiait à l'époque plus planification et ordonnancement que la programmation au sens qu'on lui donne de nos jours. En un mot, la programmation dynamique est un ensemble de règles que tout un chacun peut suivre pour résoudre un problème donné.

La programmation dynamique est similaire à la méthode diviser et régner en ce sens que, une solution d'un problème dépend des solutions précédentes obtenues des sous-problèmes. La différence significative entre ces deux méthodes est que la programmation dynamique permet aux sous-problèmes de se superposer. Autrement dit, un sous-problème peut être utilisé dans la solution de deux sous-problèmes différents. Tandis que l'approche diviser et régner crée des sous-problèmes qui sont complètement séparés et peuvent être résolus indépendamment l'un de l'autre. Une illustration de cette différence est montrée par la figure 7.2. Dans cette figure, le problème à résoudre est à la racine, et les descendants sont les sous-problèmes, plus faciles à résoudre. Les feuilles de ce graphe constituent des sous-problèmes dont la résolution est triviale. Dans la programmation dynamique, ces feuilles constituent souvent les données de l'algorithme. La différence fondamentale entre ces deux méthodes devient alors claire : les sous-problèmes dans la programmation dynamique peuvent être en interaction, alors dans la méthode diviser et régner, ils ne le sont pas. Une seconde différence entre ces deux méthodes est que la méthode diviser et régner est récursive, les calculs se font de haut en bas, tandis que la programmation dynamique est une méthode dont les calculs se font de bas en haut : on commence par résoudre les plus petits sous-problèmes. En combinant leur solution, on obtient les solutions des sous-problèmes de plus en plus grands.

La programmation dynamique est utilisée dans de nombreux domaines incluant la recherche de séquences chromosomiques dans l'ADN, les algorithmes d'intelligence artificielle des jeux d'échecs, dans certains types de chaînes de Markov, le calcul de courbes B-splines ainsi que pour la distance de Levenshtein. Pour plus d'informations relatives aux divers débouchés de cette technique de résolution de graphes, se référer à la page anglaise de

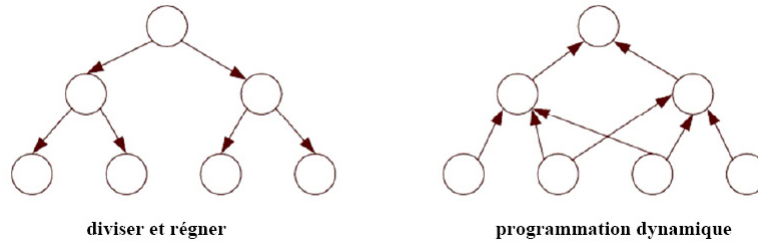


FIG. 7.2: Tiré du cours de Rebaïne. Différence entre l'algorithme diviser et régner et la programmation dynamique. L'avantage de la programmation dynamique est de permettre aux sous-problèmes de se superposer. Autrement dit, un sous-problème peut être utilisé dans la solution de deux sous-problèmes différents. Tandis que l'approche diviser et régner crée des sous-problèmes qui sont complètement séparés et peuvent être résolus indépendamment les uns des autres.

Wikipedia sous la rubrique *dynamic programming*.

7.7 Contribution - Optimisation de l'énergie par programmation dynamique

L'optimisation de l'énergie consiste à trouver le contour C_m qui minimise l'énergie totale, c'est-à-dire le facteur E_{tot} de la section 7.5, pour autant que les facteurs de pondération (α, β, γ) soient correctement paramétrés.

Dans le cas discret, l'énergie totale du système peut être représentée dans un tableau bidimensionnel $[\phi, \lambda]$ de dimension $N * M$ (voir figure 7.3). L'état initial est défini par la première colonne $\phi_0 = 0$:

$$S_0(\phi_0, \lambda_j) = E_{ext}(\phi_0, \lambda_j)$$

Cette énergie est ensuite propagée aux différentes étapes en utilisant la formule récursive suivante :

$$S_{k+1}(\phi_{k+1}, \lambda_j) = E_{ext}(\phi_{k+1}, \lambda_j) + \min_{\lambda_i \in [0, N]} \{S_k(\phi_k, \lambda_j) + E_{int}(\lambda_i, \lambda_j)\}$$

où la contrainte de régularité est $E_{int}(\lambda_i, \lambda_j) = \alpha \|\lambda_i - \lambda_j\|^2$. Le contour

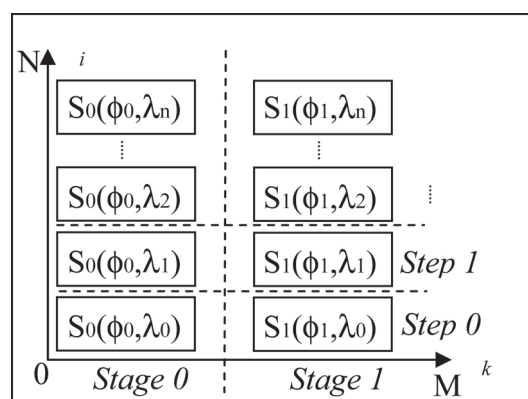


FIG. 7.3: Représentation interne durant l'optimisation par programmation dynamique. Une étape (*stage*) représente les différentes énergies externes pour un angle ϕ donné. Un état exprime l'énergie interne pour une certaine distance λ . L'énergie totale est propagée de l'étape 0 à l'étape M. La position du contour optimal peut ensuite être extrait en sens inverse, c'est-à-dire à partir du résultat de l'étape M jusqu'à l'étape 0.

optimal peut être déterminé en trouvant l'état (λ_x) présentant le résultat minimum de la dernière étape (ϕ_N) et d'ensuite revenir à l'étape précédente (ϕ_{N-1}) et ceci récursivement jusqu'à l'étape initiale (ϕ_0).

Cette méthode s'avère optimale uniquement pour des contours ouverts. Les contours fermés nécessitent l'emploi d'une contrainte supplémentaire, comme par exemple que le premier et le dernier point doivent être similaires. Dans ce cas, la programmation dynamique peut toujours être employée, pour autant que l'approche présentée soit quelque peu modifiée. Dans un premier temps, on doit fixer un premier point et trouver le meilleur contour qui se termine au même point. Après avoir testé tous les points initiaux possibles, on doit sélectionner celui qui minimise le plus l'énergie totale. La section suivante présente une approche heuristique permettant de sélectionner du premier coup le point initial.

7.7.1 Heuristique pour la sélection du point initial

Tester chaque point initial durant la phase de programmation dynamique réduit considérablement les performances de segmentation. Une façon de trouver ce point est de chercher le rayon qui possède une frontière franche entre la tête fémorale et les tissus environnants. Celui-ci est généralement

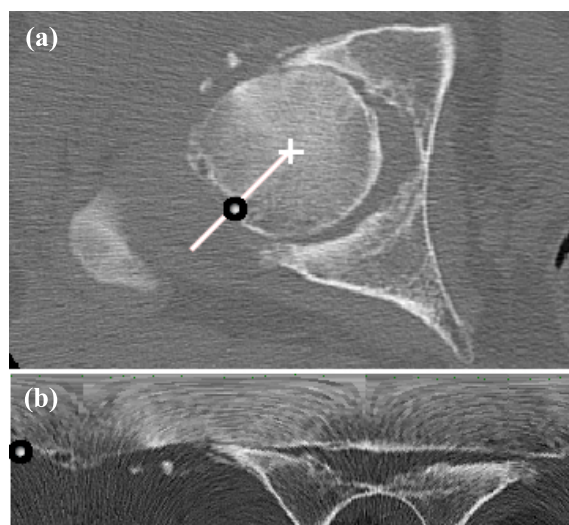


FIG. 7.4: (a) Coupe CT de la hanche. La croix blanche représente le centre du contour discret C_m , la ligne blanche est le rayon actif du premier étage de la programmation dynamique et le rond noir est la position initiale trouvée par l'algorithme de recherche du point initial. (b) Représentation radiale de la figure (a) par rapport au centre défini par la croix.

situé dans la partie externe du fémur, c'est-à-dire dans la direction du grand trochanter (voir le rayon de la figure 7.4). On peut définir pour chaque rayon le point ayant l'énergie externe minimale. Il s'agit de la frontière λ_b . Dès lors, le point initial est celui ayant le gradient le plus grand (ou similairement l'énergie du gradient - E_{grad} - la plus faible) dans l'intervalle $\lambda \in [0..\lambda_b]$. Cette heuristique s'avère très robuste et présente un taux de succès d'environ 95% sur toutes les coupes testées. La figure 7.4 montre un point initial calculé par cette méthode.

7.7.2 Paramètres

Notre méthode de segmentation par programmation dynamique a été implémentée dans le logiciel *Segmentator*. Il permet la segmentation 2D (semi-)automatique ou interactive de coupes scanners ainsi que la reconstruction tridimensionnelle des organes. Cinq paramètres de contrôle pour la segmentation du fémur peuvent être contrôlés par l'interface utilisateur (voir la figure 7.13).

1. $\omega \in [0..5]$ définit le niveau de filtrage anisotropique des coupes,

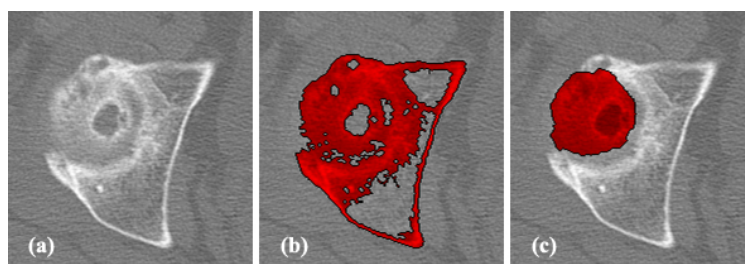


FIG. 7.5: Coupe CT de la hanche avec forte dégénération présentant de nombreux points de contact entre l'acetabulum et la tête fémorale (série 2). (a) Image originale, (b) Résultats en appliquant une segmentation par seuil (c) Algorithme présenté avec les paramètres ($\alpha = 2, \beta = 0,5, \gamma = 24, \delta = -4, \omega = 2$).

2. $\alpha \in [0..5]$ paramètre de pondération de l'énergie externe (E_{ext}),
3. $\beta \in [0..5]$ paramètre de pondération de l'énergie du gradient (E_{grad}),
4. $\gamma \in [0..30]$ paramètre de pondération de l'énergie interne (E_{int}),
5. $\delta \in [-5..5]$ paramètre de correction de la taille du contour (E_{int})

Ce dernier paramètre est utilisé uniquement lorsque le contour est approximé par la coupe adjacente. Il permet une compensation du changement de taille de la tête fémorale au moyen d'une translation sur l'axe λ du contour de référence dans la représentation radiale ou similairement un changement de rayon dans sa représentation cartésienne.

L'énergie du gradient est calculée en utilisant une valeur de pixel normalisés sur l'intervalle $[0..1]$ avec la fonction $g()$ définie comme suit :

$$g(x) = \delta_m(\phi, x) - \delta_m(\phi, x - 1)$$

$$g(x) = 0, \forall g(x) < 0$$

Les figures 7.5 et 7.6 présentent les résultats de segmentation en utilisant *Segmentator*. La figure 7.7 montre l'influence du paramètre δ sur ces résultats.

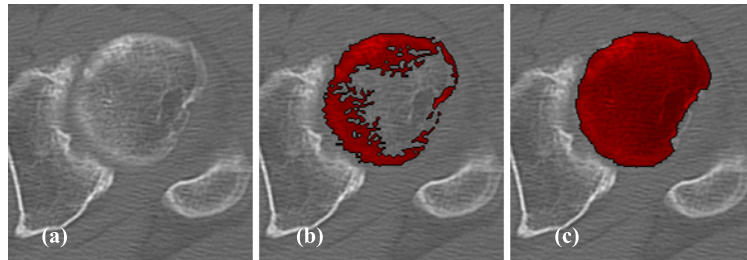


FIG. 7.6: Coupe CT de la hanche présentant une dégénérescence importante (série 3). (a) Image originale, (b) Résultats en appliquant une segmentation par seuil (c) Algorithme présenté avec les paramètres ($\alpha = 2, \beta = 0,5, \gamma = 24, \delta = -4, \omega = 2$).

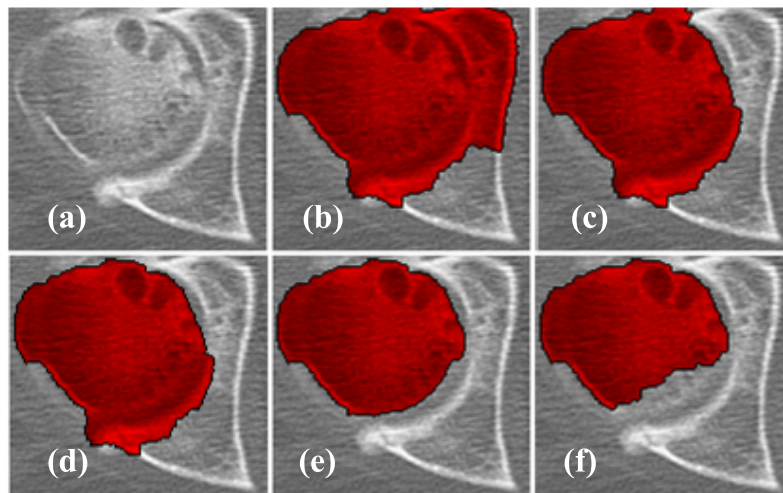


FIG. 7.7: Coupe CT de la hanche présentant une arthrose moyenne avec de multiples kystes fémoraux (série 4). (a) Image originale, influence du paramètre γ pour (b) $\gamma = 5$ (c) $\gamma = 1$ (d) $\gamma = 0$ (e) $\gamma = -2$ (f) $\gamma = -3$.

7.8 Validation

Quatre séries CT de hanches dégénérées, toutes nécessitant une arthroplastie complète ont été utilisées pour valider cette méthode de segmentation. La première série présente un homme avec une arthrose modérée de grade 3, la seconde (voir la figure 7.5) montre une hanche fortement dégénérée avec plusieurs contacts entre l’acetabulum et la tête fémorale, la troisième (voir la figure 7.6) illustre un cas standard d’articulation dégénérée et la dernière présente (voir la figure 7.7) une arthrose moyenne avec plusieurs kystes fémoraux.

7.9 Résultats

Pour comparer la précision de notre approche, nous avons analysé les quatre séries susmentionnées en utilisant le package Matlab *Active Shape Models with Multi-Resolution (ASMSnake)* [44]. Ce module est disponible gratuitement et permet une grande flexibilité dans les réglages. La coupe précédemment segmentée a été utilisée comme état initial de l’algorithme. Les paramètres du contour actif ont dû être réglés individuellement pour chaque coupe afin de garantir leur convergence. Les résultats ainsi obtenus sont très proches de la technique proposée (voir la figure 7.8). Nous avons cependant remarqué que les *snakes* manquant de stabilité, sont lents et très sensibles aux paramètres de pondération. Par exemple, la segmentation par *ASMSnake* doit être manuellement arrêtée lorsque l’algorithme diverge ou si une position initiale invalide est proposée. L’avantage principal de notre approche est qu’une solution initiale produit une solution unique pour un set de paramètres donné. Chaque résultat peut, ainsi, être validé soit par l’utilisateur, soit par un algorithme de décision automatique.

Les figures 7.5 et 7.6 présentent des résultats de segmentation avec un simple seuillage. La première figure met en évidence un phénomène de pont (*bridge*) dû à la sous-segmentation qui solidarise le bassin et le fémur. Le problème inverse, celui de l’émiettage (*crumb*) dû à une sur-segmentation est observable sur la seconde figure. Dans ce cas, le manque de corticalité fémorale génère une segmentation de l’os en plusieurs régions. En règle générale, la méthode de segmentation par seuillage est inefficace lorsque les coupes sont fines ou l’os dégénéré.

Lors de l’utilisation notre technique de segmentation sur la tête fémorale

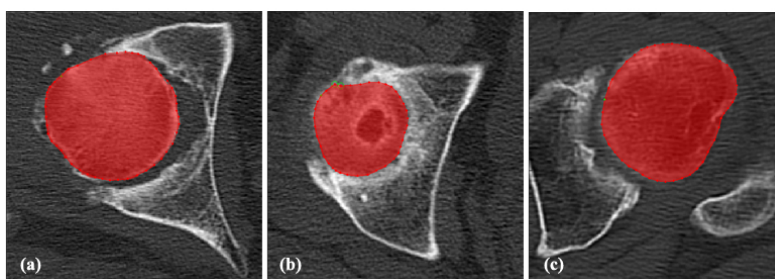


FIG. 7.8: Segmentation des séries CT présentées au moyen du package Matlab *Active Shape Models with Multi-Resolution* [44]. (a) Première série avec les paramètres optimaux après 50 itérations (b) Seconde série avec le paramètre de gonflage nul après 190 itérations (c) Troisième série avec le paramètre de gonflage nul et un seuil de 0,2 après 150 itérations. Noter que les paramètres du *snake* ont été réglés différemment pour chaque coupe afin de garantir la convergence du contour.

(avec l'initialisation du contour par la coupe adjacente), une image sur trois nécessite, en moyenne, une intervention manuelle. En cas d'échec (voir la figure 7.9), généralement quand les coupes sont très fines ou qu'elles présentent une forte ostéoporose, une modification interactive des paramètres α , β , γ , δ , ω permet de faire converger le contour actif dans presque tous les cas.

Notre méthode présente néanmoins quelques limitations si elle est appliquée sur l'ensemble du fémur. Tout d'abord, la segmentation par programmation dynamique ne permet pas de segmenter plusieurs régions sur la même coupe, ce qui est le cas dans la région du grand trochanter, où le trochanter et la tête fémorale sont tous deux visibles. Ensuite, elle ne convient pas aux coupes présentant un contour fortement concave comme dans la région du petit trochanter qui n'est segmenté que partiellement par la PD (voir la figure 7.9). Finalement, ce type de segmentation est moins bien adapté qu'une segmentation par seuillage pour analyser les coupes antérieures du corps fémoral qui sont corticales et fortement contrastées.

Ces limitations peuvent être levées si on combine les techniques de programmation dynamique et de seuillage en fonction de la région à segmenter. On remarque par exemple que l'utilisation de la PD est optimale pour la tête fémorale et le seuillage pour les zones corticales le long de l'axe intra médullaire. Combiner ces méthodes de segmentation nécessite néanmoins de savoir quelle région du fémur est présente sur chaque coupe, chose possible si on utilise le modèle déformable présenté au chapitre précédent (6). Dans ce cas,

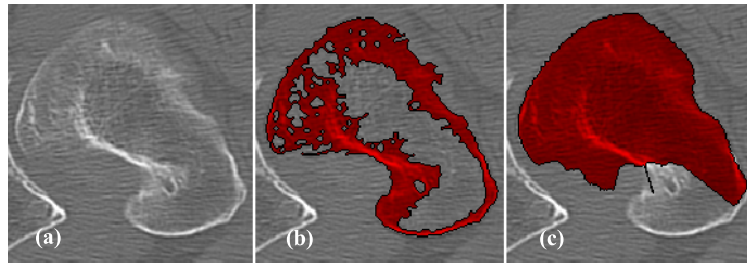


FIG. 7.9: Exemple de coupe où l'algorithme de programmation dynamique peine à converger. (a) Coupe CT de la troisième série, (b) Segmentation manuelle par seuillage, (c) Algorithme présenté avec les paramètres ($\alpha = 2, \beta = 1, \gamma = 21, \delta = 0, \omega = 2$). Les problèmes de convergence sont liés à la faible densité osseuse à gauche (ostéoporose) ou, dans la partie inférieure, où le contour du fémur ne satisfait plus la pseudo-convexité.

on peut prédéfinir des paramètres de segmentation en fonction de la coupe à segmenter, comme par exemple, commencer par segmenter la région de la tête fémorale par PD et ensuite continuer par segmenter la région de l'axe intra médullaire par seuillage. Cette astuce permet de résoudre la majorité des problèmes de segmentation précités dans les régions des trochanters.

7.10 Modèle déformable et programmation dynamique

L'utilisation des modèles déformables présente une solution élégante pour améliorer l'efficacité de la PD. Elle permet, d'une part, d'améliorer l'initialisation du contour et, d'autre part, de régler efficacement les paramètres de réglage du contour actif (pour que l'on puisse, par exemple, définir des paramètres différents en fonction de la région à segmenter). On observe, cependant, que le recalage initial réalisé uniquement au moyen d'un modèle déformable n'est pas assez précis pour une convergence optimale de la PD (voir la figure 7.11 - c). Comme ce problème ne se situe pas au niveau de la forme du modèle déformable mais de sa pose, un recalage rigide intermédiaire (voir la figure 7.11 - d) permet de repositionner correctement le modèle. Ce recalage s'effectue entre le modèle déformable et un nuage de point obtenu par une simple segmentation par seuillage des tissus osseux (voir la figure 7.11 - b). Le niveau de segmentation est obtenu au moyen d'une analyse de l'histogramme des coupes CT par une méthode de *Gaussian Mixture* (voir la section 5.4.4.1). Comme l'initialisation du recalage rigide est extrêmement proche de la solution optimale, le processus d'optimisation par LM-ICP (voir

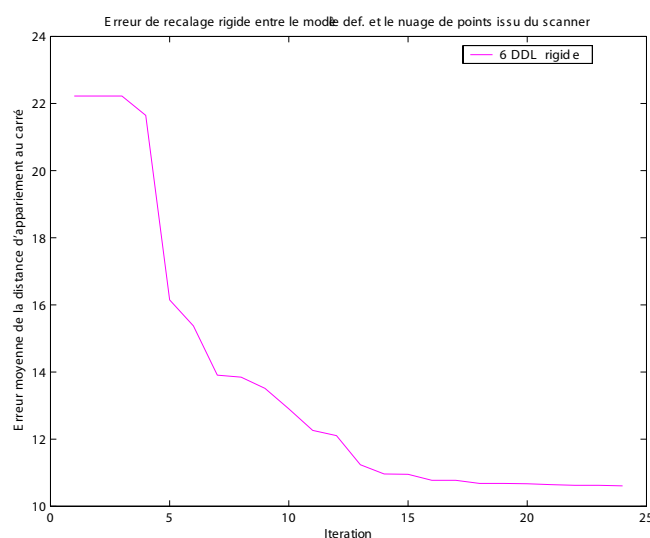


FIG. 7.10: Recalage entre le modèle déformable et le nuage de points issu de la segmentation de la série CT par seuillage. Les résultats sont exprimés en erreur quadratique en fonction des itérations. Comme l'initialisation du modèle est très proche de la position optimale, la convergence de la transformation par LM-ICP est très rapide (<1 seconde par itération) et converge en un nombre restreint d'itérations (environ 15 dans l'exemple).

le chapitre 3) est succinct et converge, à coup sûr, vers un minimum global. La figure 7.10 présente les résultats d'un tel exemple de recalage en fonction du nombre d'itérations. D'autres types de recalages présentant plus de degrés de libertés ou l'approche hiérarchique ont également été testés mais n'ont pas été retenus car ils ne présentaient pas de résultats substantiellement meilleurs qu'une transformation rigide.

La figure 7.12 présente les résultats de l'étape de recalage rigide en 3D.

Une fois la segmentation par programmation dynamique effectuée, il reste à segmenter les zones situées le long de l'axe intra médullaire. La segmentation de ces régions par seuillage est très efficace et a déjà été réalisée lors du recalage rigide. La figure 7.13 présente la segmentation compétente résultante ainsi que l'interface de *Segmentator*. La figure 7.14 présente la correspondance entre la segmentation manuelle et celle obtenue par la technique présentée dans ce travail. L'erreur moyenne de correspondance de cet exemple est de 1,36 mm (ou 4,32 pour l'erreur quadratique) avec une variance de 2,48 mm. L'erreur principale de segmentation se situe au niveau de la tête

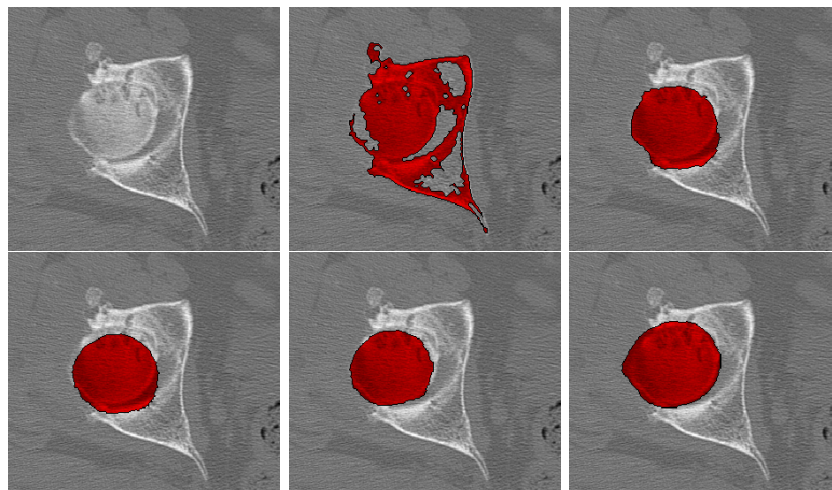


FIG. 7.11: Vue d'une coupe et de sa segmentation à différentes étapes. De haut en bas et de gauche à droite : (a) Coupe CT de la série 1, (b) Segmentation par seuillage des tissus osseux, (c) Recalage par modèle déformable, (d) Après recalage rigide du modèle déformable avec le nuage de points construit sur la base de la segmentation par seuillage (voir la figure b), (e) Résultat final après segmentation par programmation dynamique, (f) *Ground truth*.

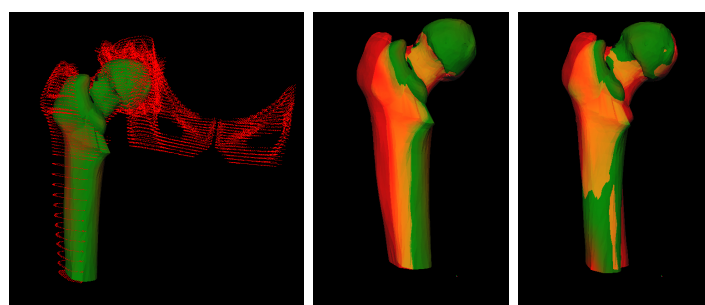


FIG. 7.12: Vue en 3D durant l'étapes de recalage rigide. De gauche à droite : (a) Position initiale du modèle déformable avec le nuage de points construit au moyen d'une segmentation par seuillage, (b) Même vue avec le fémur segmenté manuellement (*Ground Truth*) (c) Résultat après recalage rigide.

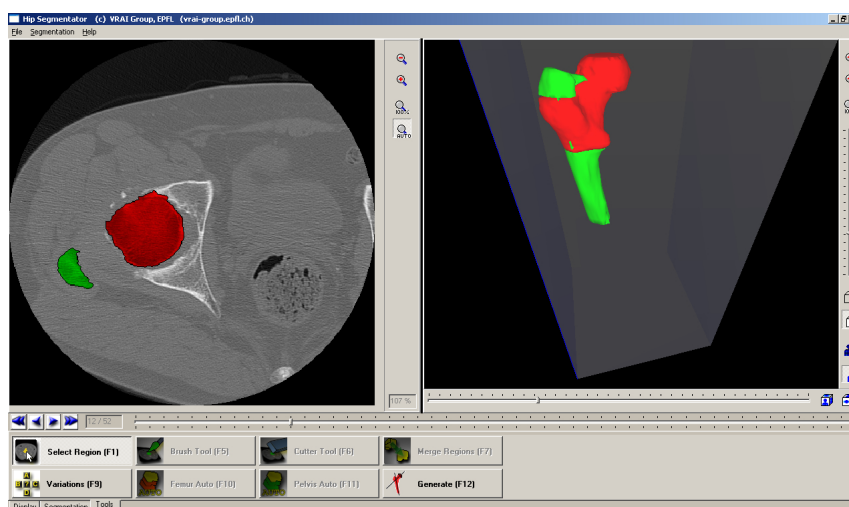


FIG. 7.13: Vue de l'interface de segmentation (*Segmentator*). En rouge, os segmenté par programmation dynamique, en vert, os segmenté par seuillage.

fémorale.

7.11 Conclusion

La segmentation par programmation dynamique est une alternative élégante aux méthodes de contours actifs. Il s'agit de la dernière contribution théorique de ce travail. Contrairement aux *snakes*, l'approche par PD a l'avantage de ne pas nécessiter une procédure itérative qui conduit généralement à un contour sous-optimal mais est, d'un autre côté, limitée à des contours pseudo-convexes. Une grande limitation des contours actifs, quelle que soit la méthode de résolution utilisée, est leur sensibilité aux paramètres de pondération de l'énergie ainsi que leur initialisation.

Combiné à un recalage par modèle déformable, comme celui décrit au chapitre 6, suivi d'un recalage rigide, il est possible d'obtenir des contours initiaux de bien meilleure qualité et plus fiables qu'en utilisant celui de la coupe adjacente. Un autre avantage lié aux modèles déformables est de connaître la région en cours de segmentation et pouvoir ainsi régler, pour chaque contour, les paramètres optimaux de pondération de la fonction d'énergie.

Mélanger la segmentation par PD et la segmentation par seuillage dans des régions spécifiques le long de l'axe intra médullaire permet d'obtenir une segmentation automatique présentant une erreur de l'ordre de 1,4 mm avec

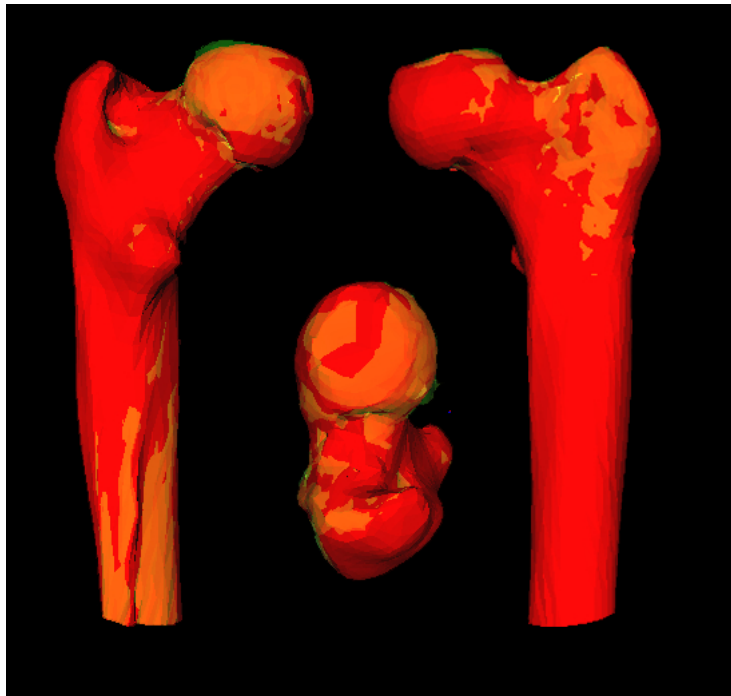


FIG. 7.14: Correspondance entre la segmentation manuelle (*Ground Truth*) et celle obtenue par la technique présentée dans ce travail. L'erreur moyenne est de 1,36 mm avec une variance de 2,48 mm. Elle se situe principalement au niveau de la tête fémorale.

une variance de 2,5 mm. Les zones les moins bien segmentées étant situées dans la région de tête fémorale et sont dus à la difficulté de dissocier le fémur et l'acetabulum dans les cas de pathologies trop prononcées.

CHAPITRE VIII

CONCLUSION

Dans ce travail, nous proposons une méthode de segmentation par modèle déformable basée sur un processus de trois étapes. Tout d'abord, la phase d'initialisation du modèle se base sur la localisation extrêmement rapide des caractéristiques anatomiques du fémur à segmenter. Ces paramètres sont obtenus uniquement par une exploration partielle du volume au moyen de l'algorithme de rayons de recherche présenté au chapitre 5. Ensuite, le modèle déformable, tiré du domaine de l'animation de synthèse et basé sur une représentation par des cylindres généralisés (voir chapitre 6), est recalé sur l'imagerie CT sans nécessiter d'itérations. A l'issue de cette étape, le modèle n'identifie que grossièrement le contour du fémur. Une phase de recalage rigide entre le modèle et une segmentation par seuillage grossière s'est avérée nécessaire pour garantir de bons résultats lors de la segmentation fine. Cette dernière étape, introduite au chapitre 7, est réalisée par une technique de segmentation d'images basée sur les contours actifs. Contrairement aux approches classiques, leur résolution est basée sur une méthode de programmation dynamique.

8.1 Synthèse

En guise de résumé, le tableau suivant présente les résultats temporels et de mise en correspondance des différentes étapes pour la segmentation de la série du premier patient :

Dans la problématique du fémur, le point faible de cette approche est lié à la nécessité de devoir effectuer un recalage rigide entre le modèle déformable et la segmentation fine. Cette étape supplémentaire implique indirectement un filtrage des coupes qui contribue à plus de 50% du temps de calcul total.

| Opération | Précision [mm] | Temps [s] |
|---|----------------|-----------|
| Détermination des paramètres anatomiques | 22 | 5 |
| Filtrage des coupes et segmentation par seuillage | - | 15 |
| Recalage rigide du modèle déformable | 10 | 15 |
| Segmentation fine | 4,5 | 10 |
| Temps total de segmentation | | 45 |

TAB. 8.1: Précision (erreur moyenne quadratique) et temps de calcul (sur Pentium 4) des différentes étapes lors de la segmentation.

Il est tout à fait envisageable de supprimer le recalage fin si l'application ne requiert pas une précision extrême mais plutôt un traitement rapide. Dans ce cas, les deux dernières étapes susmentionnées ne sont pas indispensables. Cette variante de recalage permet, par exemple, de localiser approximativement certaines structures anatomiques. Dans le contexte fémoral, on peut imaginer un logiciel d'aide au choix prothétique fonctionnant sur ce principe. Celui-ci permettrait, entre autre, de déterminer automatiquement la taille de la cupule et la forme de la tige. Dans le domaine des tissus mous, cette technique pourrait être employée pour un recalage hépatique en cours d'intervention sur la base de plans de résection préétablis.

Il faut néanmoins relativiser ces chiffres lorsqu'on les compare au temps global de segmentation par une approche interactive qui, avec les logiciels du marché, nécessite entre 20 et 45 minutes de travail! Notre approche permet un gain de temps pour un opérateur d'un facteur compris entre cinq et dix.

8.2 Contributions

En plus d'une analyse ainsi que des tests très approfondis sur les techniques de recalage rigide et non rigide, notre travail propose de nouvelles contributions dans le domaine de la segmentation par modèles déformables en introduisant une décomposition du problème en trois parties originales :

- Une phase de localisation et détermination des paramètres anatomiques de la partie sagittale du fémur par une technique de rayons de recherche. Les particularités de l'approche sont une détermination rapide de ces paramètres car le problème est résolu dans un espace unidimensionnel, un taux de fiabilité élevé et la possibilité de corriger ou spécifier ces paramètres interactivement.
- La construction d'un modèle déformable hybride entre un modèle arti-

culé et les cylindres généralisés. Ce modèle est composé d'un squelette (défini par des cylindres et des sphères) connecté à une enveloppe. Toute modification du squelette modifie son enveloppe. Ce modèle à la particularité d'être intégralement paramétré par les informations anatomiques calculées durant l'étape de localisation. Il permet de définir un fémur générique composé de 13 degrés de liberté. L'avantage principal de notre approche réside en sa méthode de résolution directe permettant de transformer le modèle déformable relativement au jeu de données sans nécessiter de recalage non rigide itératif, ce qui est le cas pour les approches traditionnelles.

- Une phase de segmentation à proprement dite. Elle est une alternative élégante aux méthodes basées sur les contours actifs. Contrairement aux *snakes*, nous utilisons une résolution par programmation dynamique. Celle-ci évite une procédure classique de résolution itérative qui, dans la majorité des cas, converge vers un résultat sous-optimal. L'initialisation du contour actif par une coupe du modèle déformable améliore le résultat de segmentation. Il permet également de définir les paramètres de pondération optimaux du contour en fonction de la région en cours de segmentation.

L'intérêt principal de nos trois contributions, comparées aux techniques de segmentation par modèles déformables les plus couramment utilisées (comme les fonctions à base radiale ou les modèles statistiques), réside en la possibilité de corriger interactivement et à tout moment la segmentation grâce aux résultats intermédiaires interprétables par un médecin. Par exemple, si l'axe diaphysaire est mal détecté, il suffit de quelques clics de souris pour le corriger. Une meilleure compréhension des phases de segmentation par l'utilisateur permet, indirectement, d'améliorer et de simplifier la phase de validation.

En conclusion, notons que la précision obtenue au terme de cette segmentation s'avère largement suffisante pour une utilisation dans des logiciels de planification prothétiques.

8.3 Travaux futurs

Tout au long de ce travail, certaines voies ont été explorées plus ou moins sommairement. Presque tous les chapitres présentent des apports originaux qui, dans certains cas, nécessiteraient une meilleure formalisation et

devraient, dans l'idéal, être sujets à des recherches plus approfondies.

Dans une continuation directe de cette étude, il nous semble que les pistes suivantes pourraient être explorées :

- La possibilité de fiabiliser et généraliser la localisation des paramètres anatomiques pour d'autres types de protocoles radiologiques et d'autres imageurs. De par ses bonnes performances temporelles, l'adaptation de notre méthode de recalage aux scanners intra-opératoires comme l'échographie nous semble intéressante ;
- Dans le cadre fémoral, l'étude d'un modèle déformable comportant plus de degrés de liberté permettrait de mieux prendre en compte la mauvaise correspondance au niveau des deux trochanters ;
- La résolution des contours actifs (*snakes*) par un processus itératif est la faiblesse qui les rend pratiquement inutilisables. Aussi, l'adaptation de notre approche aux contours concaves permettrait de supplanter cette approche traditionnelle.

La rapidité de notre approche nous fait penser que nous pourrions la transposer à la problématique du recalage temps-réel d'organes déformables comme le foie. Dans ce cas, la difficulté réside encore à redéfinir un modèle basé sur les paramètres anatomiques spécifiques de l'organe et à adapter sa détection au moyen d'un imageur temps-réel.

Mais ceci est l'histoire d'une autre thèse ! Et j'espère de tout coeur que mon successeur, qui est déjà au travail, parviendra à surmonter ces difficultés.

8.4 Publications et posters

La liste suivante présente les principaux posters et publications qui ont été réalisées dans le cadre de cette thèse.

- S. Sagbo, G. Marti, C. Di Ventuto, C. Vangenot, L. Nolte, and G. Zheng. Design and implementation of deformation algorithms for computer assisted orthopedic surgery : application to virtual implant database and preliminary results, poster session, EMBC, 2005.
- G. Marti, C. Baur, and P. Zambelli. Optimal femoral head contour segmentation in CT images using dynamic programming. *Technology Health Care.*, 12(4) :315 - 322, 2004.

- G. Marti, Fast coarse-to-fine elastic registration of medical data, poster session and presentation, CO-ME 2004.
- G. Marti, P.-Y. Zambelli, and C. Baur. Optimal Femoral Head Contour Segmentation in CT Images Using Dynamic Programming. In Poster session, Computer Assisted Radiology and Surgery, 2003.
- P.-Y. Zambelli, C. Bregand, S. Dewarrat, G. Marti, C. Baur, and P. Leyvraz. Planning and navigation solution in resurfacing hip surgery : A way to reduce the surgical approach. In Poster session, 3rd Annual Meeting of the International Society for Computer Assisted Orthopaedic Surgery, 2003.
- G. Marti, P.-Y. Zambelli, and C. Baur. Optimal Femoral Head Contour Segmentation in CT Images Using Dynamic Programming. In Poster session, 3rd Annual Meeting of the International Society for Computer Assisted Orthopaedic Surgery, 2003.

APPENDICE A

APPARIEMENT FLOU

Cette annexe présente un exemple d'appariement flou présenté au chapitre 3.3.3.

Soit un modèle composé de deux points m_1 et m_2 et d'un jeu de données comprenant trois points d_1 , d_2 et d_3 . Le graphique 8.1 présente la configuration à l'itération $iter$ une fois la transformation T_{iter} appliquée sur le modèle. La métrique utilisée est linéaire Δ .

La première étape de l'appariement consiste à construire la matrice de correspondance floue F et de lui adjoindre une *slack variable* pour la rendre carrée (voir table 8.2). Les différents éléments de cette matrice correspondent à la distance Euclidienne entre les points : $(f_{ij} = \Delta(T_{iter} \star d_i, m_j))$,

L'étape suivante est la normalisation de la matrice en fonction de la température du recuit simulé. A début de l'optimisation, la température est élevée ($\beta_{iter} \ll$) et diminue au cours des itérations ($\beta_{iter} \gg$). La matrice M suivante (8.3) est calculée en remplaçant les éléments de F avec la formule $m_{ij} \leftarrow \exp(-\beta_{iter} f_{ij})$:

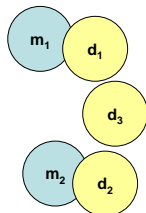


FIG. 8.1: Configuration des jeux de données. Le modèle M est composé de deux points et les données D de trois.

| F | d_1 | d_2 | d_3 |
|---------|-------|-------|-------|
| m_1 | 1,0 | 3,0 | 2,0 |
| m_2 | 2,5 | 1,0 | 1,5 |
| $slack$ | 0,0 | 0,0 | 0,0 |

TAB. 8.2: Matrice de correspondance floue F de l'exemple 8.1.

| M_{iter} | d_1 | d_2 | d_3 |
|------------|-------|-------|-------|
| m_1 | 0,90 | 0,74 | 0,82 |
| m_2 | 0,78 | 0,90 | 0,86 |
| $slack$ | 1,0 | 1,0 | 1,0 |

TAB. 8.3: Matrice de correspondance M de la matrice F (table 8.2) à l'iteration $iter$ avant normalisation. Dans cet exemple, le paramètre $\beta_{iter} = 0,1$.

En 1964, Sinkhorn ^[90] prouve qu'une matrice doublement stochastique peut être obtenue à partir de n'importe quelle matrice carrée définie positive en normalisant alternativement ses lignes et ses colonnes. Dans le cas de notre exemple, le résultat de la normalisation de la matrice M_{iter} est la matrice \tilde{M}_{iter} suivante :

| \tilde{M}_{iter} | d_1 | d_2 | d_3 |
|--------------------|-------|-------|-------|
| m_1 | 0,63 | 0,52 | 0,57 |
| m_2 | 0,52 | 0,62 | 0,58 |
| $slack$ | 0,58 | 0,58 | 0,58 |

TAB. 8.4: Matrice de correspondance \tilde{M}_{iter} de la matrice M_{iter} 8.3 après 100 normalisations alternatives des lignes et des colonnes.

On peut ainsi définir la mesure de distance "floue" entre les deux jeux de données par :

$$\Delta_{\beta}^{floue}(D, M) = \sum_{i,j} \tilde{m}_{ij} f_{ij}$$

En d'autres termes, la matrice \tilde{M}_{iter} définit le poids ou une mesure de probabilité de l'appariement entre les deux jeux de données.

Au fil des itérations, lorsque β augmente, cette matrice tend à devenir une matrice de correspondance binaire Z :

| Z | d_1 | d_2 | d_3 |
|---------|-------|-------|-------|
| m_1 | 1 | 0 | 0 |
| m_2 | 0 | 1 | 0 |
| $slack$ | 0 | 0 | 1 |

TAB. 8.5: Matrice de correspondance \tilde{M}_{iter} lorsque β augmente (dans ce cas, $\beta > 20$). Cette matrice correspond maintenant à une matrice de correspondance binaire Z .

APPENDICE B

REPRÉSENTATION DES ROTATIONS

Comme nous l'avons vu aux chapitres 3 et 4, les familles de transformation rigides et non rigides globales comprennent au moins une rotation et une translation (ou pose). Contrairement aux vecteurs de translations qui comprennent un DDL par dimension, les matrices de rotations possèdent, quant-à-elles, neuf éléments pour trois DDL. Le choix du formalisme mathématique pour définir cette matrice détermine les performances des algorithmes d'optimisations, comme par exemple, l'algorithme ICP décrit aux sections 3.3 et 3.3.5.

8.5 Angles d'Euler

La représentation la plus naturelle de la rotation est définie par ses angles d'Euler (Leonhard Euler, physicien et mathématicien Suisse 1707-1783 et entre autre inventeur du Sudoku). Il s'agit d'une généralisation des coordonnées sphériques. Une matrice de rotation R peut toujours être écrite comme le produit de trois rotations successives autour d'axes orthogonaux comme par exemple les axes \vec{e}_1 , \vec{e}_2 et \vec{e}_3 appelés également axes de roulis, tangage et lacet (voir figure 8.2) :

$$R_{e_1 e_2 e_3} = \begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\gamma & -\sin\gamma \\ 0 & \sin\gamma & \cos\gamma \end{pmatrix}$$

L'opération inverse pour l'extraction d'angles à partir d'une matrice de rotation est aisément réalisée par l'identification des coefficients de la matrice avec leur expression analytique.

Les angles d'Euler ont pour avantage de pouvoir représenter la matrice de rotation par son minimum de DDL, soit trois. Cette méthode donne, cependant, des équations non-linéaires (trigonométriques) et présente des singularités et des ambiguïtés de type *Gimbal Lock*, ce qui a pour effet un mauvais conditionnement des problèmes d'optimisation.

8.5.1 Gimbal Lock

Le problème de la perte d'un DDL, le "Gimbal Lock", arrive quand on utilise les angles d'Euler. Physiquement, ce phénomène peut être observé en astronautique dans les systèmes de fixation des gyroscopes. Anecdotiquement, une description intéressante est disponible dans le manuel de guidage et de navigation de la mission Apollo (voir le rapport de D. Hoag, E-1344 : Apollo Guidance and Navigation Considerations of Apollo IMU Gimbal Lock, MIT, 1963).

Comme la matrice de rotation finale dépend de l'ordre des multiplications, il est possible que parfois, l'une des rotations autour d'un axe soit confondue avec un autre axe de rotation (voir figure 8.2). Pire encore, il est devenu parfois impossible de tourner l'objet suivant un certain axe. C'est le Gimbal Lock. Par exemple, supposons qu'un objet est tourné dans l'ordre \vec{e}_3 , \vec{e}_2 et \vec{e}_1 et que la rotation autour de \vec{e}_2 soit de 90° . Dans ce cas, la rotation suivant \vec{e}_3 se fait correctement, puisque c'est la première. L'axe \vec{e}_2 tourne également correctement. Néanmoins, après cette rotation, l'axe \vec{e}_1 et l'axe \vec{e}_3 sont confondus. Du coup, toute rotation autour de l'axe \vec{e}_1 tourne l'objet suivant \vec{e}_3 ! Pire encore, il est devenu impossible de tourner l'objet autour de l'axe \vec{e}_1 .

Une possibilité pour remédier à ce problème est de représenter les rotations par des quaternions unitaires.

8.6 Quaternions

"Les quaternions furent inventés par William Rowan Hamilton en 1843 à partir des travaux de Carl Friedrich Gauss et au siècle précédent Leonhard Euler. Il étudiait alors l'interprétation géométrique de l'arithmétique de nombres complexes dans le plan et cherchait à obtenir des résultats analogues dans l'espace à trois dimensions.

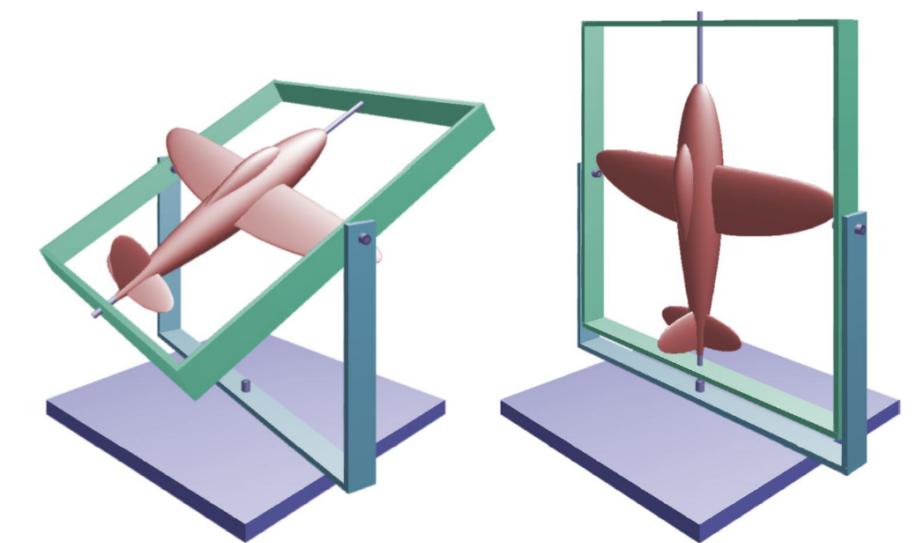


FIG. 8.2: Exemple de *Gimbal Lock* tiré du site web de Gernot Hoffmann. A gauche, représentation d'un objet avec les rotations d'Euler représentées par les différents cadrans. A droite, phénomène de *Gimbal Lock*, où la perte d'un DDL est due à la rotation du cadran intérieur qui fait se superposer les deux autres axes.

Après des années de recherches sur la construction d'une algèbre avec des "triplets" de trois nombres réels, il butait sur la multiplication, et en particulier la conservation des normes (Georg Ferdinand Frobenius a démontré en 1877 qu'une telle multiplication de triplets était impossible à définir).

Il eut alors l'idée d'utiliser des "quadruplets" en employant une dimension supplémentaire. Selon ses dires, il marchait un jour dehors le long du canal royal, avec son épouse quand soudain lui vint à l'esprit la solution sous la forme des relations : $i^2 = j^2 = k^2 = ijk = -1$. Il grava alors promptement ces relations avec un couteau dans une pierre du pont de Brougham (maintenant appelé Broom Bridge) à Dublin.", tiré de l'encyclopédie Wikipédia.

Une rotation dans l'espace 3D peut être représentée par un quaternion unitaire ^[40]. Les quaternions sont des nombres avec trois composantes complexes qui peuvent être écrits par la combinaison linéaire suivante : $a + bi + cj + dk$, où i , j et k sont les coordonnées complexes avec les propriétés décrites précédemment. Ces nombres peuvent également être représentés par

un scalaire combiné à un vecteur 3D (a, \vec{v}) . Une rotation autour d'un vecteur unitaire \vec{w} par un angle θ est décrit par le quaternion unitaire suivant :

$$q = (\cos(\frac{1}{2}\theta), \vec{w}\sin(\frac{1}{2}\theta))$$

Pour effectuer une rotation 3D d'un point \vec{m} , nous pouvons l'écrire comme le quaternion $p = (0, \vec{m})$. Le point p' qui est le résultat de la rotation est calculé par :

$$p' = q \cdot p \cdot \bar{q}$$

où \cdot est l'opérateur de multiplication dans l'algèbre des quaternions est \bar{q} le conjugué complexe. Pour de plus amples informations sur l'algèbre des quaternions, on peut se référer à l'excellent site wiki francophone (<http://fr.wikipedia.org/wiki/Quaternion>).

Les quaternions évitent le phénomène de "Gimbal Lock" mais impliquent que les techniques d'optimisation contraignent la norme de q à l'unité. La fonction objective suivante est généralement rajoutée : $k(1 - \|q\|^2)$, où k est une constante élevée. Malheureusement, cette dernière limitation augmente la complexité algorithmique et réduit la vitesse de convergence lors de l'optimisation. Aussi, la représentation par des cartes exponentielles (*Exponential Map*) ou par la formule de Rodrigues est de préférence utilisée.

8.7 Cartes exponentielles et formule de Rodrigues

Les cartes exponentielles permettent de définir la matrice de rotation comme une rotation autour d'un axe arbitraire \vec{l} , d'un angle de magnitude ω . Suivant la formule de Rodrigues, la matrice de rotation R peut s'écrire :

$$R = \begin{pmatrix} \cos\omega + l_1^2(1 - \cos\omega) & l_1l_2(1 - \cos\omega) - l_3\sin\omega & l_1l_3(1 - \cos\omega) + l_2\sin\omega \\ l_1l_2(1 - \cos\omega) + l_3\sin\omega & \cos\omega + l_2^2(1 - \cos\omega) & l_2l_3(1 - \cos\omega) - l_1\sin\omega \\ l_1l_3(1 - \cos\omega) - l_2\sin\omega & l_2l_3(1 - \cos\omega) + l_1\sin\omega & \cos\omega + l_3^2(1 - \cos\omega) \end{pmatrix}$$

Les expressions \vec{l} et ω peuvent être extraites de la matrice précédente :

$$\omega = \arccos\left(\frac{\text{trace}(R) - 1}{2}\right)$$

$$\vec{l} = \text{normalize} \begin{pmatrix} r_{32} - r_{23} \\ r_{13} - r_{31} \\ r_{21} - r_{12} \end{pmatrix}$$

Lors d'optimisations, les quatre paramètres que sont l'angle de rotation ω ainsi que le vecteur-3 \vec{l} sont généralement codés dans le vecteur-3 suivant : $\vec{l}' = \omega \cdot \text{normalize}(\vec{l})$.

La matrice de rotation R précédente peut, de manière similaire, être représentée par la série infinie suivante :

$$\exp(\lambda) = I + \lambda + \frac{1}{2!}\lambda^2 + \frac{1}{3!}\lambda^3 + \dots$$

où la matrice antisymétrique λ est définie comme :

$$\lambda = \begin{pmatrix} 0 & -l_3 & l_2 \\ l_3 & 0 & -l_1 \\ -l_2 & l_1 & 0 \end{pmatrix}$$

Il s'agit de la représentation par carte exponentielle (*Exponential Map*) qui tire son nom de la série infinie précitée. L'avantage de cette dernière représentation est d'éviter le problème du *Gimbal Lock* lié aux angles d'Euler et ne requiert pas de contraintes additionnelles lors des optimisations pour préserver la norme des quaternions. La formule de Rodrigues permet un passage simple et optimal entre la matrice de rotation et sa représentation (ω, \vec{l}) et réciproquement. Une dernière qualité, non négligeable, de ce modèle tiens à sa compréhensibilité géométrique.

APPENDICE C

LE FÉMUR

Les différentes descriptions anatomiques suivantes sont en grande partie tirées de l'encyclopédie libre Wikipédia (<http://fr.wikipedia.org/>), de l'anatomie du corps humain de Henry Gray, anatomiste et chirurgien anglais du 19ème siècle, dont les travaux sont maintenant disponibles sur internet à l'adresse <http://education.yahoo.com/reference/gray/subjects/subject?id=59>, ainsi que du site francophone (www.anatomie-humaine.com).

8.8 Anatomie du fémur

Le fémur est l'os le plus long de la jambe, le plus gros (par son volume), et le plus solide (quand on prend en compte sa résistance aux contraintes mécaniques, contraction des muscles, etc.) du corps humain. Le fémur se compose d'un corps (la diaphyse) et du col proximement (vers le haut du fémur en position anatomique) limitant l'épiphyse proximale. La partie supérieure du fémur a la forme d'un 2/3 de sphère (une énarthrose) s'emboîtant dans le bassin (acetabulum de l'os coxal). En situation proximale, on trouve deux reliefs osseux, les trochanters (le petit et le grand), permettant l'insertion des certains muscles. Plus bas, la ligne âpre permet également l'insertion d'autres muscles, dont le grand fessier et le biceps crural. Distalement, on trouve deux condyles (un médial et un latéral) et une fosse formant une articulation trochléarthrose avec le tibia et la rotule. Les figures 8.3 et 8.4 illustre bien les parties principales de l'anatomie osseuse du fémur.

8.8.1 Partie supérieure

L'extrémité supérieure du fémur (voir figure 8.5) est organisée en système de faisceaux partant de trois lames compactes qui sont : la lame compacte

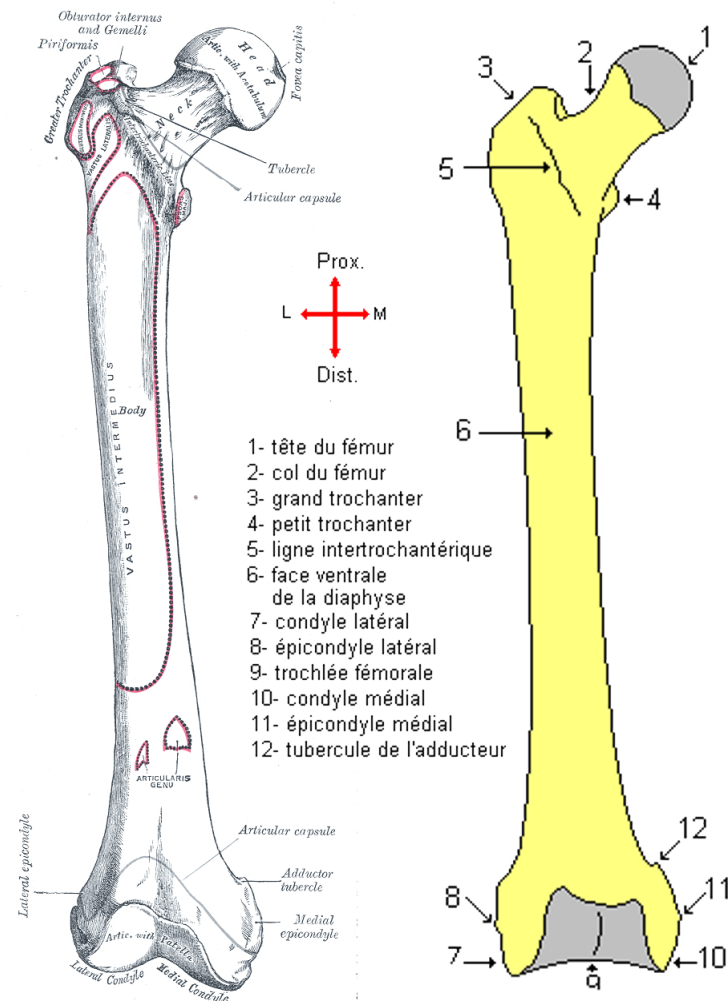


FIG. 8.3: Face antérieure du fémur. A gauche, illustration du 19ème siècle de l'anatomiste H. Gray, à droite, représentation moderne tirée du site web www.anatomie-humaine.com.

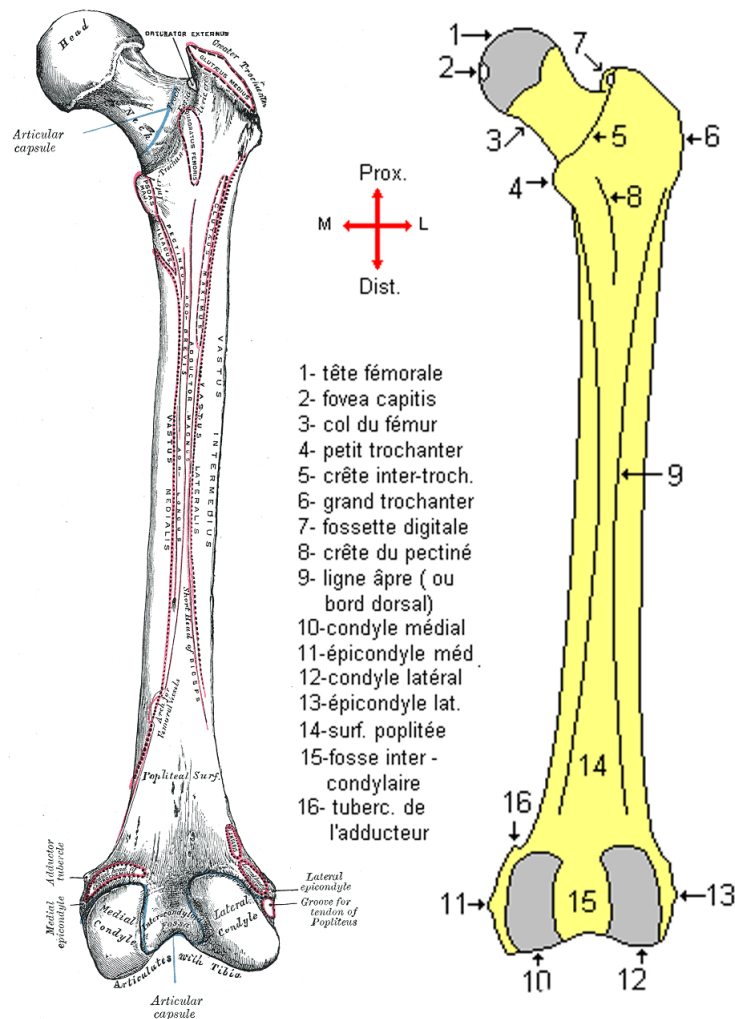


FIG. 8.4: Face postérieure du fémur. A gauche, illustration du 19ème siècle de l'anatomiste H. Gray, à droite, représentation moderne tirée du site web www.anatomie-humaine.com.

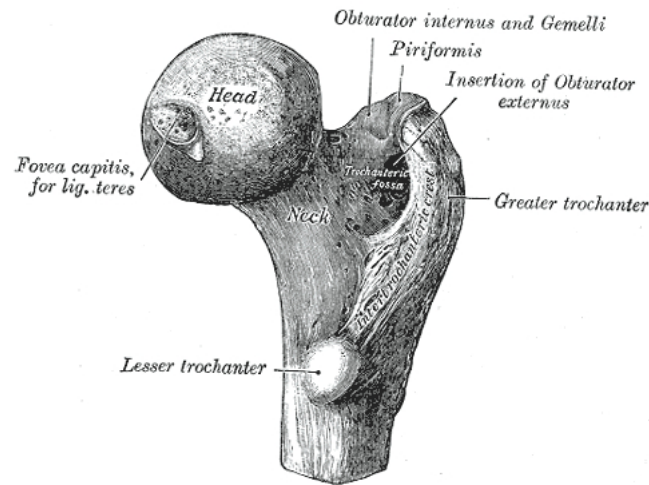


FIG. 8.5: Partie proximale du fémur selon H. Gray.

interne ou Arc d'Adams, la lame compacte externe et la lame compacte sus-cervicale. De la lame compacte interne part le faisceau trochantérien vers le grand trochanter. De la lame compacte externe, part le faisceau arciforme de Gallois vers la tête fémorale. De la lame sus-cervicale, partent des traversées osseuses interne et externe. Ce système de faisceaux délimite une zone de faiblesse appelée triangle de Ward. L'union des faisceaux arciforme et trochantérien donne un système appelé système ogival.

8.8.2 La tête (*caput femoris*)

La tête a la forme de $2/3$ de sphère et est dirigée vers le haut, vers l'intérieur et un peu vers l'avant, la majeure partie de sa convexité étant au-dessus et en avant. Sa surface est lisse, enrobée de cartilage, à l'exception d'une dépression ovoïde, la fossette de la tête (*fovea capitis femoris*, ancienne fossette du ligament rond), qui est située un peu sous et derrière le centre de la tête, et donne un attachement au ligament de la tête fémorale (*ligamentum capitis femoris*, ancien ligament rond).

8.8.3 Le col (*collum femoris*)

Le col est un os plat pyramidal, reliant la tête avec le corps, et formant avec ce dernier un grand angle (l'angle cervico-diaphysaire). L'angle est plus

grand à l'enfance et diminue durant la croissance, formant éventuellement une faible courbe rendu à la puberté. Chez l'adulte, le col forme un angle d'environ 125° avec le corps, mais variant en proportion inverse au développement du bassin. En plus de se projeter supérieurement et médialement du corps du fémur, le col se projette également antérieurement (angle d'antéversion), variablement chez les individus, allant en général de 20° à 30° .

Si l'angle cervico-diaphysaire est supérieur à 130° on parle de coxa valga, et s'il est inférieur à 110° on parle de coxa vara. Plus l'angle est petit plus il y a un risque de fracture du col du fémur (fréquence élevée chez le vieillard).

8.8.4 Les trochanters

Les trochanters sont des bosses-sphères proéminentes qui offrent un bras de levier aux muscles rotateurs de la cuisse. On retrouve deux trochanters : le grand et le petit.

Le grand trochanter (*trochanter major*) est une éminence quadrilatérale située à la jonction du col avec la partie supérieure du corps. Il est situé légèrement latéralement et postérieurement et, chez l'adulte, est environ 1 cm plus bas que la tête. Il a deux surfaces et quatre arêtes. La surface latérale sert d'insertion au muscle moyen glutéal (*gluteus medius*, ancien muscle fessier moyen). La surface médiale, plus petite que la latérale, présente à sa base une dépression, la fosse trochantérique (*fossa trochanterica*, ancienne fossette digitale), où le tendon du muscle obturateur externe (*obturator externus*). Au-dessus et en avant se trouvent le lieu de l'attachement de l'obturateur interne et des muscles jumeaux. L'arête supérieure est le lieu d'insertion du muscle piriforme. L'arête inférieure donne naissance à la partie supérieure du vaste latéral (*vastus lateralis*). L'arête antérieure est proéminente, le muscle petit glutéal (*gluteus minimus*, ancien petit fessier) s'y rattachant.

Le petit trochanter (*trochanter minor*) est une éminence conique qui varie en grosseur chez différents individus. Il projette de la partie inférieuro-postérieure de la base du col. À son sommet s'attache le tendon du muscle ilio-psoas.

8.8.5 Tubercule

Une proéminence, le tubercule du fémur, se trouve à la jonction de la partie supérieure du col et du grand trochanter. Il est le lieu d'attache de cinq muscles : muscle petit fessier latéralement, le vaste latéral en dessous et le tendon de l'obturateur interne et des deux gemilli au-dessus.

8.8.6 Ligne intertrochantérique

Descendant obliquement et médialement du tubercule, elle offre un attache au ligament ilio-fémoral de la hanche. Sa partie inférieure donne origine à la partie supérieure du vaste médial (*vastus medialis*). La ligne quadrata (*linea quadrata*) est le lieu d'origine du quadriceps fémoral (*quadratus femoris*) et de quelques fibres du muscle adducteur magnus.

8.9 L'architecture interne de la partie proximale du fémur

L'os spongieux de la partie proximale du fémur (jusqu'à la partie inférieure du petit trochanter) est composé de deux systèmes distincts de trabécules arrangées de manière curviligne. La première, qui a son origine sur la partie médiale de la diaphyse se courbe radialement jusqu'à l'extrémité opposée de l'os, tandis que l'autre, à son origine latérale et se courbe radialement jusqu'à la tête fémorale (voir figure 8.6). Si l'on effectue une analyse biomécanique de l'architecture interne du fémur ^[58], on remarque que le système dual de trabécules correspond à la position et la courbure des lignes d'efforts de compression maximales (voir figure 8.6).

8.10 Pathologies du fémur

Parmi les nombreuses maladies osseuses, on peut citer les anomalies d'origine génétique, les infections (ostéites), les troubles du métabolisme osseux (ostéoporose, ostéomalacie, hyperparathyroïdie), les tumeurs (tumeurs bénignes ou malignes, métastases osseuses, myélome), les microtraumatismes (fractures de fatigue ou de stress) ou d'autres pathologies comme la maladie de Paget.

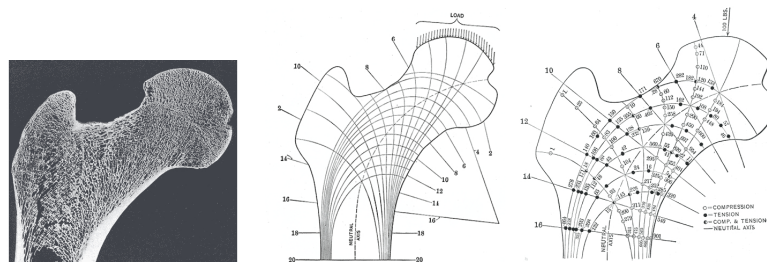


FIG. 8.6: Différentes illustrations tirées des travaux précurseurs du biomécanicien Kock [58] en 1917. A gauche, section médiane de la partie proximale du fémur. On remarque aisément les deux réseaux de trabécules dans la partie non corticale de l'os réalisant un motif en forme de nid d'abeille. Au milieu, diagramme des lignes d'efforts lorsqu'une personne se tient debout sur ses deux jambes. A droite, intensité maximale de tension et de compression pour une charge d'environ 50kg située sur la partie proximale de la tête fémorale.

8.10.1 Ostéoporose

L'ostéoporose est la perte progressive de masse osseuse, qui se traduit par une diminution de la densité osseuse et/ou une altération de la microarchitecture osseuse (voir figure 8.7). Elle est due à un déséquilibre entre l'activité catabolique et anabolique, qui s'aggrave avec l'âge, si bien que cette maladie affecte principalement les personnes âgées. Elle est souvent la conséquence d'une déficience ou d'une moins bonne absorption du calcium et de la vitamine D. L'ostéoporose s'observe facilement dans des images CT, vu que l'intensité des pixels de cette modalité est directement proportionnelle à la densité des tissus traversés. De ce fait, il est très difficile de segmenter des patients souffrant de cette pathologie.

8.10.2 Arthroses de hanche ou coxarthroses

L'arthrose est une maladie qui touche les articulations, on l'appelle aussi arthropathie chronique dégénérative. Elle est caractérisée par la douleur, mécanique et diurne et la difficulté à effectuer des mouvements articulaires. Au niveau de l'articulation, la surface du cartilage se fissure, s'effrite et finit par disparaître. Ensuite, des excroissances osseuses se forment et nuisent aux mouvements. L'arthrose est une dégénérescence du cartilage des articulations sans infection ni inflammation particulière. Cette dégénérescence conduit à une destruction plus ou moins rapide du cartilage qui enrobe l'extrémité des os. Anatomiquement, cette destruction s'accompagne d'une prolifération



FIG. 8.7: Radiographie d'une articulation de la hanche avec, à gauche, personne atteinte d'ostéoporose et, à droite, personne normale.

osseuse sous le cartilage (voir figure 8.8). C'est la maladie articulaire la plus fréquente. Les premiers symptômes apparaissent généralement à partir de 40-50 ans, mais la maladie commence souvent bien plus tôt dans la vie.

Dans les cas évolués les plus invalidants, une résection articulaire, une arthrodèse (blocage définitif de l'articulation), une prothèse totale de hanche (PTH) peuvent être proposées au malade.

Au niveau radiographique (ou CT), on constate un pincement de l'interligne articulaire (espace entre les deux os) témoignant de l'usure du cartilage, des ostéophytes (proliférations osseuses) sur le cotyle et la tête fémorale, voire des géodes (cavités) osseuses du cotyle et de la tête fémorale, signes d'une souffrance articulaire déjà très importante.

8.10.3 Fractures

Les fractures de l'extrémité proximale du fémur sont fréquentes chez les personnes âgées. Elles sont liées à l'ostéoporose. Elles surviennent souvent après quatre-vingts ans à l'occasion d'un traumatisme minime et sont plus fréquentes chez les femmes. Leur gravité vient de la fragilité des patients atteints. Globalement un tiers des patients décède dans l'année suivant l'accident, un tiers devient plus ou moins dépendant et seulement un tiers récupère

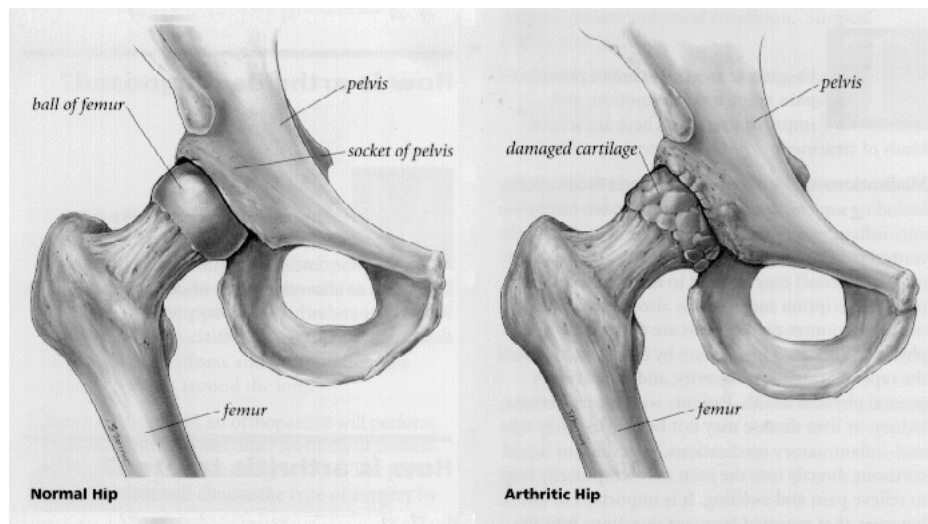


FIG. 8.8: Illustration d'une articulation de la hanche normale (à gauche) et souffrant d'arthrose (à droite). Tiré du site web www.ozarkortho.com/patiented/arthritis.htm.

une fonction normale.

8.10.4 Autres pathologies

Cette liste n'a pas l'ambition d'être exhaustive. Je pense qu'il est néanmoins encore important de citer les rhumatismes et les nécroses osseuses qui sont des formes particulières d'arthroses.

APPENDICE D

LOCALISATION DE CENTRE D'ELLIPSES PAR LANCER DE RAYONS

Soit d_1 une droite coupant un cercle en deux points distincts P_1 et P_2 . R est le centre de ce segment. Soit une autre droite d_2 ($d_2 \neq d_1$) passant par R . Elle coupe le cercle en deux points distincts P'_1 et P'_2 . R' est le centre de ce segment (voir figure 8.9).

8.10.5 Proposition

En répétant le processus précédent, les points successifs R, R', R'', \dots convergent vers le centre du cercle C .

8.10.6 Démonstration

Le segment $P'_1P'_2$ est perpendiculaire à $R'C$ car le point R' est équidistant des points P'_1 et P'_2 . Les points C, R et R' définissent donc un triangle rectangle. Selon Pythagore, nous avons $\|\vec{CR}\|^2 = \|\vec{R'R}\|^2 + \|\vec{CR'}\|^2$ ou l'inégalité suivante : $\|\vec{CR}\| \geq \|\vec{CR'}\|$. Dans le cas spécifique de l'égalité, $d_2 \perp d_1$ et l'algorithme converge en une seule itération vers le centre. Dans les autres cas, la distance des points successifs R vis-à-vis du centre C tendent à diminuer *cqfd*.

8.10.7 Proposition

La proposition précédente est également valable pour une ellipse non dégénérée.

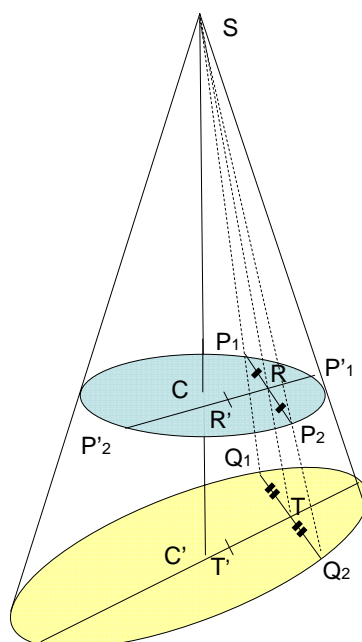


FIG. 8.9: Schéma pour les démonstrations du principe de localisation de centre d'ellipses par lancer de rayons successifs.

8.10.8 Démonstration

La démonstration suivante tente de rapporter le problème de l'ellipse au problème du cercle.

Toute ellipse (en jaune sur la figure 8.9) peut être vue comme une section d'un cône de révolution par un plan. Sur le même cône, nous pouvons représenter un cercle (en bleu). Une droite e_1 coupe l'ellipse en deux points Q_1 et Q_2 , T étant le point central du segment. Le point P_1 est défini comme étant l'intersection de la droite SQ_1 avec le cercle. S étant le sommet du cône. De manière similaire, P_2 est l'intersection de la droite SQ_2 et R l'intersection de la droite TS . Par le principe de proportionnalité, R est également le centre du segment P_1P_2 . Aussi, n'importe quelle droite q_2 ($q_2 \neq q_1$) coupant l'ellipse et passant par T correspond à une droite d_2 coupant le cercle et passant par R . Comme le point R' du cercle correspondra au point T' de l'ellipse, la distance des points successifs T, T', \dots avec le centre de l'ellipse tendra à diminuer. La démonstration précédente est donc valable pour l'ellipse cqfd.

GLOSSAIRE

Angiographie L'angiographie est une technique d'imagerie médicale dédiée à l'étude des vaisseaux sanguins qui ne sont pas visibles sur des radiographies standards. L'angiographie est un examen basé sur l'injection d'un produit de contraste lors d'une imagerie par rayons X. Sa signification littérale est "imagerie des vaisseaux". L'angiographie est un examen invasif.

Anisotropie L'anisotropie (contraire d'isotropie) est la propriété d'être dépendant de la direction. Quelque chose d'anisotrope pourra présenter différentes caractéristiques selon la direction.

Arthroplastie Réfection chirurgicale d'une articulation avec ou sans implantation d'une prothèse.

Distance de Hausdorff En topologie, la distance de Hausdorff mesure l'éloignement de deux sous-ensembles d'un espace métrique. Elle porte le nom du mathématicien allemand Felix Hausdorff.

Distance de Levenshtein La distance de Levenshtein mesure la similarité entre deux chaînes de caractères. Elle est égale au nombre minimal de caractères qu'il faut supprimer, insérer, ou remplacer pour passer d'une chaîne à l'autre. Son nom provient de Vladimir Levenshtein qui l'a définie en 1965. Elle est aussi connue sous le nom de distance d'édition ou encore de déformation dynamique temporelle, notamment en reconnaissance de la parole. Cette distance est d'autant plus grande que le nombre de différences entre les deux chaînes est grand. La distance de Levenshtein peut être considérée comme une généralisation de la distance de Hamming. On peut montrer en particulier que la distance de Hamming est un majorant de la distance de Levenshtein.

Espace multi-échelles La théorie des espaces multi-échelles est une représentation développée dans les domaines de la vision artificielle, l'analyse d'image et le traitement du signal. Il s'agit d'une théorie formelle qui permet la gestion de structures d'images à différentes échelles de telle sorte que des structures fines peuvent successivement être supprimées (en appliquant une convolution avec une gaussienne) et un paramètre d'échelle t peut être associé à tous les niveaux de la représentation multi-échelle.

Expectation-maximisation L'Algorithme espérance-maximisation (EM), proposé par Dempster et al. en 1977, est une classe d'algorithmes qui permettent de trouver le maximum de vraisemblance des paramètres

de modèles probabilistes lorsque le modèle dépend de variables latentes non observables. On utilise souvent l'espérance-maximisation pour la classification de données, en apprentissage machine, ou en vision artificielle. L'espérance-maximisation alterne des étapes d'évaluation de l'espérance (E), où l'on calcule l'espérance de la vraisemblance en tenant compte des dernières variables observées, et une étape de maximisation (M), où l'on estime le maximum de vraisemblance des paramètres en maximisant la vraisemblance trouvée à l'étape E. On utilise ensuite les paramètres trouvés en M comme point de départ d'une nouvelle phase d'évaluation de l'espérance, on l'on itère ainsi.

Fonction monotone Les fonctions monotones sont des fonctions dont le sens de variation ne change pas. Une fonction monotone sur un intervalle I est une fonction qui reste croissante ou qui reste décroissante sur cet intervalle.

Homéomorphisme Un homéomorphisme entre deux espaces topologiques est une bijection continue de l'un dans l'autre, dont la réciproque est continue. La notion d'homéomorphisme est la bonne notion pour dire que deux espaces topologiques sont à peu près le même à vue différemment. C'est la raison pour laquelle les homéomorphismes sont les isomorphismes de la catégorie des espaces topologiques.

Maillage Un maillage est la discrétisation spatiale d'un milieu continu, ou aussi, une modélisation géométrique d'un domaine par des éléments proportionnés finis et bien définis.

Matrice stochastique Une matrice stochastique (aussi appelée matrice de Markov) est une matrice carrée dont chaque élément est un réel compris entre 0 et 1 et dont la somme des éléments de chaque ligne vaut 1. Cela correspond, en probabilité, à la matrice de transition d'une chaîne de Markov finie.

Ostéosynthèse L'ostéosynthèse regroupe l'ensemble des procédés qui permettent de traiter des fractures ou des problèmes d'ordre mécanique sur le squelette à l'aide de vis, de plaques, de clous, de tiges etc. placés en interne, à même l'os.

Resurfacing Dans la prothèse classique on enlève l'extrémité supérieure du fémur et on la remplace par une tige solidement ancrée dans le fémur sur une quinzaine de centimètres. Dans le resurfacing, on retire simplement le cartilage et le moins d'os possible. La cupule fémorale vient comme un chapeau recouvrir la tête du fémur. le sacrifice osseux est minime.

Sulcatures Plis sur la surface du cerveau.

RÉFÉRENCES

- [1] R. Adams and L. Bischof. Seeded region growing. *IEEE Trans. Pattern. Anal. Machine Intell.*, 16(9) :641–647.
- [2] A. Johnson and M. Hébert. Using spin images for efficient object recognition in cluttered 3-d scenes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(5) :433–449, 1999.
- [3] A. Amini, S. Tehrani, and T. Weymouth. Using dynamic programming for minimizing the energy of active contours in the presence of hard constraints. *Computer Vision., 1988. Second International Conference on*, pages 95–99, 1988.
- [4] A. Amini, T. Weymouth, and R. Jain. Using dynamic programming for solving variational problems in vision. *IEEE Trans. Pattern Anal. Mach. Intell.*, 12(9) :855–867, 1990.
- [5] N. Arad and D. Reinfeld. Image warping using few anchor points and radial functions. *Computer Graphics Forum*, 14(1) :35–46, 1995.
- [6] K. S. Arun, T. S. Huang, and S. D. Blostein. Least-square fitting of two 3-d points sets. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9 :698–700, 1987.
- [7] N. Ayache. L'analyse automatique des images médicales : Etat de l'art et perspectives. *Rapport de recherche de l'INRIA*, 3364, 1998.
- [8] E. Bardinet, L.D. Cohen, and N. Ayache. Tracking and motion analysis of the left ventricle with deformable superquadrics. *Medical Image Analysis*, (2), 1996. also INRIA Research Report RR-2797, INRIA Sophia-Antipolis, February 1996.
- [9] R. Bellman and S. Dreyfus. Functional approximations and dynamic programming. *Mathematical Tables and Other Aids to Computation*, 13(68) :247–251, 1959.
- [10] P.J. Besl and N.D. McKay. A method for registration of 3-d shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2) :239–256, 1992.
- [11] T.O. Binford. Visual perception by computer. *IEEE Conference on Systems and Control*, 1971.
- [12] M. Black and A. Rangarajan. On the unification of line processes, outlier rejection, and robust statistics with applications in early vision. *International Journal of Computer Vision*, 19(1) :57–92, 1996.

- [13] H. Blum. A transformation for extracting new descriptors of shape. *Models for the perception of Speech and Visual Form*, pages 362–380, 1967.
- [14] F. L. Bookstein. Principal warps : Thin plate splines and the decomposition of deformations. *PAMI*, 11 :567–585, 1989.
- [15] O. Camara, O. Colliot, and I. Bloch. Computational modeling of thoracic and abdominal anatomy using spatial relationships for image segmentation. *Real-Time Imaging*, 10(4) :263–273, 2004.
- [16] S. Capell, S. Green, B. Curless, T. Duchamp, and Z. Popovi. Interactive skeleton-driven dynamic deformations. *SIGGRAPH*, 2002.
- [17] J. Carr, R. Beatson, J. Cherrie, T. Mitchell, W. Fright, B. McCallum, and T. Evans. Reconstruction and representation of 3D objects with radial basis functions. In Eugene Fiume, editor, *SIGGRAPH 2001, Computer Graphics Proceedings*, pages 67–76. ACM Press / ACM SIGGRAPH, 2001.
- [18] C. Chen, Y. Hung, and J. Cheng. Ransac-based darces : A new approach to fast automatic registration of partially overlapping range images. *Trans. PAMI*, 21(11), 1999.
- [19] Y. Chen, T.S. Huang, and Y. Rui. Optimal radial contour tracking by dynamic programming. *Proc. ICIP*, 2001.
- [20] Y. Chen and G. Medioni. Object modelling by registration of multiple range images. *Image and Vision Computing*, 10(3) :145–155, 1992.
- [21] D. Chetverikov, D. Stepanov, and P. Krsek. Robust euclidean alignment of 3-d point sets : the trimmed iterative closest point algorithm. *Image Vision Comput.*, 23(3) :299–309, 2005.
- [22] H. Chui and A. Rangarajan. A new point matching algorithm for non-rigid registration. *Computer Vision and Image Understanding*, 89 :114–141, 2003.
- [23] V. Chvatal. Linear programming. *W.H. Freeman and Company, New York*, 1983.
- [24] F. Conti, O. Khatib, and C. Baur. Interactive rendering of deformable objects based on a filling sphere modeling approach. *Robotics and Automation*, 3 :3716–3721, 2003.
- [25] Q. Delamarre and O. Faugeras. 3-d articulated models and multi-view tracking with silhouettes. *ICCV*, pages 716–721, 1999.
- [26] H. Delingette. Modélisation, déformation et reconnaissance d’objets tridimensionnels a l’aide de maillages simplexes. *Thèse de : Paris, ECP*, 1994.

- [27] J. Denzler and H. Niemann. Active rays : A new approach to contour tracking. *Int. Journal of Computing and Information Technology*, 4(1) :9–16, 1986.
- [28] J. Denzler and H. Niemann. Active-rays : Polar-transformed active contours for real-time contour tracking. *Real-Time Imaging*, 5(3) :203–13, 1999.
- [29] C. Dorai, J. Weng, and A. Jain. Optimal registration of object views using range data. *Trans. PAMI*, 19(10), 1997.
- [30] D. Dörig. Segmentation algorithms of the head of femur using a search ray object detection. *Travail de semestre EPFL*, 2006.
- [31] J. Duchon. Interpolation des fonctions de deux variables suivant le principe de la flexion des plaques minces. *RAIRO Analyse Numérique*, 10 :5–12, 1976.
- [32] O.D. Faugeras and M. Hébert. The representation, recognition, and locating of 3-d shapes from range data. *The International Journal of Robotics Research*, 5(3) :27–52, 1986.
- [33] J. Feldmar. Recalage rigide, non rigide et projectif d’images médicales tridimensionnelles. *Thèse de : Ecole Polytechnique*, 1995.
- [34] J. Feldmar and N. Ayache. Recalage rigide, affine et localement affine de surfaces gauches. *Rapport de recherche de l’INRIA*, 1994.
- [35] M. Fiebig, C. M. Straus, V. Sehgal, B. C. Renger, K. Doi, and K. R. Hoffmann. Automatic bone segmentation technique for ct angiographic studies. *J. Comput. Assist. Tomogr.*, 23(1) :155–161, 1999.
- [36] M. A. Fischler and R. C. Bolles. Random sample consensus : A paradigm for model fitting with applications to image analysis and automated cartography. *Comm. of the ACM*, 24 :381–395, 1981.
- [37] A. Fitzgibbon. Robust registration of 2-d and 3-d point sets. *Proc. British Machine Vision Conference, Manchester, UK*, 2 :411–420, 2001.
- [38] M. Fornefett, Karl Rohr, and Siegfried Stiehl. Elastic registration of medical images using radial basis functions with compact support. *Computer Vision and Pattern Recognition*, 1 :402–409, 1999.
- [39] J. H. Friedmann, J.L. Bentley, and R.A. Finkel. An algorithm for finding best matches in logarithmic expected time. *ACM Transaction on Mathematical Software*, 3(3) :209–226, 1977.
- [40] D. Gennery. Visual tracking of known three-dimensional objects. *Journal of Computer Vision*, 7(1) :243–270, 1992.

- [41] S. Gold and A. Rangarajan. New algorithms for 2-d and 3-d point matching : Pose estimation and correspondence. *Pattern Recognition*, 31(8) :1019–1031, 1998.
- [42] S. Granger. Une approche statistique multi-échelle au recalage rigide de surfaces : Application à l’implantologie dentaire. *Thèse de : INRIA - Sophia Antipolis*, 2003.
- [43] L. Greengard and V. Rokhlin. A fast algorithm for particle simulations. *J. Comput. Phys*, 73 :325–348, 1987.
- [44] G. Hamarneh. Active shape models with multi-resolution. *Matlab package*.
- [45] R. L. Harder and R. N. Desmarais. Interpolation using surface splines. *J. Aircraft*, 9(2) :189–191, 1972.
- [46] B. K. P. Horn. Closed-form solution of absolute orientation using unit quaternions. *Journal of Optical Society of America*, 4 :629–643, 1987.
- [47] B. K. P. Horn, H. M. Hilden, and S. Negahdaripour. Closed-form solution of absolute orientation using orthonormal matrices. *Journal of Optical Society of America*, 5 :127–135, 1988.
- [48] N. Houhou, V. Duay, A. Allal, and J. Thiran. Medical images registration with a hierarchical atlas. In *13th European Signal Processing Conference*, 2005.
- [49] S. Ilic. Implicit meshes : Unifying implicit and explicit surface representations for 3-d reconstruction and tracking. *Thèse de : EPFL*, 2005.
- [50] T. Mac Inerney and D. Terzopoulos. Deformable models in medical image analysis : A survey. *Medical Image Analysis*, 1(2) :91–108, 1996.
- [51] J. J. Jacq and Roux. Registration of non-segmented images using a genetic algorithm. In Springer, editor, *Computer Vision, Virtual Reality, Robotics in Medicine*, 1995.
- [52] M. Jaéhne. Automatische extraktion des herzens aus kontrastverstärkten schichtaufnahmen des thorax. *Travail de Master : Technische Universität Darmstadt, Fachbereich Mathematik*, 2005.
- [53] R. Jain, R. Karsuri, and B. G. Schunck. Machine vision. *New York : McGraw-Hill*, 1995.
- [54] G. Jiang and L. Quan. Detection of concentric circles for camera calibration. *Tenth IEEE International Conference on Computer Vision*, 1 :333–340, 2005.
- [55] A. Johnson and M. Hébert. Surface registration by matching oriented points. In *International Conference on Recent Advances in 3-D Digital Imaging and Modeling*, pages 121–128, 1997.

- [56] T. Jost. Fast geometric matching for shape registration. *Thèse de : Université de Neuchâtel*, 2002.
- [57] M. Kass, A. Witkin, and T. Terzopoulos. Snakes : Active contour models. *Int. J. Computer Vision*, 1(4) :321–331, 1998.
- [58] J. Koch. The laws of bone architecture. *American Journal of Anatomy*, 21 :177–201, 1917.
- [59] F. Lazarus. Courbes, cylindres et métamorphoses pour l'image de synthèse. *Thèse de : Université de Paris VII*, 1995.
- [60] F. Lazarus, S. Coquillart, and P. Jancène. Interactive axial deformations. *Modeling in Computer Graphics*, pages 241–254, 1993.
- [61] F. Lazarus, S. Coquillart, and P. Jancène. Axial deformations an intuitive deformation technique. *Computer Aided Design*, 26(8) :241–254, 1994.
- [62] T. Lindeberg. Scale-space theory in computer vision. *Kluwer Academic Publisher*, 1994.
- [63] Z. Q. Liu, H. L. Liew, J. G. Clement, C. David, and L. Thomas. Bone image segmentation. *IEEE Trans. Biomed. Eng.*, 46 :565–573, 1999.
- [64] C. Lorenz, T. M. Buzug, C. Fassnacht, and J. Weese. Automated detection and segmentation of lumbar vertebrae in ct and cta images based on a grey-value profile parser. In Elsevier, editor, *Computer Assisted Radiology and Surgery*, pages 209–214, 1997.
- [65] C. Lorenz and J. von Berg. Fast automated object detection by recursive casting of search rays. In Elsevier, editor, *Computer Assisted Radiology and Surgery*, pages 230–235, 2005.
- [66] C. Lorenz and J. von Berg. Towards a comprehensive geometric model of the heart. *Lecture Notes in Computer Science*, 3504 :102–112, 2005.
- [67] S. Luciano, O.R.P. Bellon, and K.L. Boyer. Precision range image registration using a robust surface interpenetration measure and enhanced genetic algorithms. *IEEE Trans. on PAMI*, 27(5) :762–776, 2005.
- [68] J. B. Antoine Maintz and Max A. Viergever. A survey of medical image registration. *Medical Image Analysis*, 2(1) :1–36, 1998.
- [69] G. Marti, C. Baur, and P.-Y. Zambelli. Optimal femoral head contour segmentation in ct images using dynamic programming. *Technol. Health Care*, 12(4) :315–322, 2004.
- [70] T. Masuda, K. Sakaue, and N. Yokoya. Registration and integration of multiple range images for 3-d model construction. In *Proc. CVPR*, 1996.

- [71] Ph. Merloz. Chirurgie orthopédique assistée par ordinateur : état actuel et perspectives d'avenir. *e-mémoires de l'Académie Nationale de Chirurgie*, 5(3) :17–21, 2006.
- [72] D. J. Michael and A. C. Nelson. Handx : A model-based system for automatic segmentation of bones from digital hand radiographs. *IEEE Trans. Med. Imag.*, 8(1) :64–69, 1998.
- [73] N. Mitra, N. Gelfand, H. Pottmann, and L. Guibas. Registration of point cloud data from a geometric optimization perspective. *Proceedings of the 2004 Eurographics/ACM SIGGRAPH symposium on Geometry processing*, 71 :22–31, 2004.
- [74] J. Montagnat. Modèles déformables pour la segmentation ou la modélisation médicale 3-d et 4-d. *Thèse de : Nice - Sophia Antipolis*, 1999.
- [75] A. W. Moore. Efficient memory-based learning for robot control. *Thèse de : Carnegie Mellon University*, 1991.
- [76] B. S. Morse. Thresholding. *Thèse de : Brigham Young University*, accessed 15 June 2003.
- [77] T. O'Donnell, T. E. Boult, X.-S. Fang, and A. Gupta. The extruded generalized cylinder : A deformable model for object recovery. *CVPR*, pages 174–181, 1994.
- [78] C. Peterson and B. Soederberg. A new method for mapping optimization problems onto neural networks. *International Journal of Neural Systems*, 1(1) :3–22, 1989.
- [79] D. L. Pham, C. Xu, and J. L. Prince. Current methods in medical image segmentation. *Annual Review of Biomedical Engineering*, 2 :315–337, 2000.
- [80] H. Pottmann, S. Leopoldseder, and M. Hofer. Registration without icp. *Comput. Vis. Image Underst.*, 95(1) :54–71, 2004.
- [81] A. Rangarajan, S. Gold, and E. Mjolsness. A novel optimizing network architecture with applications. *Neural Computation*, 8(5) :1041–1060, 1996.
- [82] B. M. Harr Romeny. (tutorial on) scale-space theory for multiscale geometric image analysis. *mpss.iop.org/1999/pdf/bart.pdf*, 1999.
- [83] J.-M. Rouet, J.J Jacq, and C. Roux. Recalage élastique 3-d de surfaces numériques par optimisation génétique. *Colloque sur le Traitement du Signal et des Images (GRETSI)*, pages 1395–1398, 1997.

- [84] D. Rueckert, L.I. Sonoda, C. Hayes, D.L.G Hill, M.O. Leach, and D.J. Hawkes. Nonrigid registration using free-form deformations : application to breast mr images. *IEEE Transactions on Medical Imaging*, 18(8) :712–721, 1999.
- [85] S. Rusinkiewicz and M. Levoy. Efficient variants of the ICP algorithm. In *Proc. 3rd Int. Conf. on 3-D Digital Imaging and Modeling (3DIM)*, pages 145–152, 2001.
- [86] T. B. Sebastian, H. Tek, J. J. Cristo, S.W.Wolfe, and B. B. Kimia. Segmentation of carpal bones from 3-d ct images using skeletally coupled deformable models. *MICCAI 98*, pages 1184–1194, 1998.
- [87] T. Sederberg and S. Parry. Free-form deformation of solid geometric models. In *SIGGRAPH '86 : Proceedings of the 13th annual conference on Computer graphics and interactive techniques*, pages 151–160, 1986.
- [88] J. Serra. Image analysis and mathematical morphology. *Academic Press, London*, 2, 1988.
- [89] D. Simon. Fast and accurate shape-based registration. *Thèse de : Carnegie Mellon University*, 1996.
- [90] R. Sinkhorn. A relationship between arbitrary positive matrices and doubly stochastic matrices. *Ann. Math. Statist.*, 35 :876–879, 1964.
- [91] C. Stewart, C.-L. Tsai, and B. Roysam. The dual-bootstrap iterative closest point algorithm with application to retinal image registration. *IEEE Transactions on Medical Imaging*, 22(11) :1379–1394, 2003.
- [92] C. Studholme, J. Little, G. Penny, D. Hill, and D. Hawkes. Automated multimodality registration using the full affine transformation : application to mr and ct guided skull base surgery. *Visualization and Biomedical Computing*, pages 601–606, 1996.
- [93] N. Sugano, T. Nishii, K. Nakahodo, T. Sasama, Y. Sato, S. Tamura, K.T. Sakai, K. Haraguchi, S. Nishihara, A. Ohzono, K. Yonenobu, and T. Ochi. Combined acetabular and femoral navigation for resurfacing total hip arthroplasty. *CARS 2000*, pages 226–231, 2000.
- [94] R. Szeliski and S. Lavallée. Matching 3-d anatomical surfaces with non-rigid deformations using octree-splines. *IEEE Workshop on Biomedical Image Analysis*, pages 144–153, 1994.
- [95] M. Teistler, J. Dormeier, M. Krosche, C. Braune, T. Pohlemann, and D. P. Pretschner. A software tool supporting ct-based classification of acetabular fractures. *CARS 2000*, pages 257–261, 2000.
- [96] J.-P. Thirion. New feature points based on geometric invariants for 3-d image registration. *Rapport technique INRIA*, 1910, 1993.

- [97] J.-P. Thirion. Non-rigid matching using demons. *Computer Vision and Pattern Recognition*, 1996.
- [98] E. H. Thompson. A method for the construction of orthogonal matrices. *Photogrammetry Record*, 3 :55–59, 1958.
- [99] P. A. van den Elsen, E. J. D. Pol, and M. A. Viergever. Medical image matching - a review with classification. *IEEE Engineering in medicine and biology*, 12(1) :26–39, 1993.
- [100] Z. Xie and G.E. Farin. Image registration using hierarchical b-splines. *IEEE Transactions on Visualization and Computer Graphics*, 10(1) :85–94, 2004.
- [101] P. Yan and K. W. Bowyer. Icp-based approaches for 3-d ear recognition. *SPIE Conference on Biometric Technology for Human Identification*, 2005.
- [102] P.J. Yim, J.J. Cebal, R. Mullick, H.B. Marcos, and P.L Choyke. Vessel surface reconstruction with a tubular deformable model. *IEEE Transactions on Medical Imaging*, 2001.
- [103] A. L. Yullie and J. J. Kosowsky. Statistical physics algorithms that converge. *Neural Computation*, 6 :341–356, 1994.
- [104] P.-Y. Zambelli, C. Bregand, S. Dewarrat, G. Marti, C. Baur, and P. Leyvraz. Planning and navigation solution in resurfacing hip surgery : A way to reduce the surgical approach. *Poster session, 3rd Annual Meeting of the International Society for Computer Assisted Orthopaedic Surgery*, 2003.
- [105] M. Zerroug and R. Nevatia. Quasi-invariant properties and 3-d shape recovery of non-straight, non constant generalized cylinders. *CVPR*, pages 96–103, 1993.
- [106] A. P. Zijdenbos and B. M. Dawan. Brain segmentation and white matter lesion detection in mr images. *Critical Reviews in Biomedical Engineering*, 6 :401–465, 1993.
- [107] R.A. Zoroofi, Y. Sato, T. Sasama, T. Nishii, N. Sugano, K. Yonenobu, H. Yoshikawa, T. Ochi, and S. Tamura. Automated segmentation of acetabulum and femoral head from 3-d ct images. *IEEE Transactions on Information Technology in Biomedicine*, 6 :329–343, 2003.